12-17-2019

Profiling Ubiquitin Chain Specificity and Topology Using Di-ubiquitin Probes

Han Zhou
Georgia State University

PROFILING UBIQUITIN CHAIN SPECIFICITY AND TOPOLOGY USING DI-UBIQUITIN PROBES

by

HAN ZHOU

Under the Direction of Jun Yin, PHD

ABSTRACT

Ubiquitin chains of specific linkages are assembled by ubiquitin conjugating enzymes (E2s) and ubiquitin ligases (E3s) to encode unique signals in the cell. Nevertheless, ubiquitin chains are often trimmed by deubiquitinases to reverse encoded signals. Given that there are altogether eight linkage types of polyubiquitin chains, how E2s, E3s and DUBs mediate ubiquitin chains in a linkage-specific manner is still an open question. In this thesis, an interdisciplinary approach including organic synthesis, genetic code expansion, protein engineering has been investigated to stress this intriguing question. To this end, linkage-specific diUB featuring a thiol group embedded near the isopeptide bond site was prepared in large scale and it was readily used for activity-based protein profiling to uncover more linkage-specific ubiquitin chain regulators by

converting it into a probe with an electrophilic trap, or for structure determination by conjugating it to corresponding chain-specific E2 or E3s. Taken together, these linkage-specific diUB probes should provide access for structural biologists to map dynamic, transient, weak protein-protein interactions in UB transfer, therefore to uncover detailed mechanism underlying chain specificities of E2, HECT E3 or DUB with direct evidences and to reconstitute full pictures of E2, E3 mediated ubiquitination on substrates and linkage-specific chain elongation.

INDEX WORDS: Ubiquitin chain synthesis, Linkage specificity, Genetic code expansion, Native chemical ligation, Activity-based protein profiling, Crystallography

PROFILING UBIQUITIN CHAIN SPECIFICITY AND TOPOLOGY USING DI-UBIQUITIN PROBES

by

HAN ZHOU

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2019

PROFILING UBIQUITIN CHAIN SPECIFICITY AND TOPOLOGY USING DI-UBIQUITIN PROBES

by

HAN ZHOU

Committee Chair: Jun Yin

Committee: Binghe Wang

Luo Ming

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2019

DEDICATION

This dissertation is dedicated to my parents, whose unconditional love has empowered me working day and night to pursue and complete this research; and to those, who I thought would stand by my side for the rest of my life.

ACKNOWLEDGEMENTS

I would like to express my largest gratitude to my advisor, Prof. Jun Yin, for his guidance and support during all these years. Dr. Yin has opened up a brand-new academic world to me, addressing problems, doing research, interpreting results: something I really enjoy doing for the rest of my life. He also encouraged me when I am frustrated, to try things out, to "leave no stone unturned"; that's something I learned from him and will benefit me to go further.

I would also like to thank our collaborators, Prof. Ashton Crop from Virginia Commonwealth University, Prof. Sonja Lorenz from University of Wurzburg, Prof. Nicolas Brown from University of North Carolina at Chapel Hill, Prof. Ning Zheng from University of Washington, Prof. Ming Luo from Georgia State University, Prof. Courtney Aldrich from University of Minnesota at twin cities. All of them have given me useful suggestions and provided valuable materials to my project.

Nevertheless, I would like to thank my past and current Yin lab members for your time working with me, especially people working on diUB project. I would thank everyone, including undergraduate students who devote their time and effort on this project. Without your help, I cannot finish this thesis.

TABLE OF CONTENTS

LIST OF SCHEMES

LIST OF FIGURES

1 INTRODUCTION

As a fundamental mechanism to modulate protein function in the cell, post-translational modification (PTM) regulates most if not all biological processes[1, 2]. The side chain of lysine with a primary amino group often serves a reactive handle for several important PTMs including methylation, acetylation and ubiquitination[1, 2]. Among them, ubiquitination is particularly intriguing in regards of structures and functions; unlike other PTMs such as phosphorylation, methylation, acetylation, in which the modifying groups are relatively simple and small, ubiquitination involves attachment of one or even more entire proteins, namely ubiquitin, covalently to a desired substrate, rendering it very challenging yet interesting to study.

Ubiquitin was first known to post-translationally target eukaryotic proteins for degradation in the late 1970s.[3-6] This discovery was recognized by the awarding of the 2004 Nobel Prize in Chemistry to Aaron Ciechanover, Avram Hershko, and Irwin Rose. Until the mid-90's, this PTM was believed to conduct primarily proteasomal signals; however, a large portion of non-proteasomal functions in, for example, endocytosis and lysosomal targeting, subcellular localization of proteins, autophagy, DNA repair, and kinase activation has been reported. Nevertheless, malfunction of the UB system plays causal roles in diseases such as cancer, inflammatory diseases, and neurodevelopmental and degenerative disorders.[7-11]

1.1 Ubiquitin and E1-E2-E3 cascade

Ubiquitin (UB), the 76-residue protein riding on a E1-E2-E3 enzymatic cascade, is a key note in the tunes of cell signaling of ubiquitination. In this process, the C-terminus of ubiquitin is covalently attached to a ε-amine of a lysine forming an isopeptide bond. (Figure 1-1).

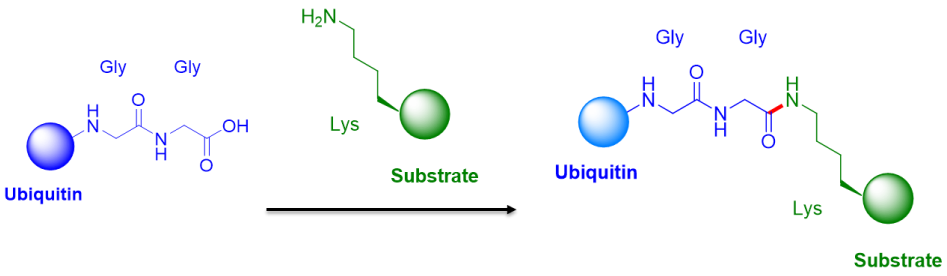


Figure 1-1 Isopeptide formation in substrate ubiquitination

In nature, amide bond forming is usually hard and requires additional activation of the carboxyl beforehand. Fortunately, a ubiquitin pathway involving three different sets of enzymes, known as E1-E2-E3 cascade, elegantly facilitates the conjugation between ubiquitin and the desired substrate. (Figure 1-2) [12-16] In this cascade, ubiquitin and ATP will firstly bind to a 110 kD enzyme, namely ubiquitin activating enzyme (E1), in an adjacent manner so that a condensation reaction will happen yielding UB-AMP and pyrophosphate. In this first half reaction E1 catalyzed, the C-terminus of ubiquitin is activated through an energy consuming process and the newly-formed AMP ester is prone to nucleophilic attack by a catalytic cysteine on E1 forming a thiol ester bond often known as UB~E1 where as "~" designates a thioester bond. On the second half of the reaction, E1 will transfer this charged ubiquitin onto ubiquitin conjugating enzyme (E2) which is bound to E1. Upon Ub~E2, a third component of enzymes, ubiquitin ligases or E3s with their unique substrate binding domain will recruit corresponding E2s charged with ubiquitin and carry out the transfer to a specific lysine of the target protein to achieve ubiquitination with an isopeptide bond.

Ubiquitin ligases (E3) play a very crucial role in precisely delivering ubiquitin to pre-defined targets and fall into different classes based upon their working modes to engage the UB~E2 conjugate. HECT-type E3s with the name coming from Homologous to the E6-AP Carboxyl Terminus bear a similar catalytic cysteine like E2s that can form Ub~E3 thus indirectly

ubiquitinate substrates, whereas RING-type E3s (Really-Interesting-New-Genes) often serve as a platform and recruit corresponding E2s loaded with UB and substrate at the same time to facilitate a direct transfer.

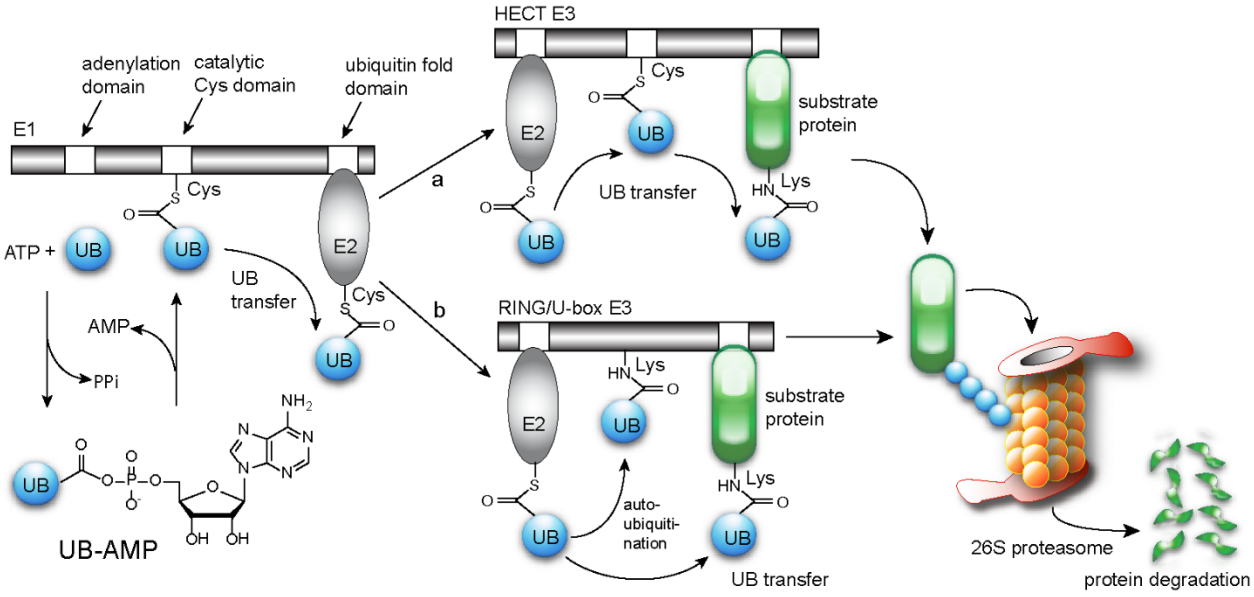


Figure 1-2 E1-E2-E3 cascade in ubiquitination pathway

E1-E2-E3 cascade enables regulations from different steps of ubiquitin transfer and ensures the accuracy of modifications on desired substrates. There are 2 E1s, 40 E2s, and more than 600 E3s encoded in the human genome that assemble a complex network of UB transfer. More detailed E1, E2, HECT or RING type E3 structures and functions are discussed in later chapters. (Chapter 5.1, Chapter 1.3 and Chapter 4.3).

1.2 Ubiquitin chains: topologies and functions

Within E1-E2-E3 cascade for ubiquitin transfer, the assignment of each E1, E2, E3 to facilitate ubiquitination of a target substrate at a desired position of a lysine residue is very

complicated, resulting in the complexity of ubiquitination from one dimension.[17] Nevertheless, aiming at one particular substrate, single UB can be attached to one specific position of lysine residue (mono-ubiquitination) or several different positions (multiple-ubiquitination); several UBs can modify the same lysine residue in a manner of forming homotypic or heterotypic UB chains (poly-ubiquitination).[18-20] In addition to these, UB on the substrates can also be post-translationally modified with acetylation[21] or phosphorylation[22-24]. Adding together, ubiquitination patterns provides another dimension of complexity in ubiquitination, also known as ubiquitin code[19, 20]. Deciphering the ubiquitin code is extremely difficult yet very interesting due to the complexity of ubiquitin code and availability of homogeneous ubiquitin chains.

1.2.1 Ubiquitin chains with different linkage types and their structural features

Based upon the sequence and structure of ubiquitin (Figure 3), there are altogether seven lysine, namely K6, K11, K27, K29, K33, K48, K63 and one methionine at N-terminus (M1) that can potentially serve as ubiquitin chain acceptors. Previous proteomics studies suggested that all possible linkage types co-exist in cells with very different abundances (K48 51.7%, K63 37.8%, K29 7.8%, K11 2.2%, K6 0.5% M1 $<$ 0.5%, K27 $<$ 0.5% K33 $<$ 0.5% respectively). [25-29] Consequently, M1, K6, K11, K27, K29 and K33 linkages are usually referred as “atypical” ubiquitin chain types.

The variety of cellular processes initiated and regulated by ubiquitination has been indicated in part by the structural diversity of different ubiquitin linkage types. Although the ubiquitin chains are overall dynamic, structural analysis including X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, has revealed topologies of ubiquitin dimers. [30-36] (Figure 1-4) The conformations of ubiquitin dimers fall into two categories: K6, K11, K29, K33

and K48 type chains tend to form a “closed” topology with intact intramolecular interfaces while K63 and M1 type chains adopt “open” architectures without much interaction region except for the linkage point.[37] However, the structure of di-Ubiquitin with K27 linage has not been resolved yet.

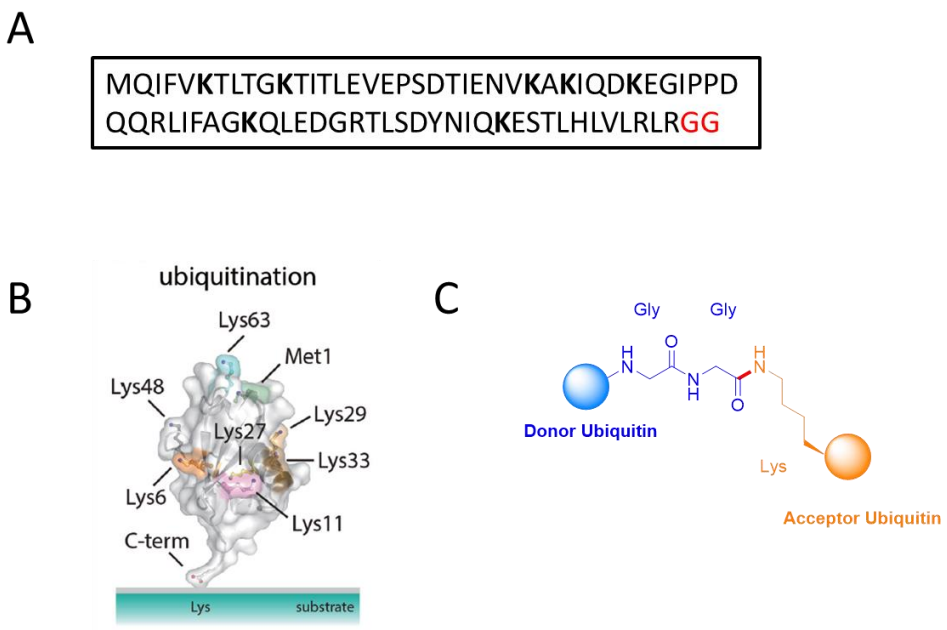


Figure 1-3 Ubiquitin chain formation

(A) Ubiquitin protein sequence, lysines that can be conjugated with another molecule of ubiquitin are bolded and a conserved diGly motif at C-terminus of UB is colored in red. (B) Structure of ubiquitin highlighting the eight sites of ubiquitination[18]. (C) Isopeptide formation between donor and acceptor UBs.

Ubiquitin features several surface patches, for example, hydrophobic regions around I36 and I44, that are crucial elements recognized by ubiquitin chain regulators. These patches are positioned uniquely due to the topologies of linkage-specific polyubiquitin chains with distinct projection patterns of hydrophobic areas (Figure 1-4), which serve as structural bases for signal specificity of various cellular event summarized in the next section.

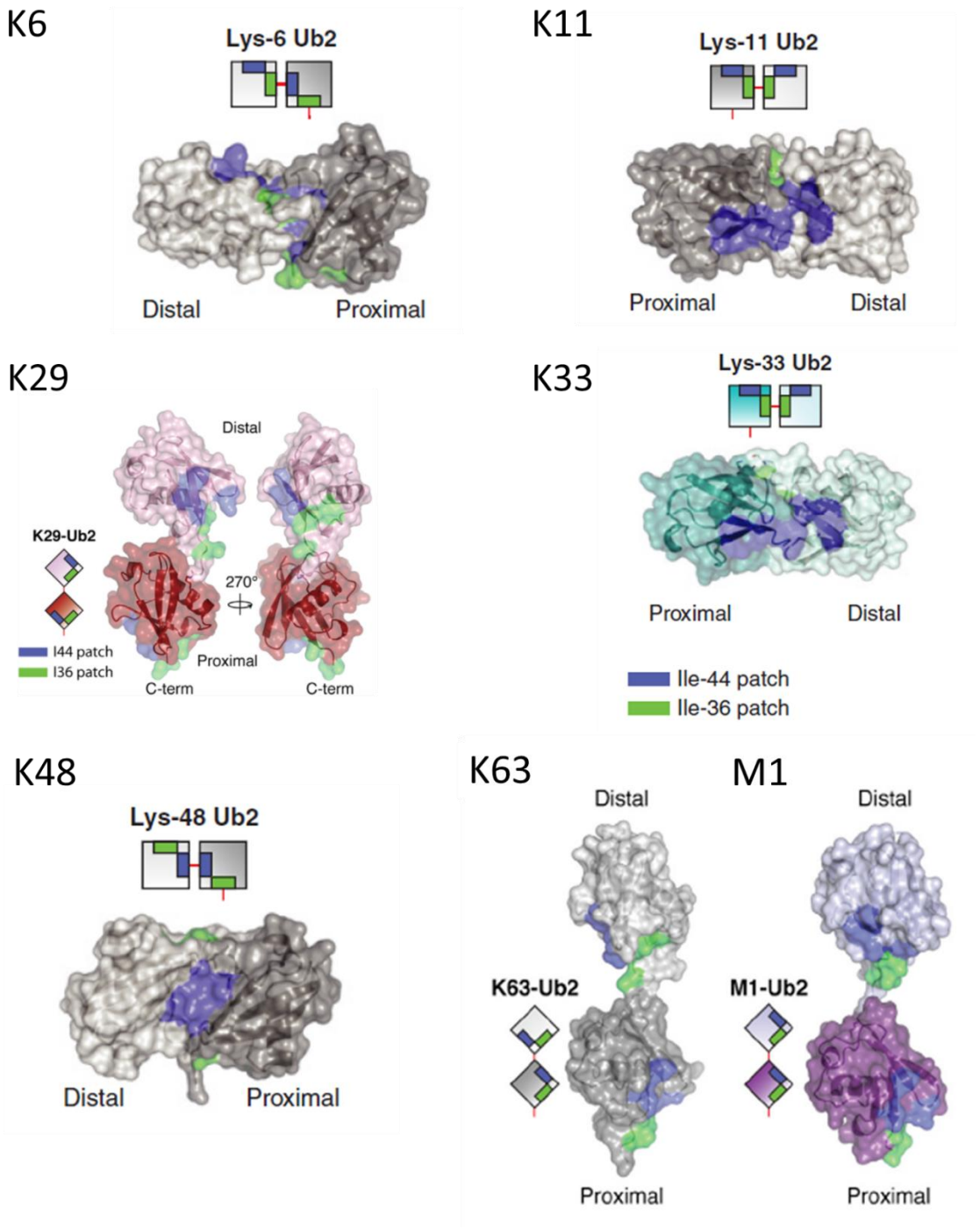


Figure 1-4 Ubiquitin chain topologies

PDB number: K6 2XK5, K11 1AAR, K29 4S22, K33 2XK5, K48 3N0B, K63 2JF5, M1 2W9N

1.2.2 Biological significances of linkage-specific ubiquitin chains

Ubiquitination study had been limited to K48-and K63-linked polymers for many years until the late 2000s when innovative technologies in molecular biology and mass spectrometry reveal the biological significances of "atypical" ubiquitin chains.[18, 19] (Figure 1-5)

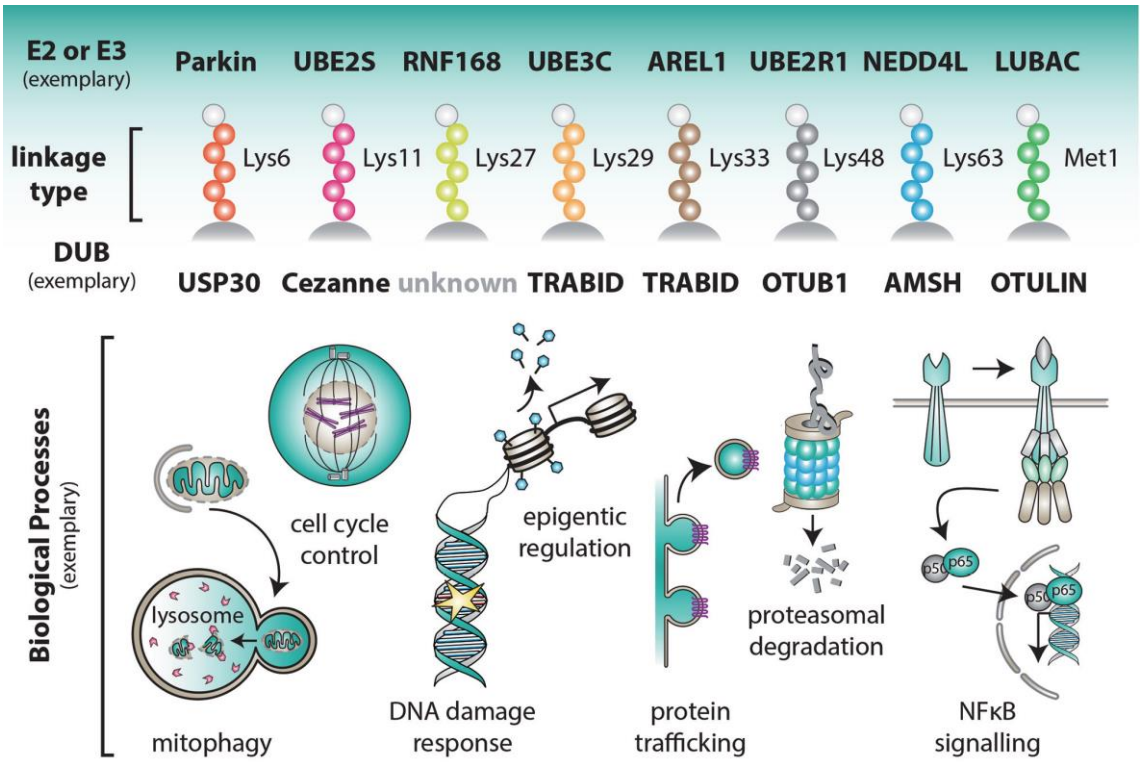


Figure 1-5 Physiological roles associated with individual chain types.

A small selection of E2 or E3 enzymes that assemble and DUBs that disassemble ubiquitin chains with linkage preferences is indicated. Below, cartoons illustrate some of the biological processes that particular linkage types have been linked with as discussed in the section.

1.2.2.1 M1 linkages

The role of linear ubiquitin chains (M1) is heavily implicated in regulation of innate immune signaling associate with NF-κB nuclear translocation and activation. (Figure 1-6) Upon cytokine receptors and toll-like receptors (TLRs) activation, downstream kinases and E3 ligases

are further activated to phosphorylate or ubiquitinate κB kinase (IKK) complex comprised of NEMO, IKKα and IKKβ subunits; activated IKK complex will phosphorylate and then cause degradation of NF-κB inhibitor, or IκB with ubiquitination by an E3 ligase SCF(β-TrCP), resulting in the release and translocation of NF-κB to induce transcription for inflammatory and immune response. [38, 39] In the process, ubiquitination of IKK complex by TRAF6 forming K63-type poly-ubiquitin chains on IKKα and IKKβ and more importantly ubiquitination by LUBAC, a multi-subunit RBR or Ring-Between-Ring E3 ligase[40] comprising HOIP, HOIL-1L and SHARPIN, forming M1-type poly-ubiquitin chains on NEMO are critical for IKK complex full activation[41, 42] in a way that M1-linked chains will bind to a linear-chain specific ubiquitin binding domain (UBD) on NEMO to recruit TAK1 and then to further activate IKKα and IKKβ.[43]

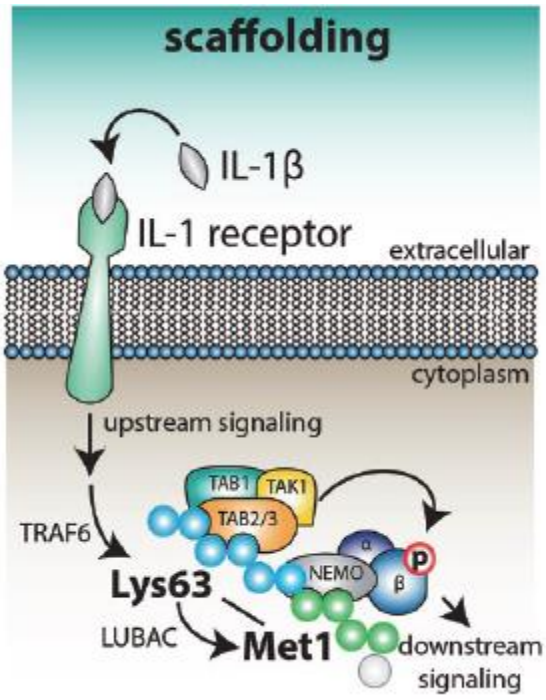


Figure 1-6 M1 and K63 ubiquitin chains in NF-κB activation[18]

K63-polyubiquitination by TRAF6 and M1-polyubiquitination by LUBAC (HOIP) are crucial components for kinase inhibitory kinases TAK, IKK activation to release and translocate NF- κB.

1.2.2.2 K11 linkages

So far K11-linked ubiquitin chains are mainly associated with cell cycle regulation. During mitosis, the abundance of K11 linkages increases dramatically when metazoan anaphase-promoting complex, or APC/C is active.[35] It was believed that APC/C initiates substrate ubiquitination and chain formation together with an E2 enzyme UBE2C where mainly K48 or K63 chains are assembled.[44] Subsequently, another E2 enzyme UBE2S further elongates the UB chains of K11 linkage.[45] The ubiquitinated substrates by UBE2C, UBE2S and APC/C are targeted for proteasomal degradation, which results in mitotic exit. A detailed mechanism in the event revealed by recent structural studies is discussed in later chapter. (Chapter 4.1 and Chapter 4.3)

1.2.2.3 K48 linkages

The K48 linkages of ubiquitin chains are the most abundant one as previously described. And the major role is targeting proteins for proteasomal degradation.[20] Traditionally, tetra-ubiquitination of K48 linkage is regarded as a minimum requirement to be efficiently targeted to 26S proteasome.[46] However, the dogma was challenged by recent studies showing that multiple-ubiquitination with various short chains can promote proteasomal degradation. [47] Nevertheless, some non-degradative purposes of K48 linkages are reported where protective UBDs seem to play a role. [48, 49]

1.2.2.4 K63 linkages

As discussed in M1 linkages section, the K63 type ubiquitin chains are found significantly in immune response of NF-κB pathways where TRAF family E3 ligases catalyze K63-polyubiquitination for “upstream” kinase TAK1 and “downstream” kinase IKK activation. (Figure 1-6) Notably, there are UBD on TAK1 for chains of K63- linkage, called TAB2 and TAB3.[50, 51]

Along with similar K11 poly-ubiquitin chain recognizing UBD on NEMO (UBAN domain), these features serve as structural basis for ubiquitin chain formation and NF- κB activation.[52]

The role of K63 chains is also heavily implicated for DNA double-strand break (DSB) repair. DSBs are detected by the Mre11-Rad50-Nbs1 (MRN) complex, which recruits and activates the protein kinase ATM at the site of DNA damage (Figure 1-7) [53].ATM phosphorylate H2AX and MDC1; phosphorylated MDC1 is recognized by the FHA domain of RNF8, a RING domain E3 that functions together with Ubc13 to catalyze K63 polyubiquitination of target proteins including H2AX and H2A. The K63 polyubiquitin chains bind to the RAP80 UIM domain, thereby recruiting the BRCA1 complex that includes Abraxas, BARD1, and BRCC36.[54]

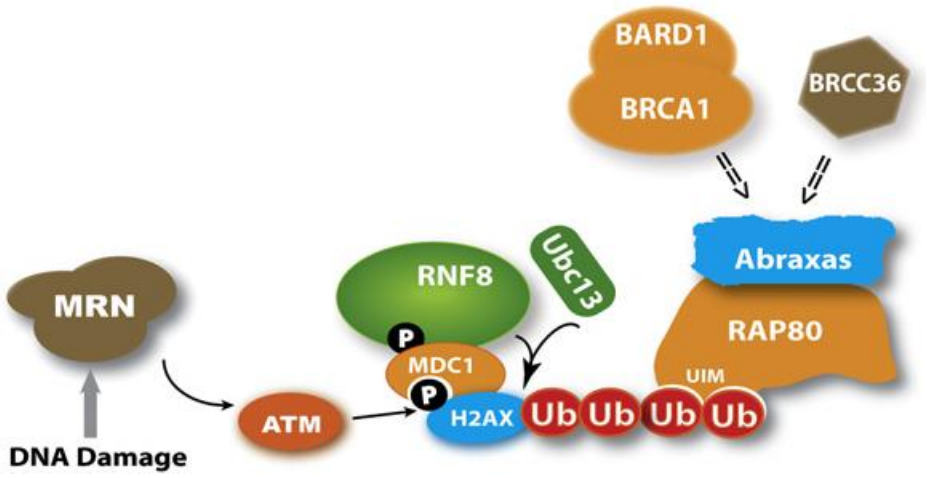


Figure 1-7 Role of K63 polyubiquitination in DNA damage repair[54]

1.2.2.5 K6, K27, K29, K33 chains

The significances of other "atypical" ubiquitin chains are uncovered gradually by recent studies with still a lot more to learn. DNA repair associated heterodimeric E3 ligase BRCA1-BARD1 aforementioned has been reported to assembly K6 poly-ubiquitin chains with no protein stability implication.[55, 56] Another evidence of K6 poly-ubiquitination was found in mitochondrial

outer membrane (MOM) proteins upon depolarization of the organelle, most likely assembled by Parkin E3 ligase. [57]

K27 chains are reported[58] to be formed on histone H2A proteins in the process of DNA damage response; crucial mediators in this event, including 53BP1, Rap80, RNF168 and RNF169 were known to recognize K27 polyubiquitination H2A.

Recent studies show that Smad ubiquitination regulatory factor 1 (Smurf1) can modify Axin with non-degradable K29-linked ubiquitin polymers. Axin is a scaffold protein in Wnt/β-catenin signaling pathway; Axin bearing K29 ubiquitin chains perturbs the ligand bindings for Wnt co-receptors thus inhibits this process.[59]

K33- ubiquitin chains are associated with negative regulation of both T-cell antigen receptor (TCR) [60] and AMP-activated protein kinase (AMPK)-related protein kinases[61]. Another paper suggested K33 linkages are important for post-Golgi protein trafficking.[62]

Taken together, polyubiquitin chains mediate various signal pathways and cellular events beyond proteasomal degradation. New functions of canonical ubiquitin chains are still under heavy investigation, while the roles of atypical chains remain largely unclear.

1.3 Factors that associate with ubiquitin chains

The dynamic features of almost all post-translational modifications facilitate sophisticated yet flexible regulations of biological processes to maintain steady states for all cells.[1, 2] In the case of ubiquitination, while distinct architectures of polyubiquitin chains are the structural basis of various cellular functions, recognition and regulations of the lengths and linkage types of ubiquitin chains are critical when building-up, removing and deciphering ubiquitin codes so that the ubiquitination signals can be precise and reversible. The nature of these chains is determined

largely by specific E2s and E3s involved, which together with E1, “write” the UB signal on proteins. Ubiquitination is a reversible process, deubiquitinating enzymes (DUBs; deubiquitinases), can either completely reverse or alternatively “edit” substrate-bound chains by cleaving linkages with various degrees of specificity. The outcome of protein ubiquitination is largely a function of the UB chain linkage in the context of the substrate and its recognition by the many proteins containing UB-binding domains (UBD), which translate or “read” the UB code on a particular protein.

1.3.1 E2 and E3 as ubiquitin code “writers”

The E1-E2-E3 cascade (Figure 1-2) “writes” the ubiquitin code by building up various ubiquitin polymers on substrates. In this process, E2 and E3 are important players to determine chain linkage types as E1 mainly focuses on activating ubiquitin and transferring loaded ubiquitin to E2.

1.3.1.1 Ubiquitin conjugating enzymes (E2s)

E2s are at the crossroads regarding the linkage specificity of ubiquitination. Members of this ~ 40-member family have a conserved UBC fold that is critical to interactions with both E1 and the various families of E3s.[16, 63] This fold includes a catalytic Cys through which UB is bound in a thioester linkage (E2~UB) after transesterification from E1. The N-terminal helix of UBC is a key element for interacting with the UB-fold domain (UFD) of the E1 enzyme [64]. Besides the UBC, various E2s may have extended N or C termini that affect their pairing with various E1 and E3 enzymes.

E2s are the decision makers for ubiquitin chain initiation on a lysine of a substrate or elongation on a lysine of an acceptor ubiquitin. Most of the time, different E2s are assigned for

initiation or elongation. In the case of UBE2C, UBE2S and APC/C described in chapter 1.3.1.1, UBE2C ubiquitinates APC/C substrates with short K48 ubiquitin chains before UBE2S further elongates K11 chains specifically. [65] A similar example is BRCA1-BARD1 ubiquitination mentioned earlier where UBE2W or UBE2E2 initiates ubiquitination and heterodimeric UBE2N-UBE2V1 or UBE2K specifically promote chain elongation[66]. Nevertheless, aforementioned TRAF6 ubiquitination follows the same pattern with UBE2D to initiate and UBE2N–UBE2V1 to elongate Lys63-linked ubiquitin chains during NF-κB activation.[67] An exception in this case is yeast E2 Cdc34, which initiates and elongates exclusive K48 polyubiquitin chains by itself when associates with E3 SCF (Figure 1-8)[68].

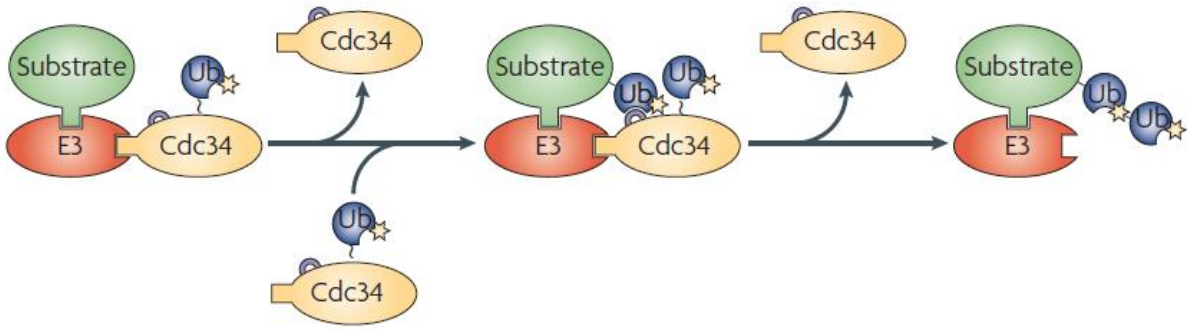


Figure 1-8 Cdc34 catalyzed ubiquitin chain initiation and elongation

Interestingly, E2s that initiate ubiquitination, especially UBE2D family members, lack the selectivity of lysine residues on substrate. However, UBE2S elongates exclusive K11-linked chains, UBE2N-UBE2V1 elongates K63 chains and UBE2K elongates K48 chain respectively. While the molecular basis for K11 linkage specificity of UBE2S remains unclear, it has been noticed that the efficient elongation of K48-specific the ubiquitin chain requires an interaction between an acidic loop located near the active site in Cdc34 and the substrate loaded ubiquitin to orient the attacking ε-amino group of Lys48. [69, 70] In addition, the crystal structure of yeast

homologues UBE2N-UBE2V1, Ubc13-Mms2 bound with both donor and acceptor ubiquitin provided direct evidence on the selectivity of lysine that can be conjugated with donor ubiquitin. (Figure 1-9) [71] In Ubc13-Mms2, the donor ubiquitin is covalently linked to the catalytic Cysteine in Ubc13, whereas the acceptor ubiquitin is noncovalently bound to the back of Mms2. The interaction with Mms2 positions the acceptor ubiquitin relative to Ubc13 so that Lys63, but no other Lys residue, can attack the thioester bond between the donor ubiquitin and Ubc13.

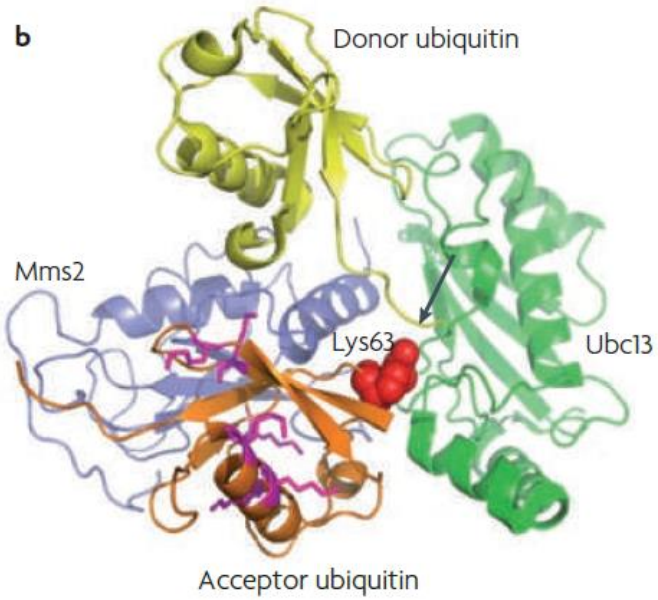


Figure 1-9 Crystal structure of Ubc13~donor Ub-Mms2 (PDB 2GMI)

1.3.1.2 HECT-type E3s

Another important "writer" of various ubiquitin chains is HECT E3s and RBR E3s with a catalytic cysteine that can form a thioester bond with donor ubiquitin. In humans, this category includes ~28 HECT E3s and 12 RING-Between-RING (RBR) E3s[14, 72]. Previous LC/MSMS data suggested most HECT E3s have preference of linkage types during ubiquitination (Figure 1-10) [34], somehow direct molecular evidence is still missing. Some recent studies addressing this question are discussed in the later chapter. (Chapter 4.2)

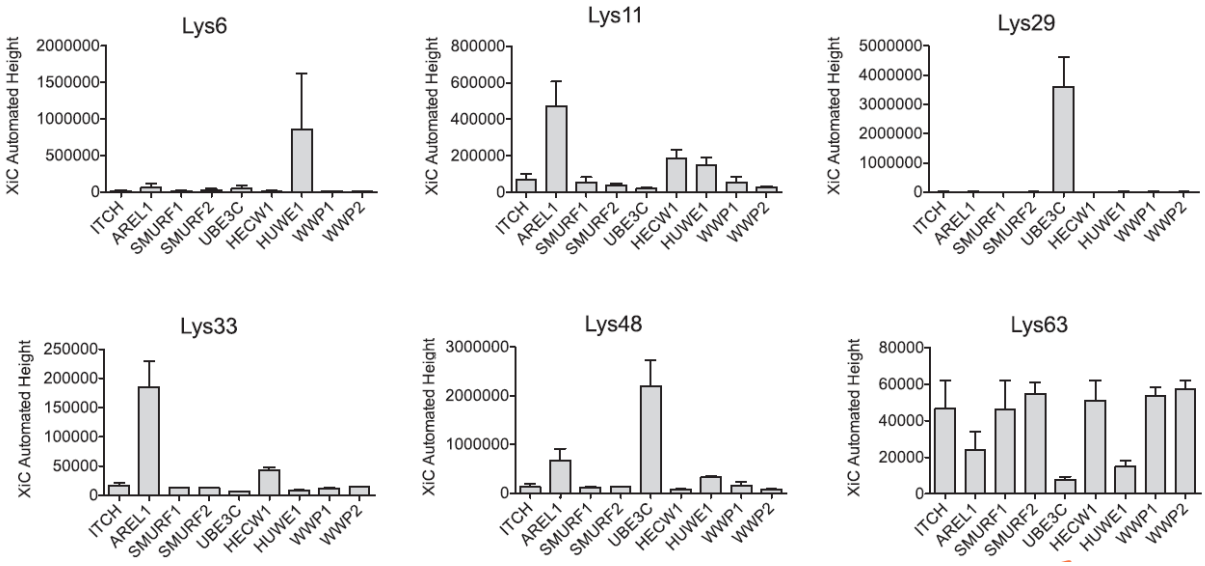


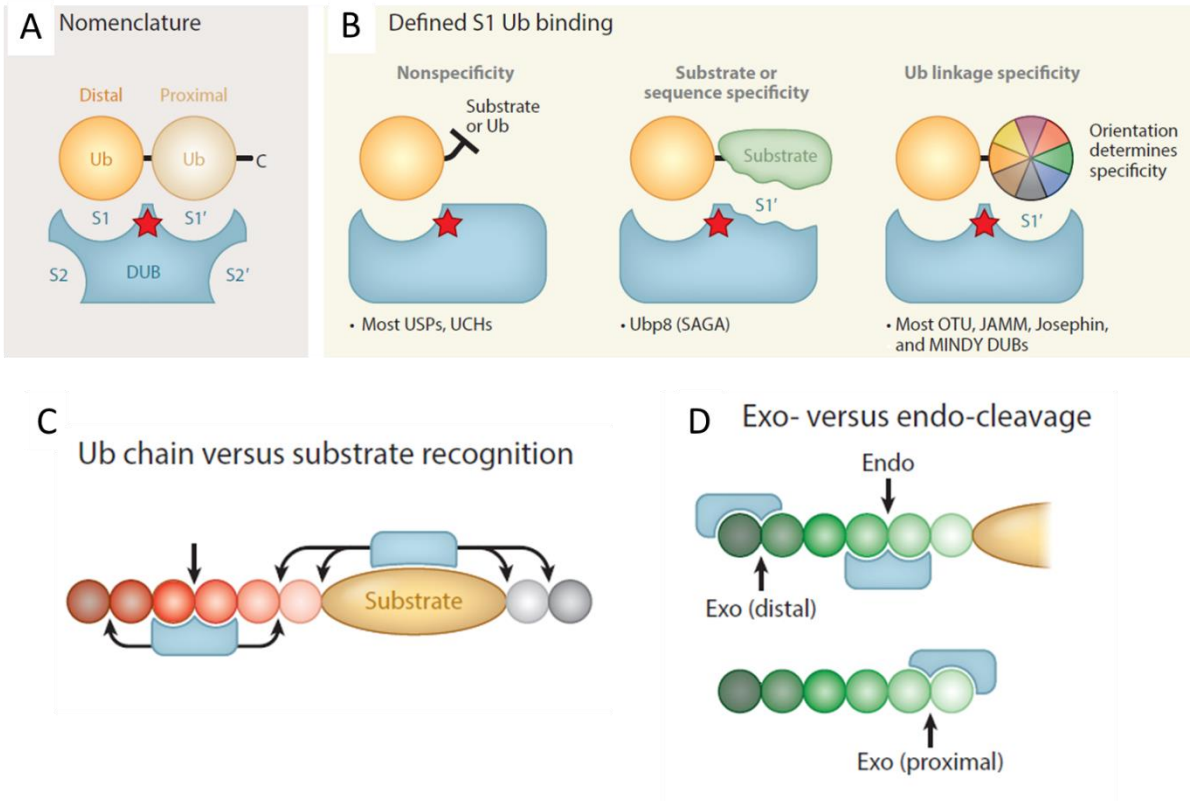
Figure 1-10 LC-MSMS data suggested specificity of polyubiquitin chain formed by HECT[34]

1.3.2 Deubiquitinating enzymes (DUBs) as ubiquitin code "erasers"

Ubiquitin chains are constantly trimmed or removed by DUBs, which account for more than 100 genes in the human genome. [73, 74]. A majority of DUBs are Cysteine proteases including UB-specific proteases (USPs, 54 members in humans), ovarian tumor proteases (OTUs, 16 members), UB C-terminal hydrolases (UCHs, 4 members), Josephin DUBs, and as of late, the motif interacting with ubiquitin (MIU)-containing novel DUB family (MINDYs, 4 members), except for a sixth family that is Zn-dependent metalloproteases of the JAMM family (6 members). [75, 76] The catalytic domains of DUBs bear a shared primary recognition site, the S1 site, that can bind with ubiquitin, which is the most important feature for DUB activity of editing ubiquitin chains.[77]

However, how DUBs select which modifier to target, how DUBs approach their substrates or ubiquitin chains are largely depend on other binding sites, such as S1', S2, S2'. (Figure 1-11A and B) These features result in different cognition and cleavage mechanism of DUBs, thereby explain different preferences of DUB function. Some of DUB remove the endo ubiquitin or

ubiquitin chains regardless of their length and linkages, regenerating unmodified substrates. (Figure 1-11C) With few exception, USP superfamily of DUBs falls into this category by taking



off the substrate-bound proximal ubiquitin without much selectivity of ubiquitin linkage types. [78]

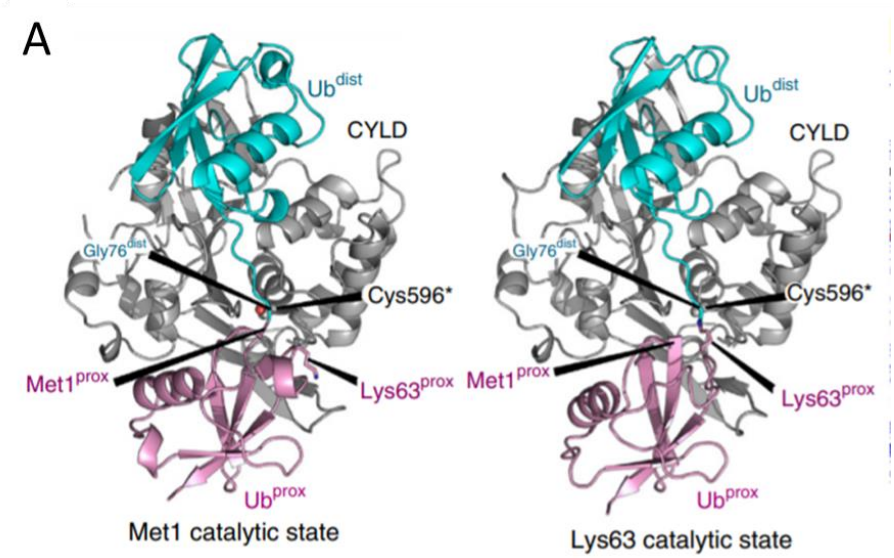
Figure 1-11 Principles of DUB substrate recognition and linkage specificity
(A) Basic nomenclature for DUB ubiquitin-binding sites and a model diubiquitin substrate. (B) Defined S1 sites can sufficient to recruitubiquitinated substrates while DUBs without S1' site are non-specific (USP and UCH family members). S1' sites interacting with substrate proteins or proximal ubiquitin are the structural basis for substrate or ubiquitin chain specificity. (OUT, JAMM, MINDY family members) (C) The binding features of DUB determine the cleave sites and chain specificity. (D) Arrangement and nature of UB-binding sites in DUBs suggest the cleavage modes: whether polyubiquitin is cleaved from the distal or the proximal end (exo-cleavage) or within a chain (endo-cleavage).

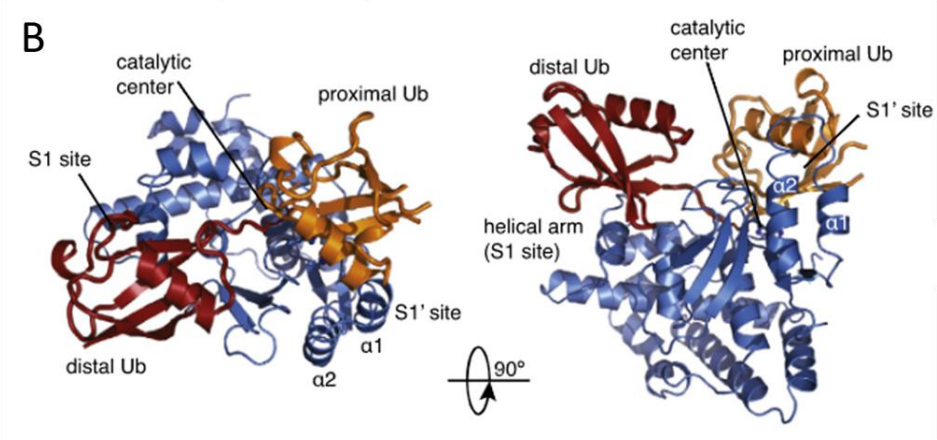
A completely different types of DUBs recognize and cleave predefined ubiquitin chains without much concern of the substrate (Figure 1-11D) whereas some DUBs of this category prefer removal the distal ubiquitin one at a time, known as "exo"-cleavage while some DUBs, especially members in OTUB superfamily, cleave the "endo" ubiquitin with a higher chain removal efficiency[75]. Commonly, these DUBs feature an additional proximal ubiquitin binding site, S1'

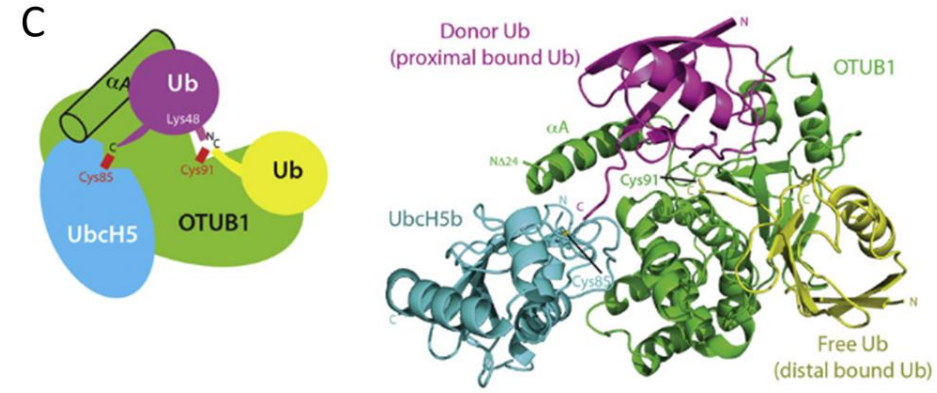
site, that can further orient the ubiquitin chain and facilitate the cleavage of an exposed C-terminus of donor ubiquitin in a linkage-specific fashion.

The degree of linkage specificity varies between DUB families; MINDY DUBs are K48 specific whereby most JAMM metalloproteases cleave K63 chains exclusively. As describes earlier, UCH and Josephin family members of DUB are promiscuous ubiquitin chain editors without a defined S1' site; therefore, they can barely cleave diubiquitin effectively. [79] Notably, USPs without the same S1' site are regarded as non-discriminable ubiquitin code "erasers" except for CYCD (cylindromatosis-associateed DUB)[80] and mitochondrial DUB USP30.[81] CYCD is a tumor-suppresser protein implicated in NF-κB pathway (Figure 1-6) where it selectively cleave K63- and M1- polyubiquitin chains assembly by TRAF5 and LUBAC.[82] Consequently, CYCD deactivate NF-κB to negatively regulate NF-κB pathway and prevent tumor cell survival or oncogenesis. A recent structural studies of CYCD bound of K63 or M1- di ubiquitin validates the role of S1 and S1' sites in ubiquitin chain recognition and specificity in ubiquitin chains. (Figure 1-12A)[83]

The most intriguing DUB class regarding their specificity towards various ubiquitin chains is OTU family.[84] OTUBLIN, an enzyme that equally crucial in NF-κB response, binds and cleaves M1-linked ubiquitin chains exclusively.[85] In addition, OTUB1 is one of the most abundant DUB in cells with a remarkable specificity for K48 linked polyubiquitin chains[86] whereas another member of OTU family, Cezanne, is the only know DUB that target and cleave K11 chains with high specificity[87]. Like CYCD, the crystal structures of DUBs in complex with their corresponding di-ubiquitins revealed direct evidence of their chain specificity. (Figure 1-12B, C, D)


A
Ub^dist
CYLD
Gly76^dist
Cys596*
Met1^prox
Lys63^prox
Ub^prox
Met1 catalytic state
Ub^dist
CYLD
Gly76^dist
Cys596*
Met1^prox
Lys63^prox
Ub^prox
Lys63 catalytic state


B
catalytic center
proximal Ub
S1 site
S1' site
α1
α2
distal Ub
90°
catalytic center
proximal Ub
distal Ub
S1' site
α2
α1
helical arm (S1 site)


C
αA
Ub
Lys48
Cys85
Cys91
Ub
UbcH5
OTUB1
Donor Ub (proximal bound Ub)
OTUB1
αA
NΔ24
Cys91
UbcH5b
Cys85
Free Ub (distal bound Ub)

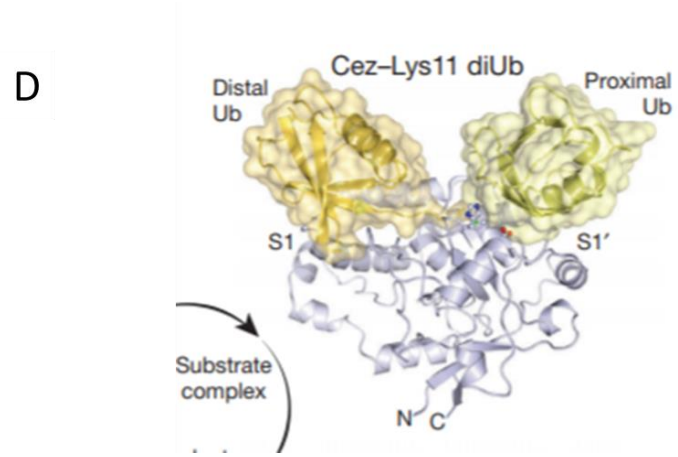


Figure 1-12 Crystal structure of DUBs bound preferred di-Ub

(A) CYCD bound with K63- and M1- diUB respectively (PDB 3WXE, 3WXG)
(B) OTULIN bound with M1 diUB (PDB 2ZFY)
(C) OTUB1 bound with K48- diUB (PDB 4DDI)
(D) Cezanne bound with K11- diUB (PDB 1UBQ)

1.3.3 UB-binding domains (UBDs) as ubiquitin code "readers"

The various topologies of ubiquitin chains and ubiquitinated substrates provide the molecular basis of ubiquitin codes that to be read so that unique modification will be interpreted into specific outputs. This is accomplished usually through binding to over 20 specific Ubiquitin-binding domains, or UBDs. [88, 89] Depending on the recognition unit on UBDs, they mainly divide into two superfamilies. The first group, including Ubiquitin-interacting motifs (UIM), Ubiquitin-associated domains (UBA) and Coupling ubiquitin to ER degradation domains (CUE), are made from one to three α helixes. The linear polyubiquitin binding motif, UBAN, that described in Chapter 1.2.2.1, also belongs to this group.[43] Another superfamily of UBD utilizes zinc fingers that recognition specific ubiquitin chains and these are comprised of NZF (novel zinc finger), ZnF_UBP (UBP-type zinc finger), ZnF_A20 and UBZ. These readers of signals can recognize various structural features on UB such as the classic hydrophobic patch surrounding Ile44, the TEK box (Thr9-Glu10-Lys11), the loop region of Glu51-Lys63 and the C-terminal Gly of UB [89]

However, most UBDs bind to ubiquitin with low affinity, with the binding constants ranging from several micromolar to several millimolar, which suggests that additional interactions or components are important in the specificity recognition.[53]

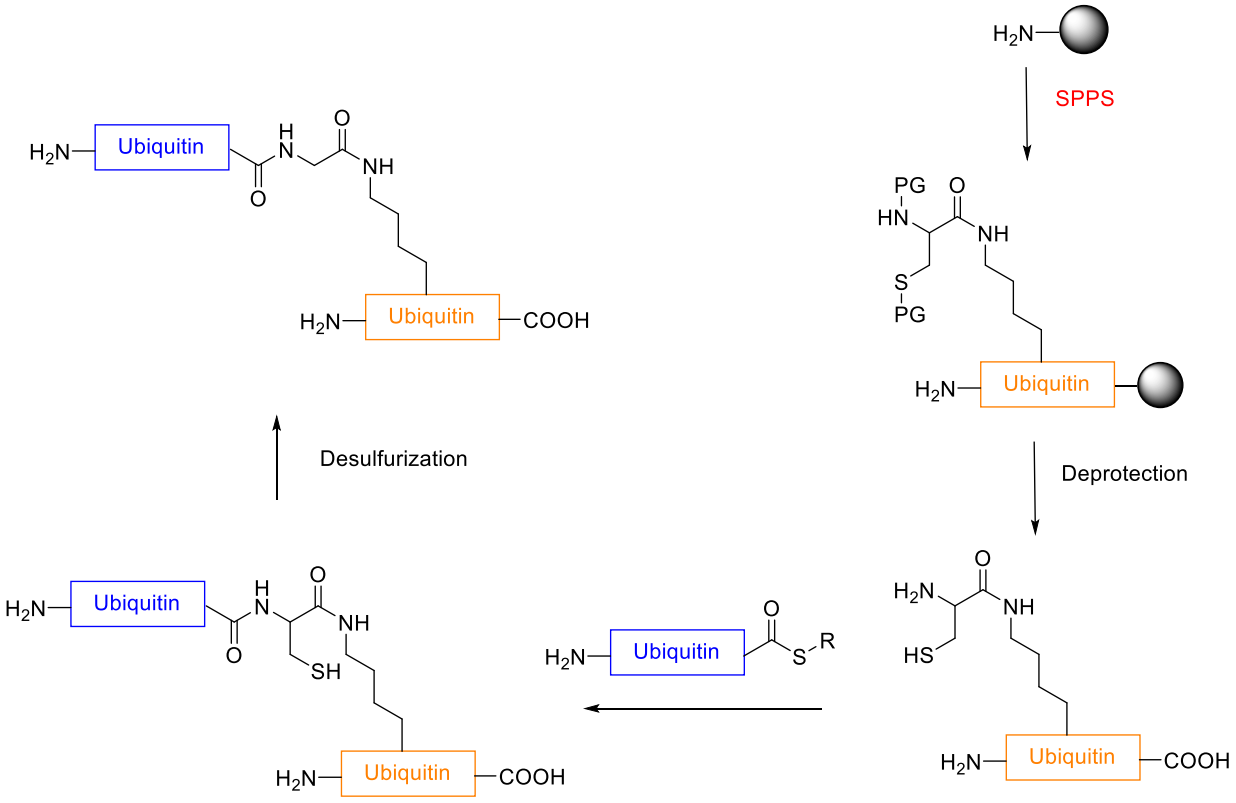
1.4 Bioorganic and biochemical approaches to understand ubiquitin code

The complexity of ubiquitin codes provides sophisticated manipulation and tight regulation of ubiquitin transfer within E1-E2-E3 cascade and specific ubiquitin chain formation, edition and recognition, which in part explains the significance of ubiquitination as a ubiquitous post-translational modification and errancy in ubiquitination pathway always results in serious diseases.[90-92] However, these different layers of complexity rise the difficulty in understand the mechanisms in ubiquitination pathway and its regulation. Fortunately, to decipher the ubiquitin code, previous biochemist and organic chemist had utilize various elegant approaches to accomplish at least some part of this task.

1.4.1 Ubiquitin chain synthesis

1.4.1.1 Synthetic ubiquitin chains with solid-phase peptide synthesis

The availability of ubiquitin chains with predefined length and linkage types is fundamental in understanding specific signals they convey. Unfortunately, utilizing linkage-specific E2s, such as E2-25K, which is known for building up K48-polyubiquitination, is not a good choice *in situ*, simply because the length of ubiquitin chains cannot be manipulated without extra mutations in donor ubiquitin to block further chain elongation. [93-95] Therefore, solid-phase peptide synthesis (SPPS) has been widely used initially for this purpose. The general scheme for ubiquitin chain synthesis via SPPS was shown in Scheme 1-1.

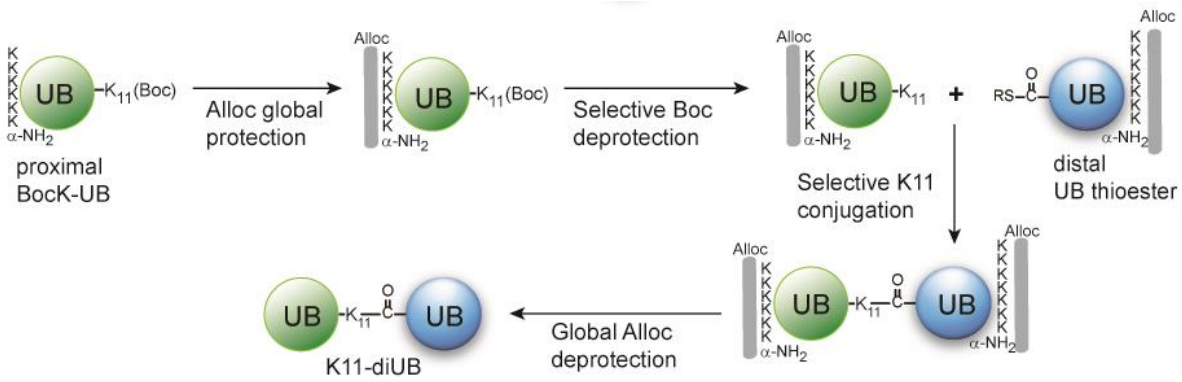


Scheme 1-1 The general synthetic route for ubiquitin synthesis with solid-phase peptide synthesis

Following this approach, total synthesis of diUB[96, 97] and tetraUB[98] of K48 linkage was reported by Brik group and Ovaa group respectively, who are the pioneers for SPPS of ubiquitin chains with predefined yet native linkages.[99] Nevertheless, SPPS enables design and incorporation of various reporting groups as warhead embedded in synthetic ubiquitin chains for labelling[100] (with fluorophores of FRET pairs, for example) or probing[101] applications. However, the pure synthetic approach is believed to be tenuous and material-consuming requiring extra steps of ligation and purification: the comfort length for SPPS is under 40 amino acids; synthesizing one ubiquitin moiety usually requires one or two steps of native chemical ligation.

1.4.1.2 Semi-synthetic ubiquitin chain with UAA incorporation.

More recently, Chin group and Fushman group [102, 103] separately reported to use a semi-synthetic approach to generate linkage-specific ubiquitin dimers, which they name as "genetically encoded orthogonal protection and activated ligation (GOPAL)". GOPAL takes advantage of the pyrrolysinyl-tRNA synthetase (PylRS) to incorporate a lysine analogue with protected ε-amino group, namely BocK or CbzK to ensure the specificity of position to be modified; the remaining amino functionalities (the N-terminal amine and six lysine residues) are globally protected with orthogonal Cbz or Alloc groups. Acid treatment selectively removes the Boc group to unmask the reactivity of the designated lysine, which is then used to couple with the intein-derived thioester of donor UB that is also globally protected on the amines. After coupling, the protecting groups are removed to afford diUB of a native linkage. Such an approach generates native isopeptide linkages between the two recombinant UBs produced in *E. coli*. (Scheme 1-2)



Scheme 1-2 GOPAL approach for linkage-specific ubiquitin chain semi-synthesis

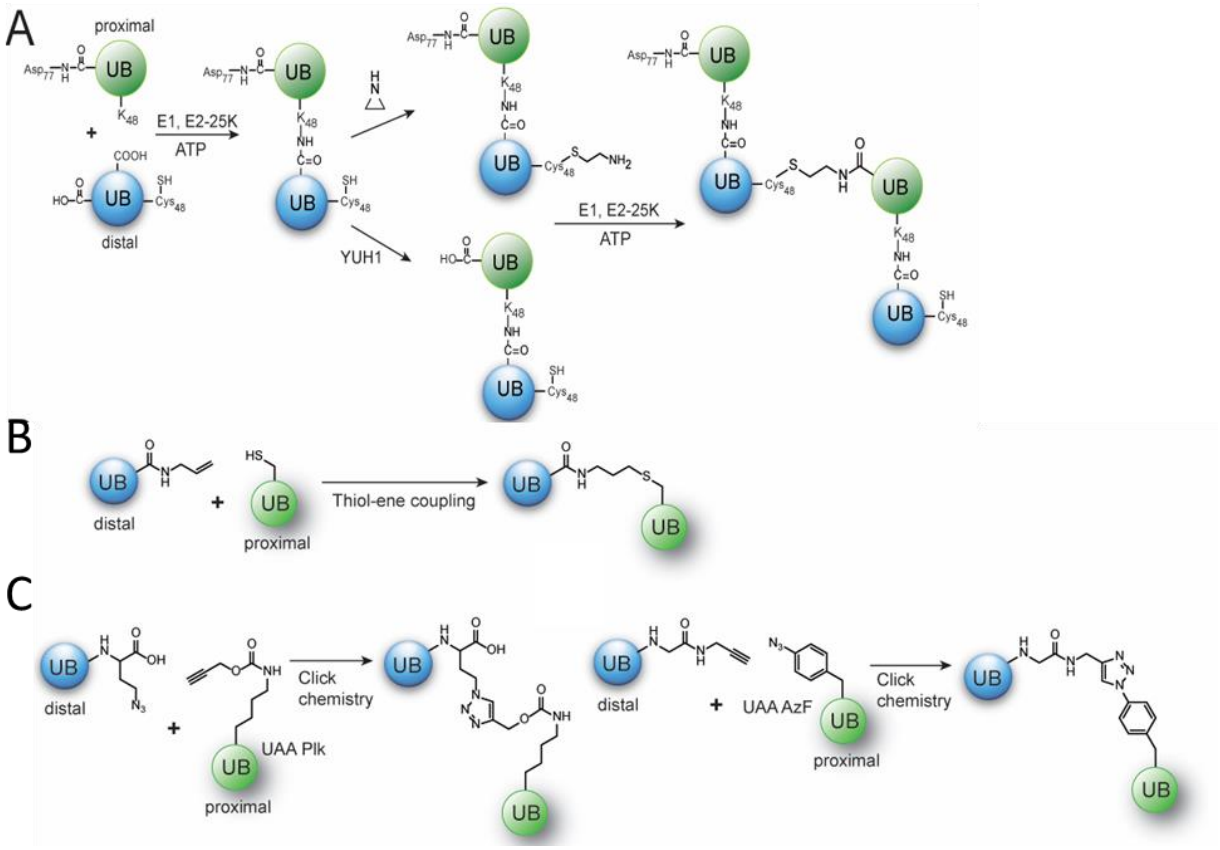
Another similar and even more elegant method for ubiquitin chains synthesis is incorporating unnatural amino acid into acceptor ubiquitin for direct ligation to activated donor ubiquitin without additional protection and deprotection on other amines.[104] This method utilizing

directed evolution for UAA incorporation, known as genetic code expansion (GCE) and will be discussed in Chapter 2.2 extensively.

1.4.1.3 Engineered ubiquitin chains of nonnative linkages

Ubiquitin chains of nonnative linkages are a close mimic of native chains that can be employed to reveal chain topology and identify reader proteins of ubiquitin chains; these linkages are relatively easy to assemble with a wider choice of conjugating chemistry.

As previously mentioned, linkage-specific E2s provides a facile way to assemble tetra UB chains or even longer conjugates with a periodic thioether linkage [105-107] (Scheme 1-3A). In such a method, to control the length of ubiquitin chain, the very lysine on donor ubiquitin is mutated into cysteine to block further elongation; the resulting diUB then reacts with ethyleneimine to



Scheme 1-3 Nonnative ubiquitin chain synthesis

generate a thioether C-terminus that can be conjugated to another diUB bearing a free C-terminus by treating with a yeast DUB YUH1. The two parts are combined for E2-catalyzed UB chain assembly to generate tetra UB chains with a thioether linkage. Such a process can be repeated to generate higher order of UB chains.

Non-enzymatic reactions such as thiol-ene coupling (Scheme 1-3B) have been used to assemble diUB and UB chains where a photoinitiator was used to induce radical formation on the thiol group of the Cys side chain. [108] Other researchers have been utilizing copper-catalyzed click chemistry (Scheme 1-3C) to make ubiquitin chain analogues. [109-111] Despite of the convenience these coupling chemistries bring, these linkages are regarded as nonnative with variations in the atomic length of chains or orientation of two ubiquitin moieties.

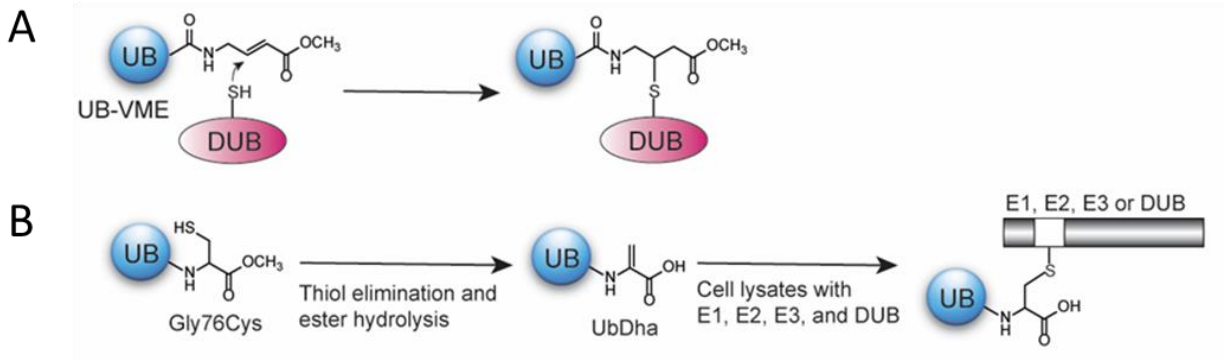
1.4.2 Activity-based protein profiling by ubiquitin probes

To profile the specificity in both ubiquitin transfer process and ubiquitin chain recognition process of define linkage, protein engineering tools have been developed to generate ubiquitin or ubiquitin conjugates equipped with reactive functionalities. These activity-based probes, often made available by protein engineering based on expressed-protein ligation (EPL) or unnatural amino acid incorporation, have helped researchers understand the writing, editing and reading of ubiquitin code.

1.4.2.1 UB with C-terminal electrophiles as probes

The C-terminus of ubiquitin can be conveniently derivatized by EPL with new functionalities most frequently electrophiles acting as DUB traps [112, 113]. A couple of examples described below involve generating an activated C-terminal carboxylate in the form of thioester bond by EPL, namely, UB~MES.[114] For instance, ubiquitin with a C-terminal vinylmethylester (UB-VME, Scheme 1-4A) has been generated through UB~MES as probe to identify upregulated

DUBSs in tumor cells or virally-infected cells. [115, 116] More recently, UB-Dha (Scheme 1-4B) with a dehydroalanine at the C-terminus was reported to probe catalytic cysteines in ubiquitin transfer cascade; E1, E2 and HECT E3 were able to be labelled. [117]

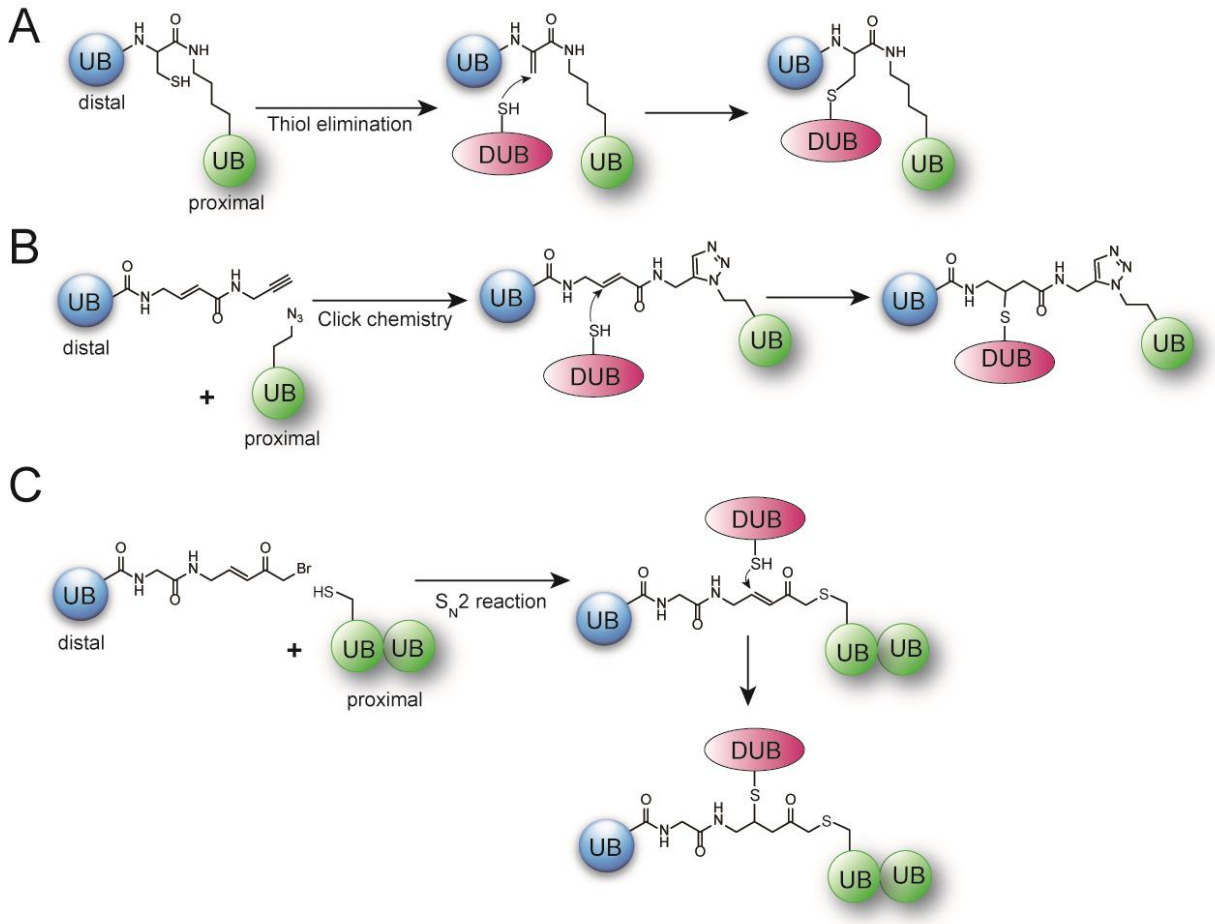


Scheme 1-4 mono UB probes

1.4.2.2 diUB probes with Cys traps to target DUB and E3s

To understand the ubiquitin chain specificity of DUBs or E3s, diUB probes are made with defined linkage types. Similar to the monoUB probes described in last section, a Gly to Cys mutation is inserted so that an electrophilic trap can be created. [118, 119] Following this idea, Brik et al synthesized linkage-specific diUB by total peptide synthesis to introduce a dehydroalanine at the linkage site as traps for catalytic Cys residues of DUB (Scheme 1-5A) [96, 120]; a linear dimer of ubiquitin embedded with a dehydroalanine was also reported to selectively label M1-polyubiquitin reactive DUB, OTULIN.[121] Nevertheless, the linkage between distal and proximal ubiqiuitins was also generated with click chemistry while the distal ubiquitin containing alkyne group was made available from aforementioned EPL and azido group on proximal ubiquitin was incorporated with genetic code expansion. (Scheme 1-5B) [122]. A Michael acceptor bearing α, β-unsaturated ketone was embedded and served as a reactive warhead towards corresponding DUBs. Furthermore, linkage-specific triUB probes could be assembled by coupling a distal UB functionalized with a

C-terminal α, β-unsaturated ketone linker with a Cys residue replacing a specific Lys on the diUB (Scheme 1-5C). The triUB probe has been used to investigate the recognition of UB chains by the DUB USP9X [123, 124]



Scheme 1-5 DiUB probes for DUB chain specificity profiling

2 DECIPHERING THE COMPLEXITY OF UBIQUITIN CHAINS: DESIGN AND SYNTHESIS OF DI-UBIQUITIN PROBES VIA GENETIC CODE EXPANSION

Understanding the ubiquitin chain specificity in ubiquitin transfer, elongation and edition processes proves to be essential to decipher ubiquitin code. To this end, tremendous work has been done to generate linkage-specific ubiquitin chains but ubiquitin polymers themselves didn't explain much of the specificity without corresponding regulators, such as chain builders, erasers or readers. Nevertheless, structural or biochemical studies of E2, E3 or DUB provides little knowledge towards chain specificity if their interactions with ubiquitin chains are ignored.

2.1 Mechanism-based Di-Ubiquitin probes design

Ubiquitin transfer, chain elongation and edition can all be simplified into a three-component event based on the molecular reaction mechanism. In these processes, an addition-elimination mechanism featuring a nucleophilic attack is highlighted: for ubiquitin transfer, a catalytic cysteine on HECT E3 will nucleophilic attack the corresponding E2~UB conjugated with a thioester bond (Figure 2-1A). Similarly, in chain elongation process, a pre-selected lysine residue on the acceptor UB or proximal UB will nucleophilic attack the corresponding E2~UB where C-terminus of the donor or distal UB is fully activated (Figure 2-1B) whereas in chain edition or ubiquitin removal process, the catalytic cysteine on DUB will nucleophilic attack the amide of the isopeptide bond (peptide bond for M1 linkage) between distal and proximal UBs (Figure 2-1C). Taken together, the specificity implicated in these processes or reactions is determined by the intact interactions in between these three components.

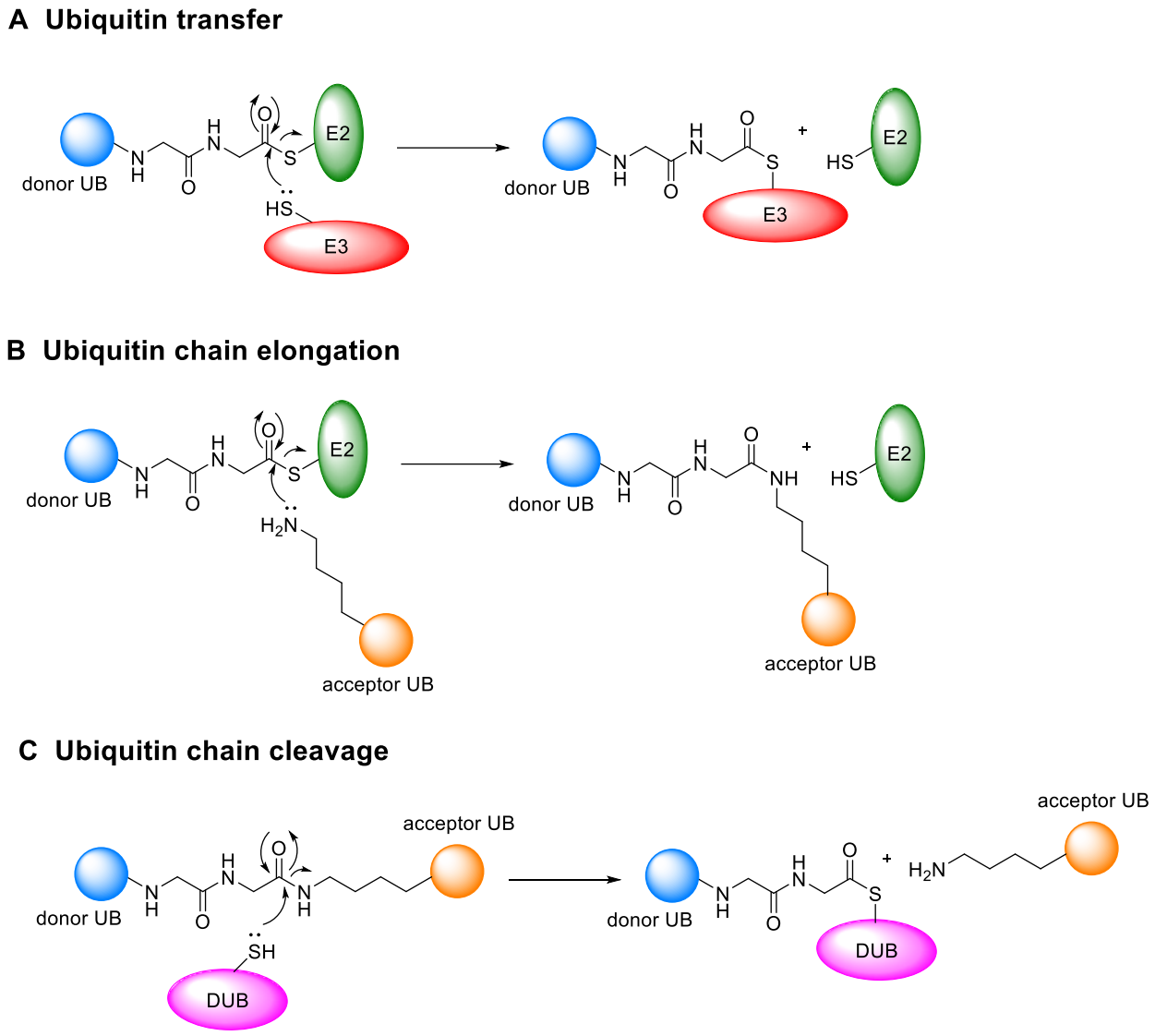


Figure 2-1 Mechanism of ubiquitin transfer, chain elongation and removal

In this thesis, the molecular logic of linkage specificity revealed by E2, E3 and DUB is focused. Taken E2 as an example, the ubiquitin chain elongation adapts a typical additional-elimination mechanism. Initially, donor UB is covalently bound to a particular E2 through a thioester bond (**1**); the thioester bond is prone to be nucleophilic attack by an acceptor UB forming a tetrahedron intermediate (**2**). (Figure 2-2) At this rate-determine step, the interaction between the acceptor UB and E2 is crucial for recognition of these components in that E2 shall orientate

acceptor UB and potentially select a specific lysine residue (M1, K7, K11, K27, K29, K33, K48, K63) that facilitates the nucleophilic attack by its ε-amino group. Following this logic, it was hypothesized that if one can snapshot this transient tetrahedron intermediate (**2**) by trapping it with stable mimics, the intact interactions between acceptor UB and its corresponding E2, donor UB and acceptor UB, donor UB and E2 can be captured and visualized simultaneously.

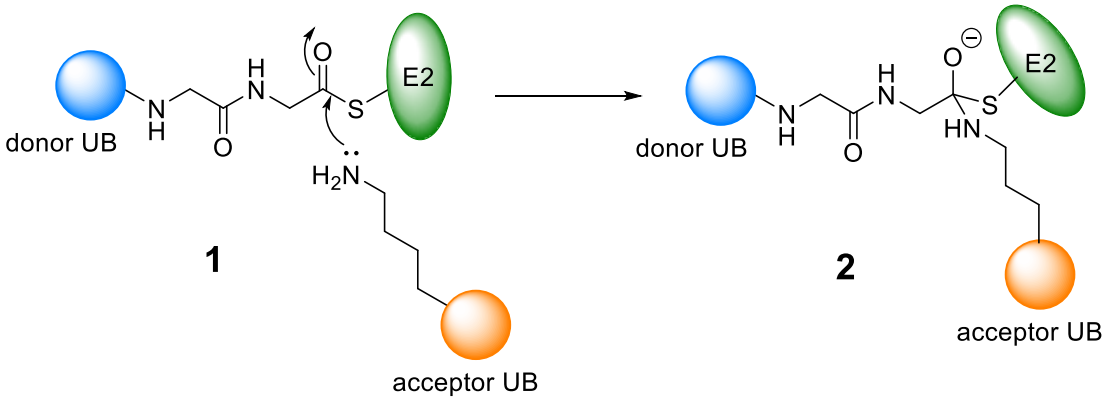


Figure 2-2 Tetrahedron intermediate formation in ubiquitin chain elongation

In these interfaces imprinted by this snapshot, the interaction between acceptor UB and E2 is the most important yet hard to capture, given that it reveals the direct molecular evidence that E2 orientate acceptor UB to ensure linkage specificity but it was believed the binding affinity between acceptor UB and E2 is low to mM level. In addition, the interaction between the two UB moieties also potentially plays a role to maintain chain specificity especially those ubiquitin chains adopting a "closed" conformation (K48, K11 for instance, Chapter 1.2.1). Nevertheless, reported crystal structures of donor UB~E2 suggest a flexible conformation featuring E2 links to the long tail of donor UB C-terminus; a third component (either an acceptor UB, HECT E3 or substrate) is critical to solidate the conformation of donor UB~E2 into a rigid complex of three to facilitate ubiquitin chain elongation or ubiquitin transfer to substrates. This hypothesis serves as the

guideline for diUB probe design in the thesis and will be further discussed in later sections (Chapter 4.1).

Some criterion need to meet when design mimics for protein-protein interaction (PPI) mapping: firstly, the protein complex should be chemically or enzymatically accessible with bio-compatible ligation reactions; additionally, protein complex should be rigid and stable for characterization and further analysis; most importantly, with bio-organic tools available, this mimic should be as close resemblance to native one as possible with the least perturbance in structures or intact interfaces. To this end, a novel complex featuring a disulfide linkage (**4**) or a thioether linkage (**6**) is designed (Figure 2-3B). To probe the chain specificity of E2 (as an example shown in Figure 2-3B), HECT E3 or DUB, a di-UB probe with defined linkage (**3**) embedded a G76C mutation is utilized in such a way that the inserted cysteine at the linkage site can serve as a reactive warhead to form a disulfide bond with corresponding E2; therefore, the resulting three-component complex (**4**) is a stable mimic for the transient tetrahedron intermediate (**2**). Compared to **2**, the mimic counterpart **4** is very close to native with the catalytic thiol on E2 three-atom away from the catalytic center for this addition-elimination reaction; the isopeptide bond formed between the C-terminus of donor UB and ε-amino group on the lysine of the acceptor UB is completely native.

Nevertheless, this diUB probe (**3**) can be easily converted into an electrophilic trap by thiol elimination; the resulting diUB probe featuring a dehydroalanine next to the catalytic center for addition-elimination reaction (**4**) can be used as a 1,4-Michael addition acceptor to report any catalytic thiol on E2/HECT E3/DUB that recognize this ubiquitin chain type. Notably, the carbon for thiol addition is only two-atom away from the native catalytic center.

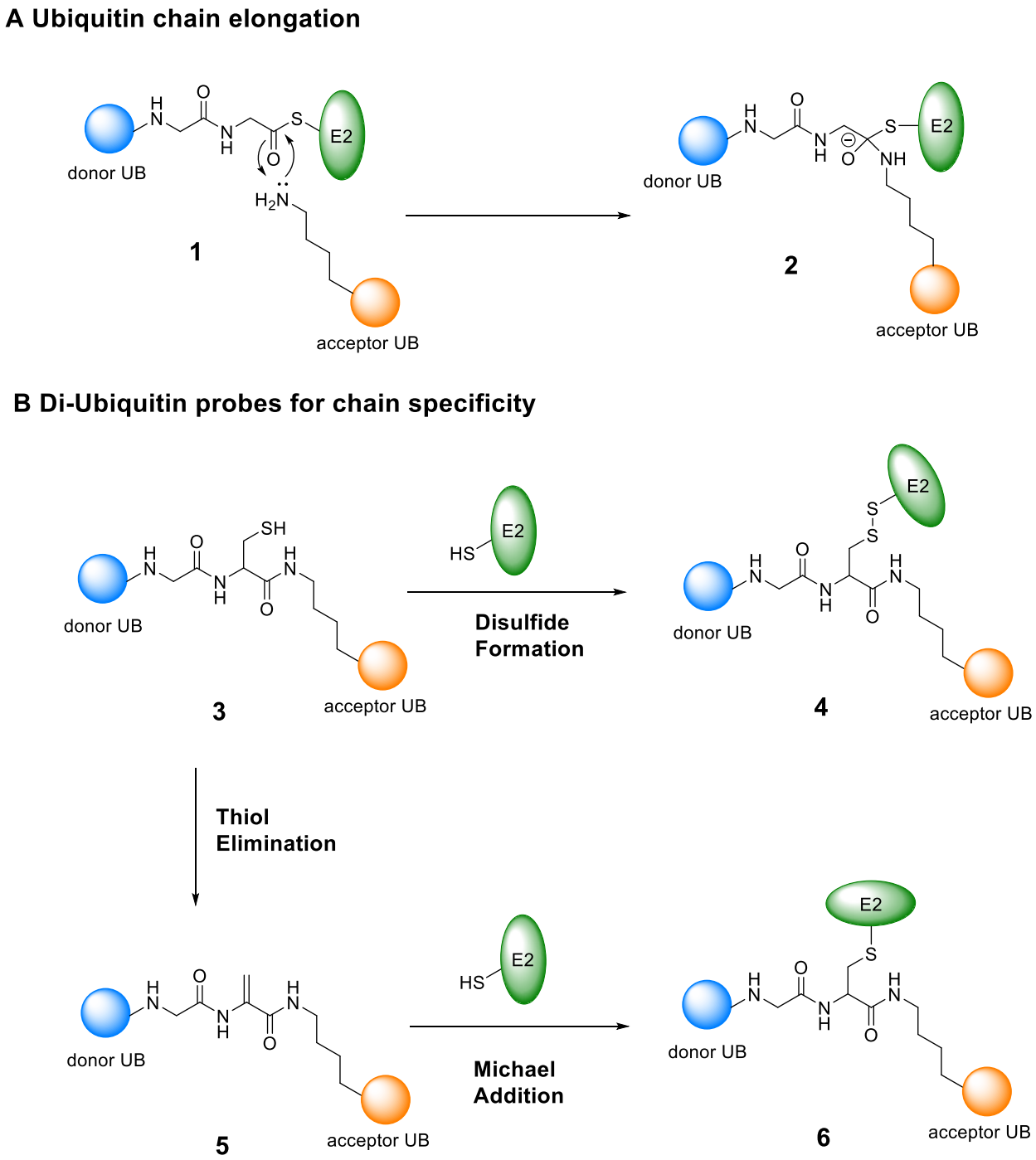


Figure 2-3 Mechanism-based design of Di-ubiquitin probes

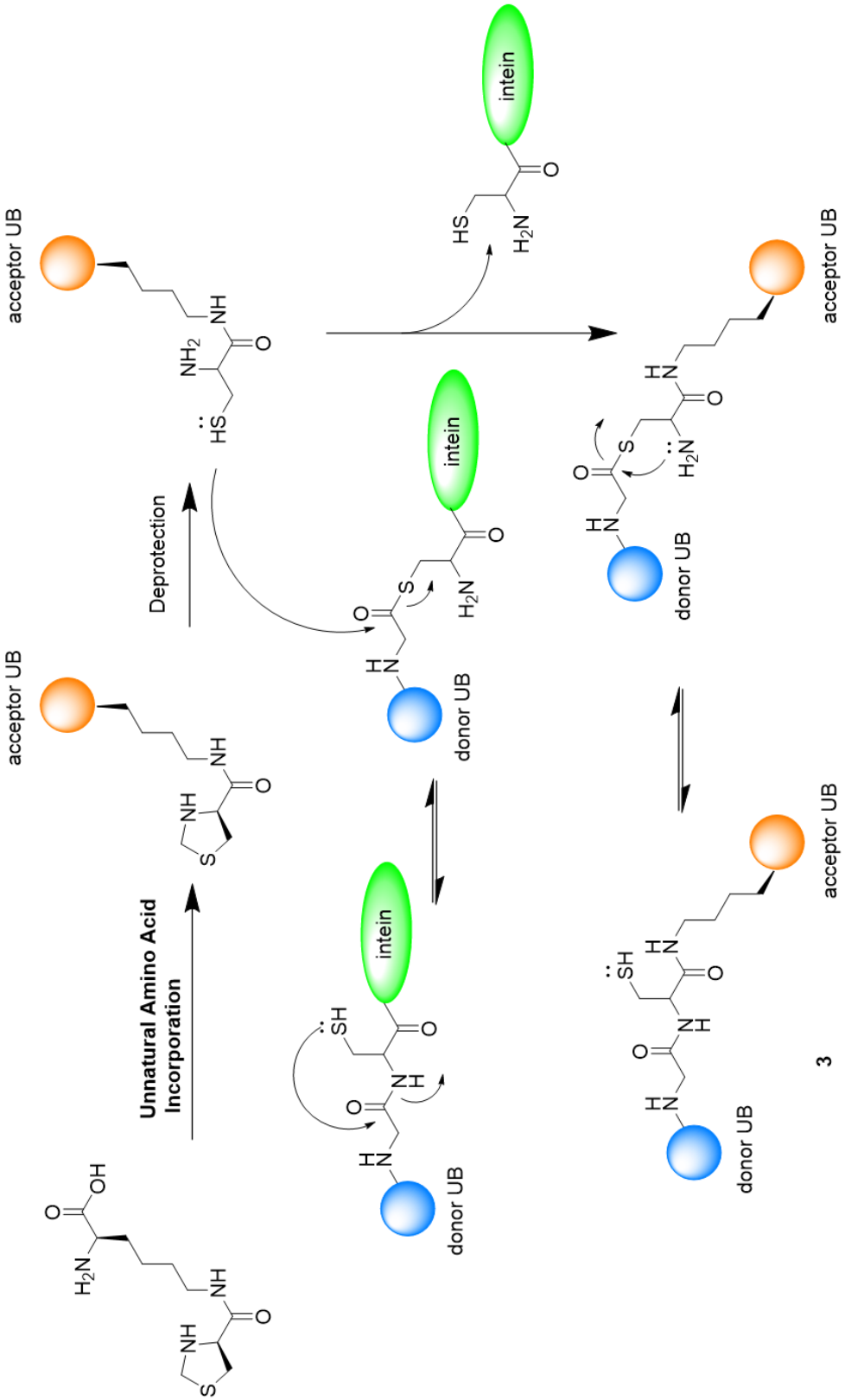
In a retro-synthetic perspective, the question narrows down to the accessibility of linkage-specific UB dimer with a thiol reactive group (**3**). Although several synthetic or semi-synthetic approaches have been summarized in Chapter 1.4, a more convenient method via genetic code

expansion has been investigated. (Scheme 2.1). Instead of introducing the chain specificity (selected lysine to be modified) by SPPS, the amber codon position for unnatural amino acid incorporation ensures the specificity of UB chain assembly by nature; unlike GOPAL approaches that requires tenuous steps of protection and deprotection, the 1-amino-2-thiol- reactive functionality is directly introduced in the form of synthesis-compatible unnatural amino acid, namely N^{ε}-L- thiaprolyl-L-lysine (L-ThzK), by an pre-engineered pyrrolysinyl tRNA synthetase derived from directed evolution.

Upon the UAA incorporation, a deprotection reaction is used to convert UB-ThzK into UB-CysK to reveal the 1-amino-2-thiol reactive motif; CysK is not incorporated directly because 1-amino-2-thiol functionality was reported to react with intrinsic pyruvate to form an adduct.[104, 125] Donor UB is engineered as an intein fusion so that a reactive thiolester can be readily formed after the intein cleavage. Finally, the native linkage of an isopeptide bond is achieved through native chemical ligation.

2.2 Resolving the linkage specificity: genetic code expansion approach

Unnatural amino acid (UAA) incorporation into proteins via genetic code expansion has greatly expanded the structural diversity and chemical reactivity of the functionalities on a protein scaffold. [126-129] UAA-incorporated proteins have become powerful tools for biological studies - they can generate precise posttranslational modification patterns[130] or be site-specifically attached with customarily made chemical probes to enable photocrosslinking[131], biorthogonal labeling[132], photocaging[133] or enzymatic profiling capacities[134] on the proteins.



Scheme 2-1 Synthetic scheme for diUB probe

Scheme 2-1 Synthetic scheme for diUB probe

In principle, unnatural amino acid incorporation hijacks the native translational machinery for nascent peptide synthesis and transform the ambor stop codon (UAG) into a sense codon by introducing another orthogonal pair of aminoacyl-tRNA and aminoacyl-tRNA synthetase (RS) to insert UAA at UAG position. (Figure 2-4)[135, 136] Notably, this extra pair of aminoacyl-tRNA and its cognate synthetase shall be orthogonal to intrinsic ones in a way that the extra synthetase introduced can charge pre-selected UAA onto its corresponding tRNA exclusively from any of the 20 natural ones and vice versa. To this end, a general strategy for generating orthogonal tRNA/aminoacyl-tRNA synthetase pairs in bacteria makes use of orthologues from archaeal bacteria or eukaryotic organisms considering that some prokaryotic tRNA/aminoacyl-tRNA

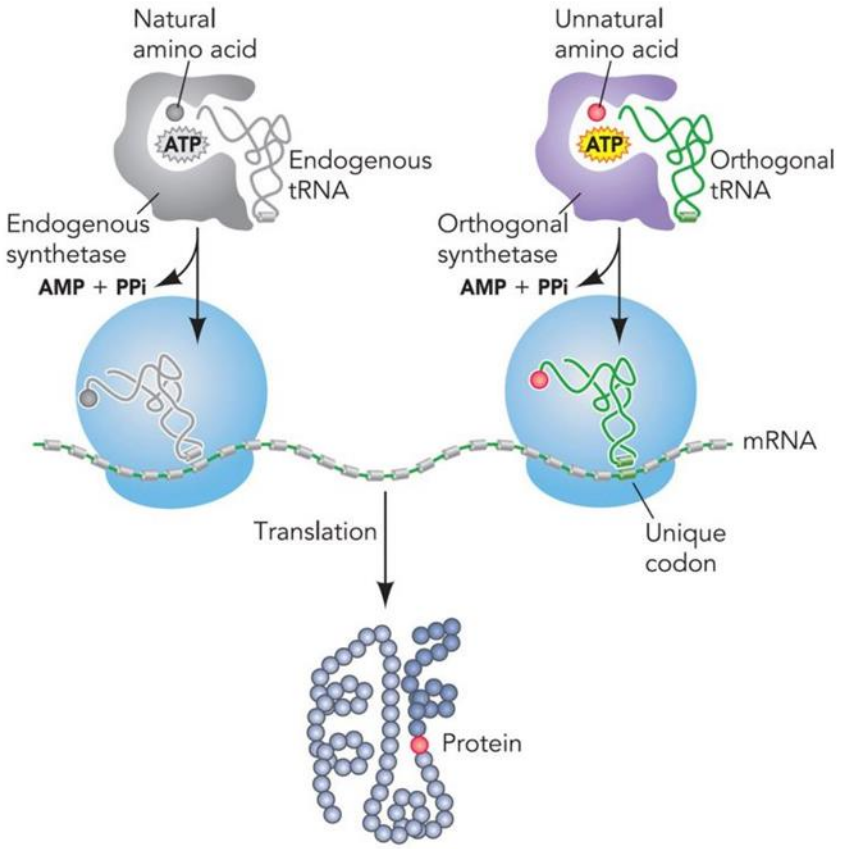
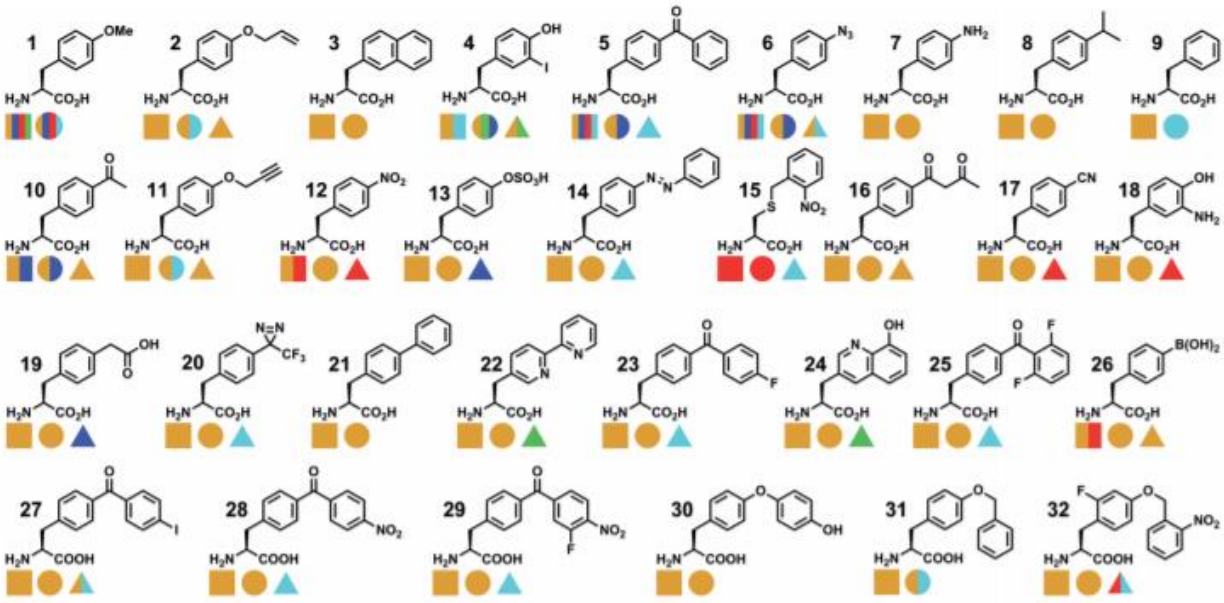
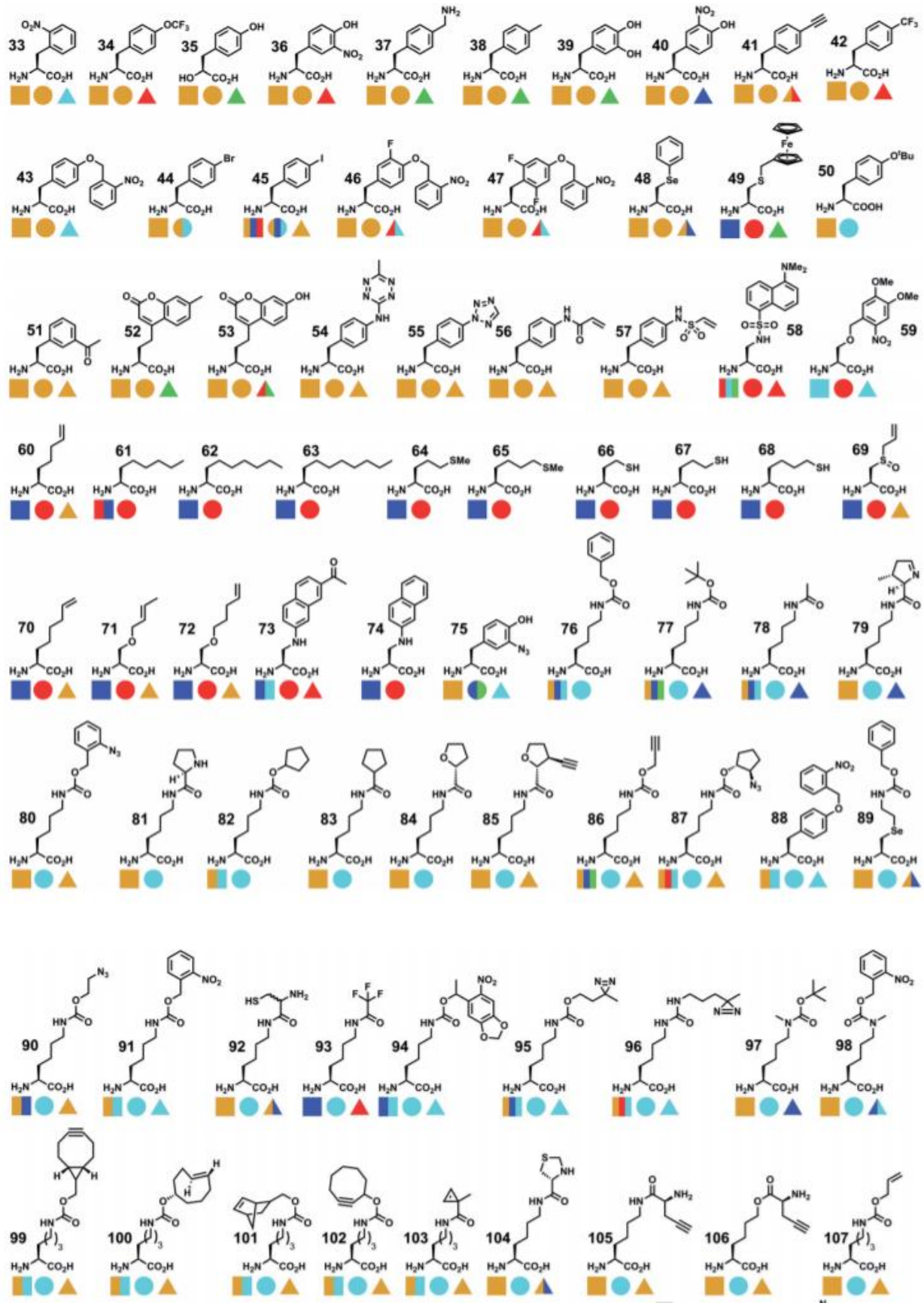


Figure 2-4 UAA incorporation through ambor codon suppression

synthetase pairs do not cross-react to any significant degree with their eukaryotic counterparts, as a result of differences in tRNA identity elements, primarily in the acceptor stem and variable arm[137]. Therefore, the first orthogonal E. coli tRNA/synthetase pair to be generated from archaeal bacteria was derived from the tyrosyl pair from *Methanococcus jannaschii*[135]. However, the wide-type Mj tyrosol-tRNA synthetase can only take tyrosine as a cargo, hence extra mutations around the amino acid binding region on the Mj tyrosol-tRNA synthetase need to be introduced to kill the original activity of charging wide-type tyrosine and compensate it with capacity of desired unnatural amino acid. The mutations are selected from several rounds of positive and negative selection of a tyrosol-tRNA synthetase library randomizing the key residues for amino acid recognition[128, 138]; this directed evolution approach was awarded Nobel Prize in Chemistry in 2018.

The first generation of UAAs in the GCE toolbox is mainly tyrosine derivatives inserted by Mj tyrosol-$tRNA_{CUA}$/synethetase pair. Recently, a large number of lysine derivates were reported to be successfully incorporated into different organelles by *Methanosarcina barkeri* or *Methanosarcina mazei* pyrrolysyl-synthetase (Pyl-RS)/ pyrrolysyl tRNA (Pyl-tRNA) pair. (Figure 2-5) [139-143]




33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107

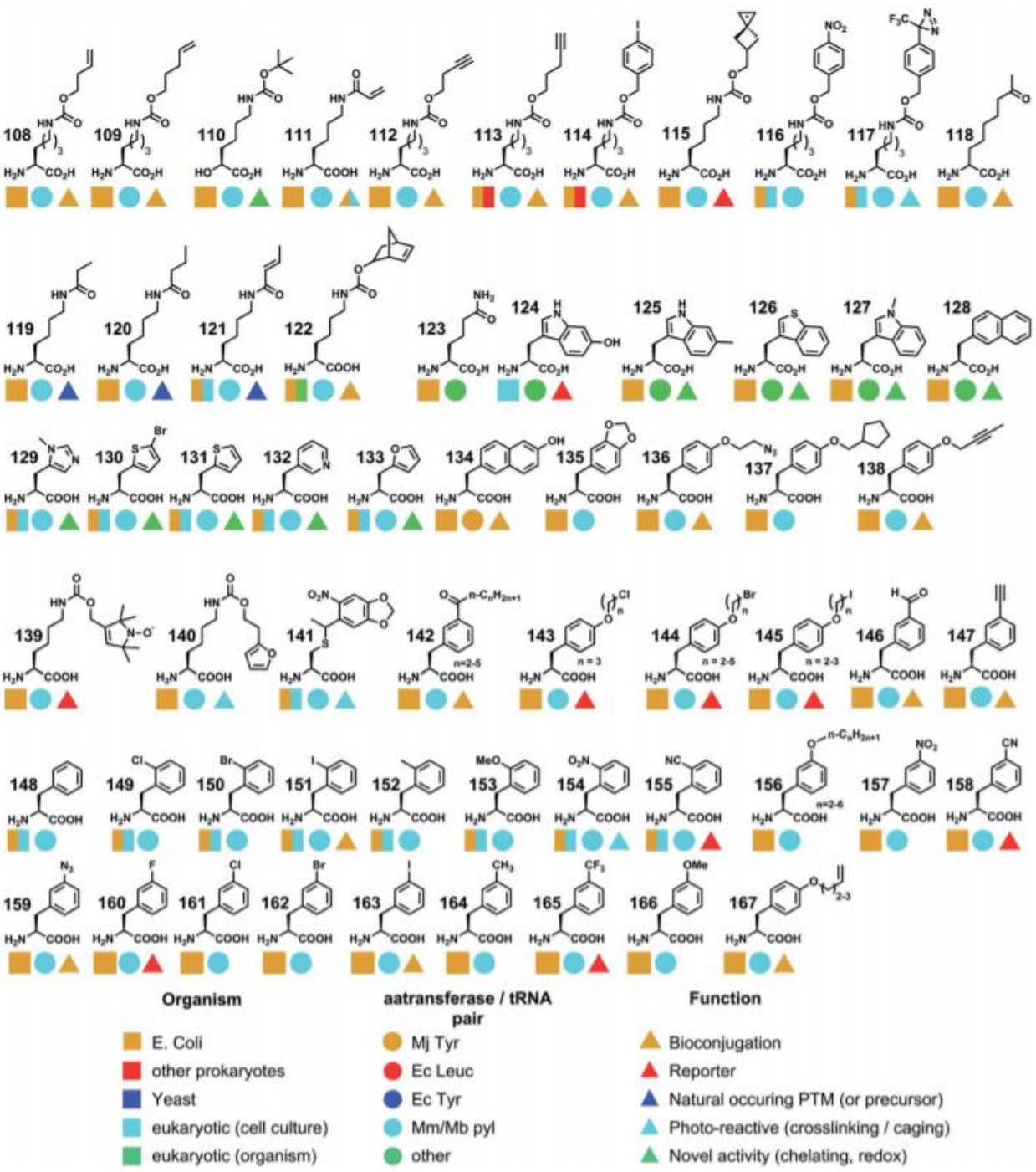


Figure 2-5 Unnatural amino acids that have been genetically encoded in proteins to date.

Notably, incorporation of lysine derivatives has largely enriched GCE toolbox to investigate post-translational modifications on lysine, such as acetylation, methylation and ubiquitination.[130] A more convenient GCE method to generate native and site-specific ubiquitin

chain was reported utilizing δ-thiol-Nε-(p-nitrocarbobenzyloxy)-L-lysine into acceptor UB (Figure 2-6). [104] Unlike SPPS or GOPAL methods discussed in Chapter 1.4.1, genetic code expansion featuring ambor codon suppression ensures incorporation at a pre-defined position for ubiquitin elongation. Furthermore, direct insertion of a protected 1-amino-2-thiol functionality bypasses several steps of protection and deprotection because other amines don't require pre-treatment in native chemical ligation.

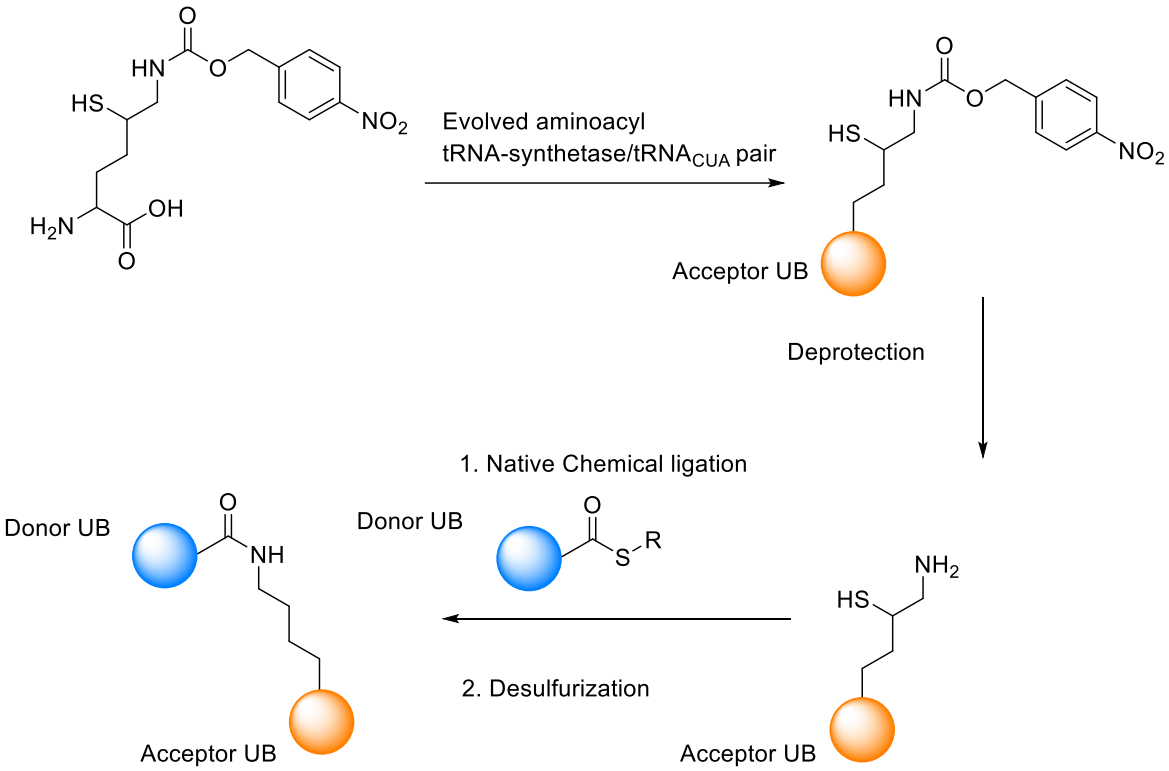


Figure 2-6 GCE approach for traceless and site-specific diUB synthesis

2.3 Incorporation of L-Thiazolidine Lysine (L-ThzK) into Ubiquitin via GCE

Similarly, L-thiazolidine-L-lysine was reported to be incorporation into protein via GCE to facilitate orthogonal protein labeling.[125] L-thiazolidine-L-lysine is also a 1-amino-2-thiol functionality precursor that can be condensate with Cyanobenzothiazole.(Figure 2-7). In this

paper, Chin lab has reported an evolved Pyl-RS to incorporate L-ThzK bearing which contains three programmed mutations: A267S, C313V, M315F and one nonprogrammed mutation D344G. Consequently, n this thesis, we decided to utilize this UAA incorporation to generate acceptor UB bearing a 1-amino-2-thiol motif for native chemical ligation (NCL).(Scheme 2-1) Notably, this inserted thiol serves both as a nucleophilic trigger for NCL and a reactive warhead for enzymes conjugation or profiling. (Figure 2-3)

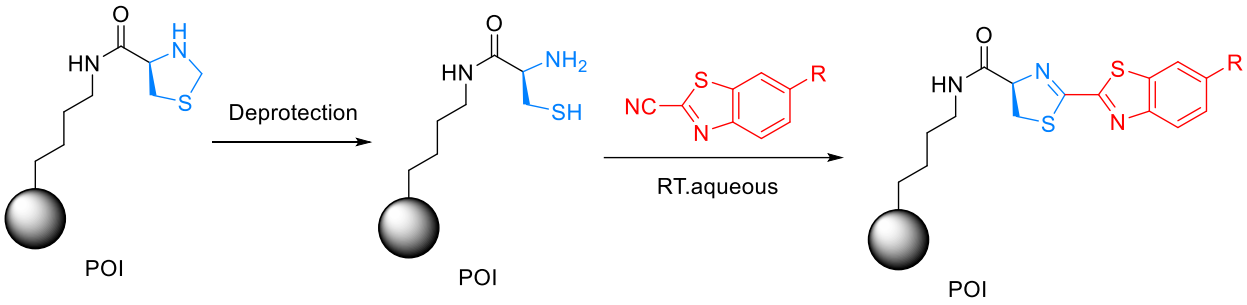
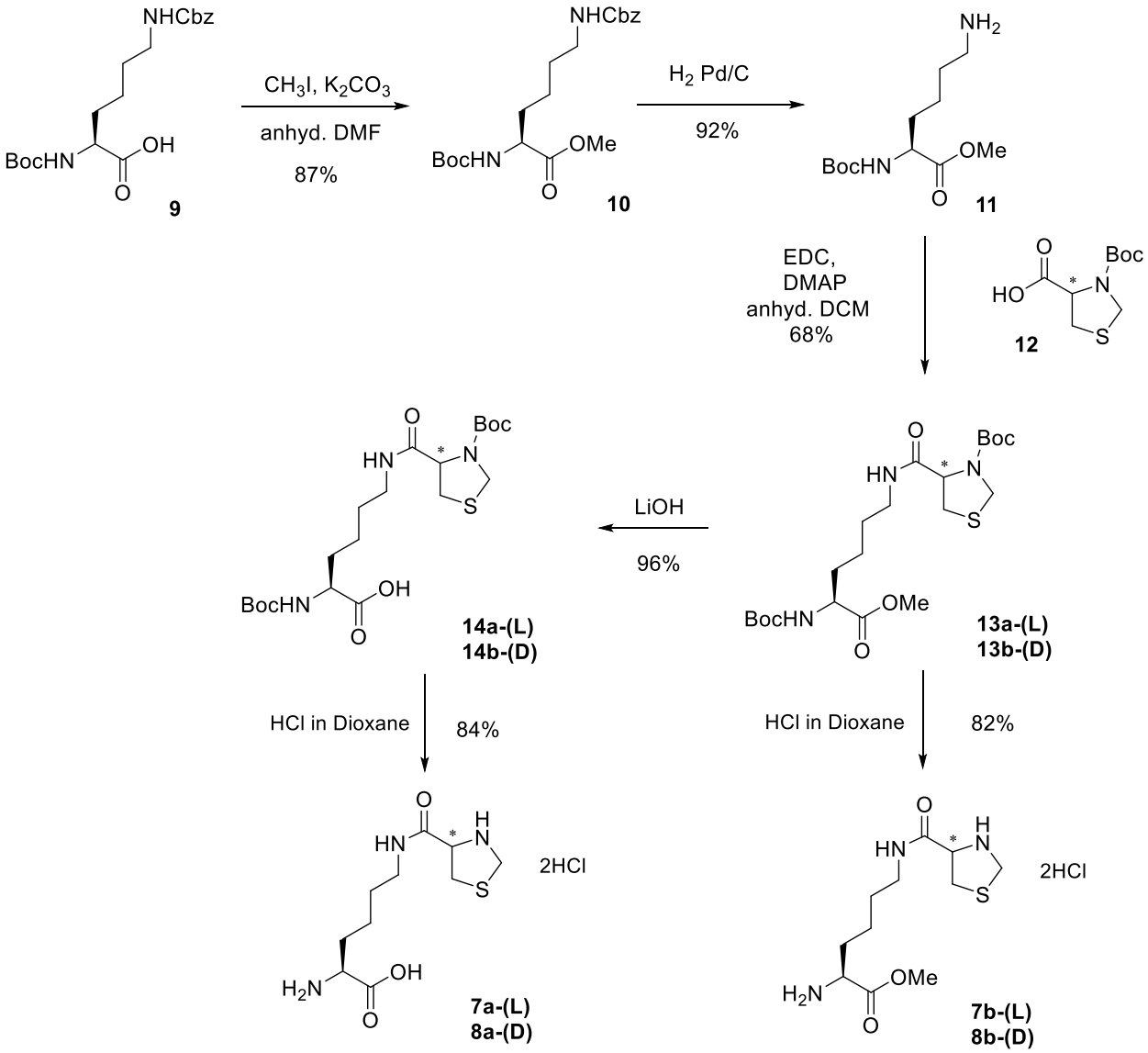


Figure 2-7 Incorporation of L-ThzK for bioorthogonal labelling: original reported application for ThzK incorporation

2.3.1 Synthesis of N^{ε}-L- thiaprolyl-L-lysine (L-ThzK)

Instead of preparing ThzK from N^{ε}-Boc-L-Lys-OMe **11**, [125] we chose N-Boc-Lys(Cbz)-OH **9,** a cheaper starting material to begin our synthesis (Scheme 1B). We first methylated **3** to generate protected Lys methylester **10** and then deprotected Cbz group on **4** by hydrogenation to afford **11**. We then coupled **11** with L or D form of Boc-thiazolidine-4-carboxylic acid **12** to afford Boc and methylester protected ThzK (**13a** and **13b**). We either removed all the protecting group on **13a** and **13b** to generate the free acid forms of L/Z-ThzK-OH (**7a** and **8a**) or only removed the Boc group to generate the methylester forms of ThzK (ThzK-OMe, **7b** and **8b**). Through this procedure, we achieved an overall yield of 25% for the synthesis of ThzK-OMe and similar yields

for ThzK-OH. Nevertheless, the final product was purified with HPLC (preparative-C12 column). The usage of ThzK-OMe will be discussed in the next section. Detailed procedure of synthesis with full characterization is in Chapter 2.6.



Scheme 2-2 ThzK synthesis

2.3.2 Incorporation of L-Thiazolidine-L-Lysine(L-ThzK) into Ubiquitin

Genetic code expansion relies on extra vectors encoding orthogonal tRNA and tRNA synthetase to compensate additional UAA incorporation. To express ubiquitin incorporating L-ThzK, ubiquitin gene featuring a stop codon (TAG) at K11, K48 or K63 position was inserted into original pBAD mother vector kindly provided by Chin lab; the very vector also encodes Pyrrolysinyl $tRNA_{CUA}$. Additionally, pBk vector also from Chin lab encoding engineered L-ThzK Pyrrolysinyl synthetase was co-transformed into DH10B cells. L-ThzK at 1 mM was added when the cell culture reached the OD 0.6. Detailed procedure of protein expression is documented in Chapter 2.6.

Following the described protocol, 1 liter of UB expression would yield ~6 mg of UB with L-ThzK incorporation after purifying the 6×His-tagged UB by a Ni-NTA affinity column (Figure 2-8). Although UBK11ThzK was able to be purified and well characterized by mass spectrum (MALDI), unfortunately, the yield was not decent enough to provide sufficient protein for the following steps. A few modifications on the original reported protocol was examined including increasing UAA concentration to 5mM, increasing induction temperature from 30 degree to 37 degree, changing vectors (pCDF vector encoding ubiquitin gene or pBk vector encoding ThzK-PylS), but the yield didn't increase a lot.

2.4 Methyl ester form of Unnatural amino acids that increases incorporation Efficiency in GCE

Previously Wang and co-workers demonstrated the beneficial effect of acetoxymethyl ester form of a Tyr UAA analogue in enhancing its incorporation into GFP in mammalian HEK293

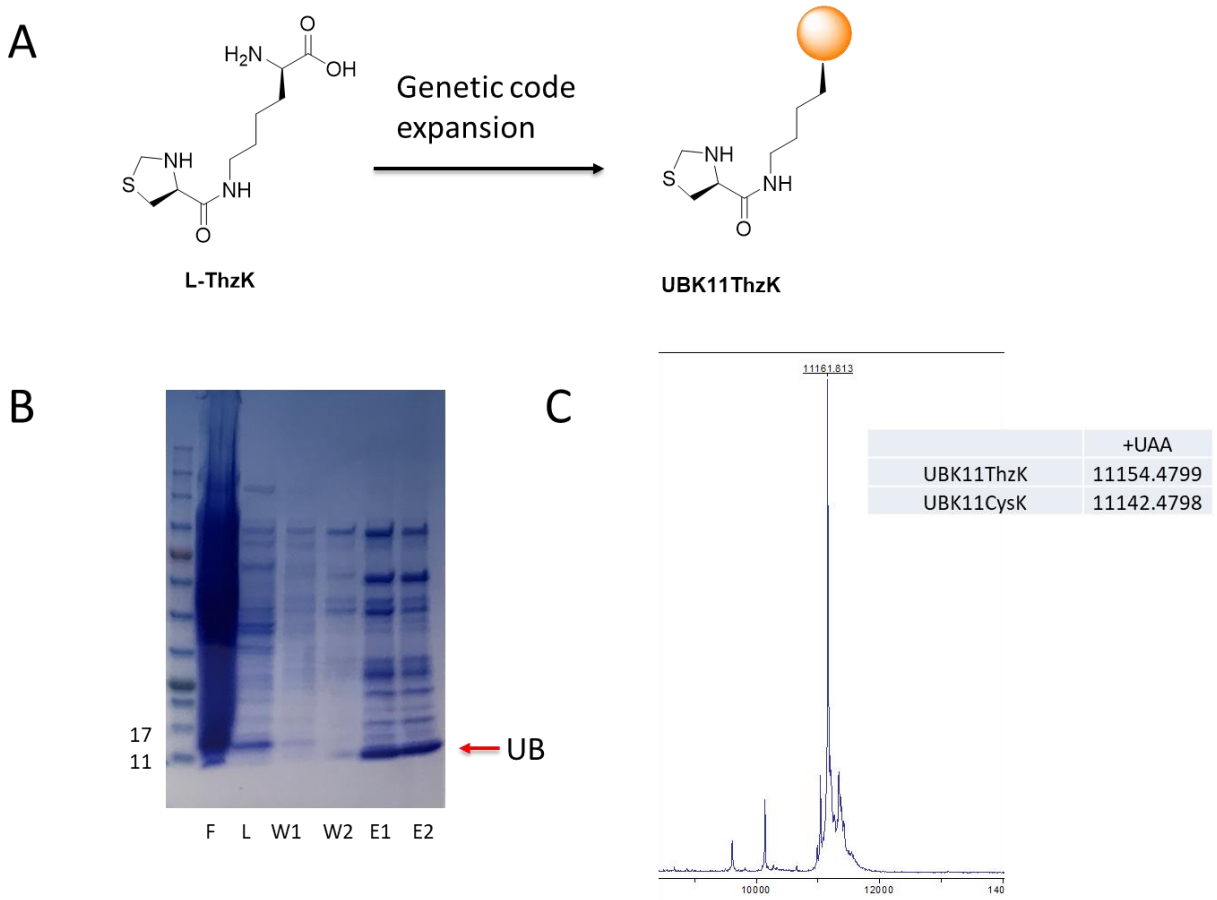


Figure 2-8 Genetic code expansion for L-ThzK protein expression

(A) Scheme for L-ThzK incorporation into UB yielding UBK11ThzK (B) Coomassie gel of fractions in Ni-NTA affinity purification. F: Flow-through; L Lysis buffer wash; W1, W2 Wash buffer wash; E1, E2 Elution buffer wash with a UB band at around 11kD. (C) MALDI-LP shows the correct molecular weight (11161), corresponding to calculated molecular weight (11154)

cells. [144] Liu's group reported the use of methylester form of L-ThzK to achieve good yield of incorporation into UB in *E. coli* cells. [145, 146] We suspected the L- and D-ThzK-OH in their free carboxylic acid forms could have poor membrane permeability to the *E coli* cell and so they would not be readily available for incorporation to the target proteins. We thus tried methylester forms of L and D-ThzK-OMe (**7b** and **8b**) for ThzK incorporation into K11 of UB.

To test the incorporation efficiency, we first used the engineered L-ThzKRS[125] to express UB with L-ThzK incorporated at K11 of UB. We supplied the cell culture with 1 mM L-ThzK-

OH (**7a**) UAA and found 1 liter of UB expression would yield ~6 mg of UB with L-ThzK incorporation after purifying the 6×His-tagged UB by a Ni-NTA affinity column (Figure 2-9A). We also attempted D-ThzK-OH (**8a**) incorporation into UB with ThzKRS originally engineered for L-ThzK incorporation and found a similar yield of UB expression (Figure 2-9B). We repeated the experiment for GFP incorporation and found the incorporation efficiency of either L- or D-ThzK-OH into GFP was much lower than that of ε-Boc-Lys (BocK) when wt PylRS of the *Methanosarcina mazei* (*Mm*) or *Mb* origin were used for protein expression (Figure 2-10A and Figure 2-10B). When we used the engineered ThzKRS derived from *Mb* PylRS for incorporation,

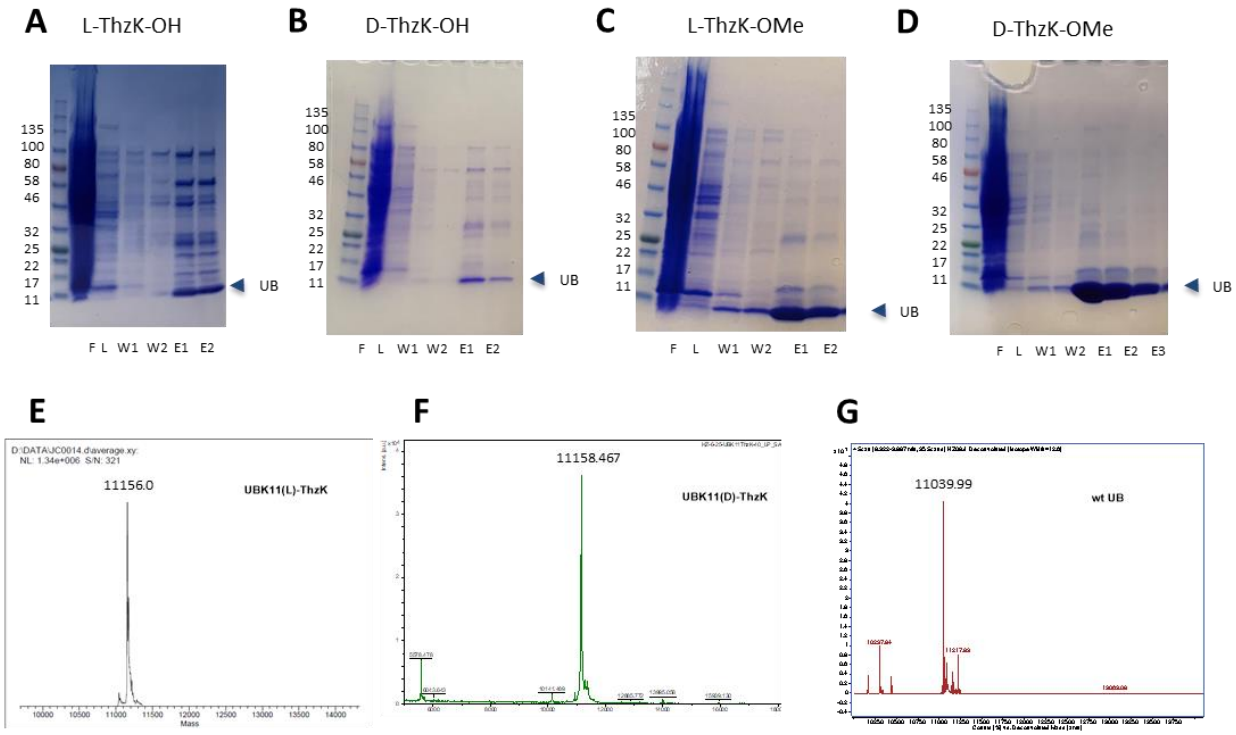


Figure 2-9 Incorporating L- or D-ThzK into UB at K11 with the free acid or methylester forms of the UAA

(A) UB expression with *Mb* ThzKRS and L-ThzK-OH. The Coomassie stained gel of protein samples from the flow through of the Ni-NTA column (F), lysis buffer wash (L), twice wash buffer wash (W1 and W2), and twice elution buffer wash (E1 and E2) were shown. (B) UB expression with ThzKRS and D-ThzK-OH. (C) UB expression with ThzKRS and L-ThzK-OMe. (D) UB expression with ThzKRS and D-ThzK-OMe. (E), (F) and (G) MALDI spectra of UB with L-ThzK, D-ThzK and Lys incorporated at K11, respectively. Calculated MS for each specie: 11154, 11154, 11039. The difference in the MW of UB with ThzK and with Lys incorporated is 115, matching the addition of the thiazolidine carboxylate to the Lys residue at K11.

the yield of L/D ThzK incorporation into GFP was still low as judged by the weak bands of GFP on the SDS-PAGE gel stained with Coomassie blue (Figure 2-9C). BocK UAA is known for its high efficiency of incorporation into various proteins by wt PylRS so such a result was expected. [147] Correspondingly, the GFP fluorescence from the cell culture also showed significantly higher fluorescence for incorporating BocK into GFP with wt PylRS of *Mm* and *Mb* origin. In contrast, the GFP fluorescence with L or D-ThzK-OH incorporation was low across the board with either wt PylRS or ThzKRS (Figure 2-9D).

We supplied 1 mM L-ThzK-OMe to the cell culture to compare the efficiency of incorporation with the same concentration of ThzK-OH. We found we could express an average of 33 mg of L-ThzK-incorporated UB from 1 liter of cells while the yield of L-ThzK incorporation into UB with L-ThzK-OH was only 6 mg (Figure 2-9A and 2-9C). The results of protein expression were consistent among three trials. We found similar improvement when we compared the yield of D-ThzK incorporation into UB with D-ThzK-OMe and D-ThzK-OH. The methylester form of UAA yielded 37 mg of UB from 1 liter of cell culture while the free acid form yielded 6.3 mg of UB, 17% of the methylester UAA (Figure 2-9B and 2-9D). The correct incorporation of ThzK into UB using L- or D-ThzK-OMe was confirmed by MALDI spectrometry comparing the difference of the MWs of UB with either ThzK or Lys incorporated (Figure 2-9E, 2-9F and 2-9G).

Our work showed that the yield of ThzK incorporation into UB could be increased by as much as 6-fold with the use of the methylester form of the UAA. With the use of the same tRNA synthetase and *E coli* strain for incorporation, the enhancement of the yield is most likely due to the improved membrane permeability of the methylester forms of the UAA to *E coli* cells comparing to the free acid forms of the UAA.

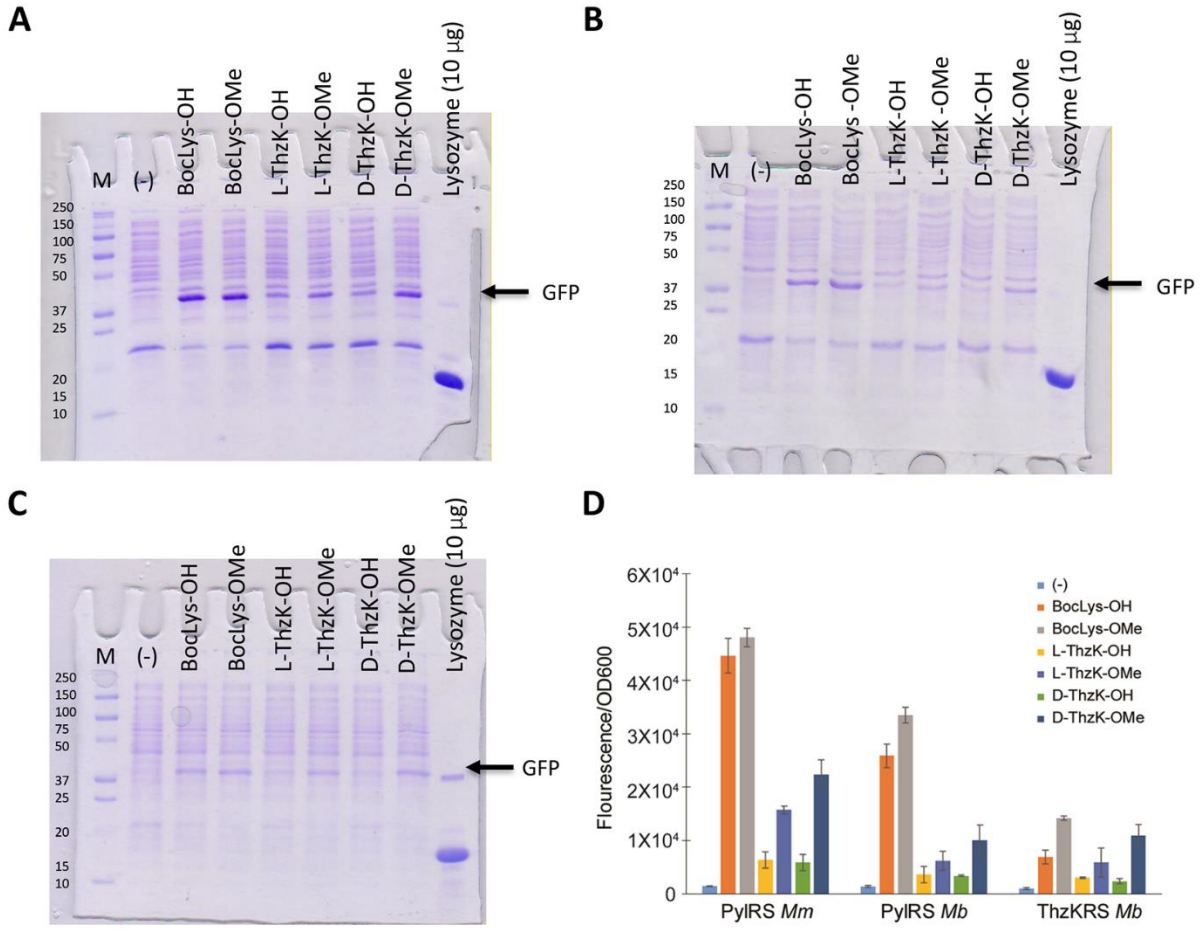


Figure 2-10 Comparing the yield of BocK and ThzK incorporation into GFP with wt PylRS and ThzKRS.

(A) Using wt *Mm* PylRS to incorporate various forms of UAA into GFP. Efficiency of incorporation was evaluated based on the intensity of the GFP band on the gel. (B) Using wt *Mb* PylRS to incorporate various forms of UAA into GFP. (C) Using ThzKRS engineered for L-ThzK incorporation to measure the yield of GFP expression with various UAA. (D) The intensities of fluorescence signals from cells expression GFP using various UAA and tRNA synthetase pairs. The results are the average of three trials.

Here we carried out detailed comparison of the yields of L- and D-ThzK incorporation between the methylester and free acid forms of the UAA and found ThzK methylester could enhance UAA incorporation by 6 folds. We suggest the methylester forms of UAA could benefit the incorporation if the free acid forms of UAA give poor yields. Our results may help to revive the interest in some UAA-engineered tRNA pairs by enhancing the permeability of UAA in its methylester form. On the other hand, there is only moderate increase of BocK incorporation into

GFP when the BocK-OMe form is used comparing to the free acid form (Figure 2-10D). So if the free acid form of UAA is incorporated well into proteins, further enhancement in yield with the methylester form is not obvious.

2.5 Building up ubiquitin chains: native chemical ligation (NCL)

2.5.1 Finalize acceptor UB: transforming UBThzK into UBCysK

Incorporated thiazolidine needs to be deprotected in order to reveal the 1-amino-2-thiol functionality. For this purpose, methoxyamine works as a competitor for the aldehyde group[148],

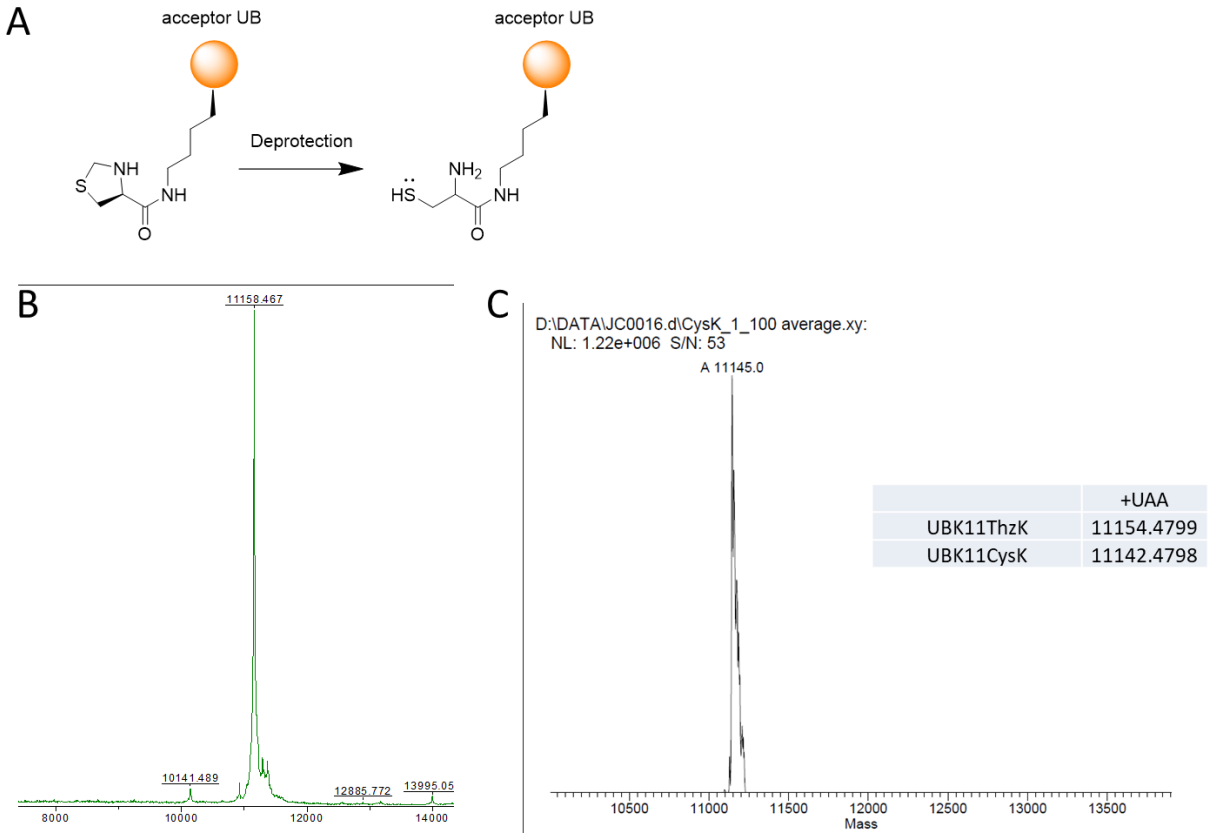


Figure 2-11 Conversion from UBThzK into UBCysK

(A) Scheme for deprotection reaction. (B) HPLC chromatogram for purifying UBCysK, retention time: 16-17 min (C) MALDI-LP shows the molecular weight of UBThzK 11158 (calculated 11155) (D) ESI shows the molecular weight of UBCycK 11145 (calculated 11142)

that rapidly reacts to form an oxime, conveniently releasing the captured peptide or protein. [149, 150] Therefore, methoxyamine is extensively used in SPPS for protein synthesis when several peptides need to be ligated together sequentially, such as erythropoietin[151] and modified histones[152, 153]. Detailed protocol is described in Chapter 2.6. After the reaction, the molecular weight reduces 12 Da, indicating the full conversion from UBThzK into UBCysK. (Figure 2-11)

To confirm the full conversion of UBThzK into UBCysK, a labeling assay utilizing a thiol-reactive reagent, Biotin-Maleimide was performed. Theoretically, UBCysK after the deprotection can react with Biotin-Maleimide whereas UBThzK cannot. Indeed, the molecular weight increase detected by MALDI-LP was expected. (Figure 2-12)

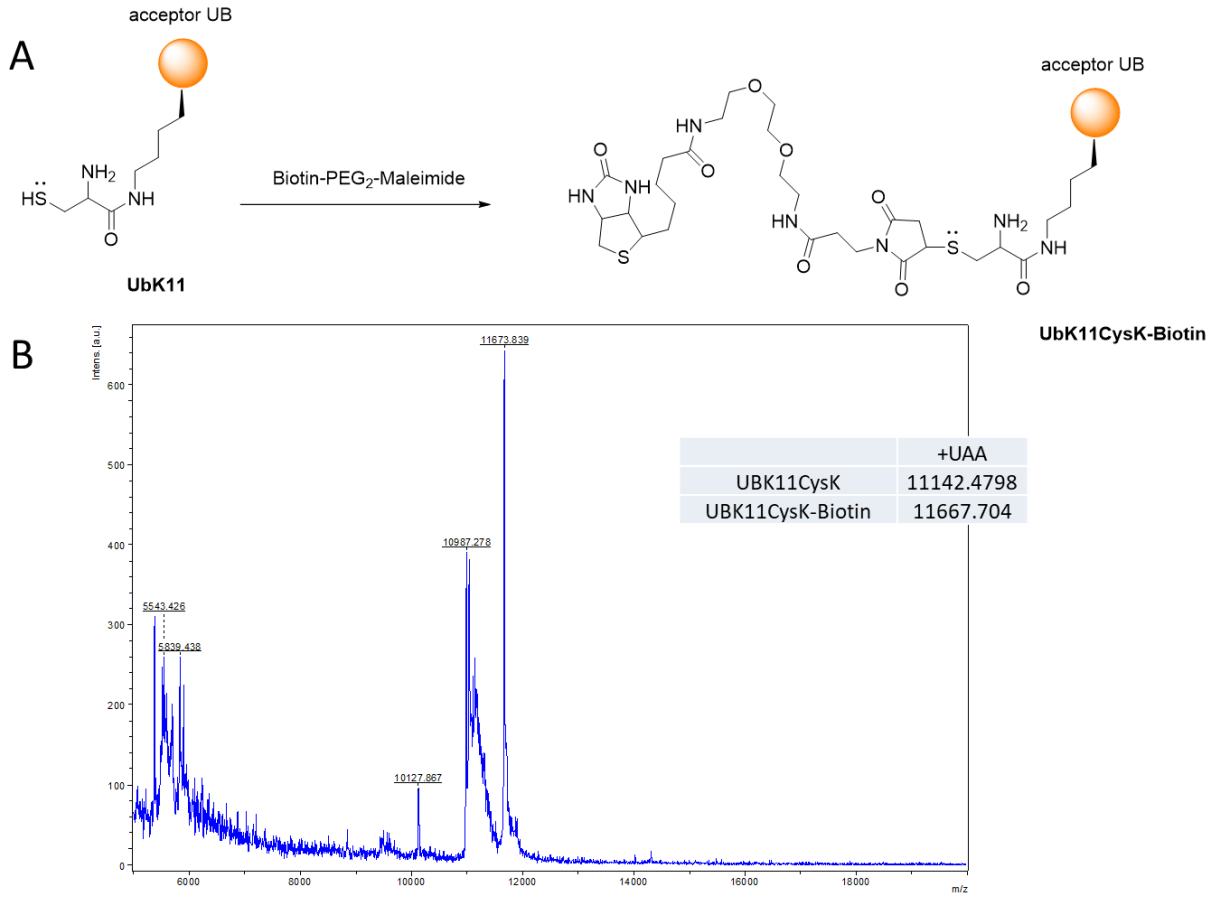


Figure 2-12 Biotin labeling assay for UBCysK

(A) UBCysK after the deprotection reveal a thiol group that can react with Biotin-Maleimide. (B) MALDI-LP shows the correct molecular weight of Biotin conjugated UB 11673. (calculated 11667)

2.5.2 Construct donor UB: generating a reactive C-terminus

The intein carries out protein slicing, a unique auto-processing event where it excises itself out from a larger precursor polypeptide through the cleavage of two peptide bonds. This rearrangement usually occurs post-translationally for protein maturation.[154, 155] In addition, intein-mediated protein splicing is spontaneous; it requires no external factor or energy source, only the folding of the intein domain.

Hijacking this natural feature, intein was utilized by protein chemists intensively centered around the cleavage and/or formation of peptide bonds.[156, 157] Among them, the most frequently used ones are intein-facilitate tag-free protein expression and purification[158], *in vitro* protein synthesis like expressed protein ligation. (EPL)[159]. (Figure 2-13) Ubiquitin probes bearing an electrophilic C-terminus for DUB profiling described in Chapter 1.4.2 (Scheme 1-4) was generated based upon intein-mediated EPL.

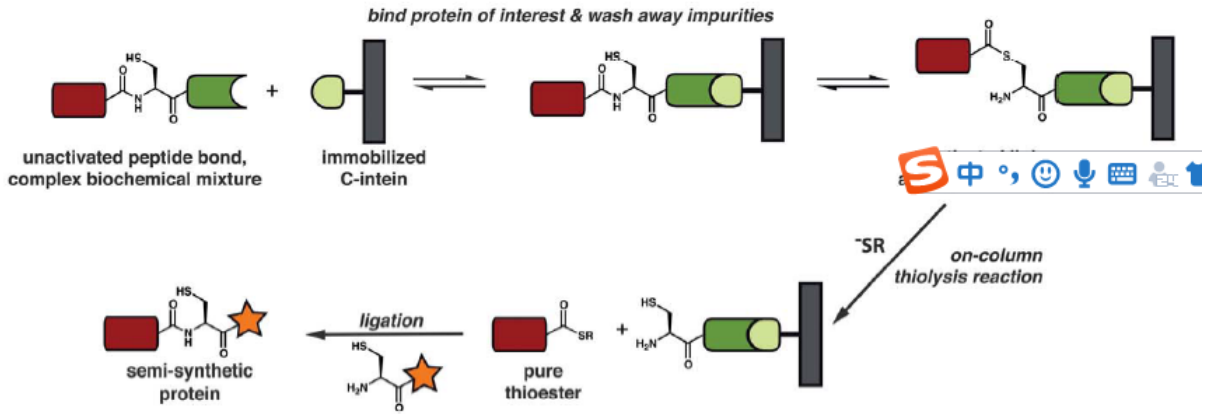
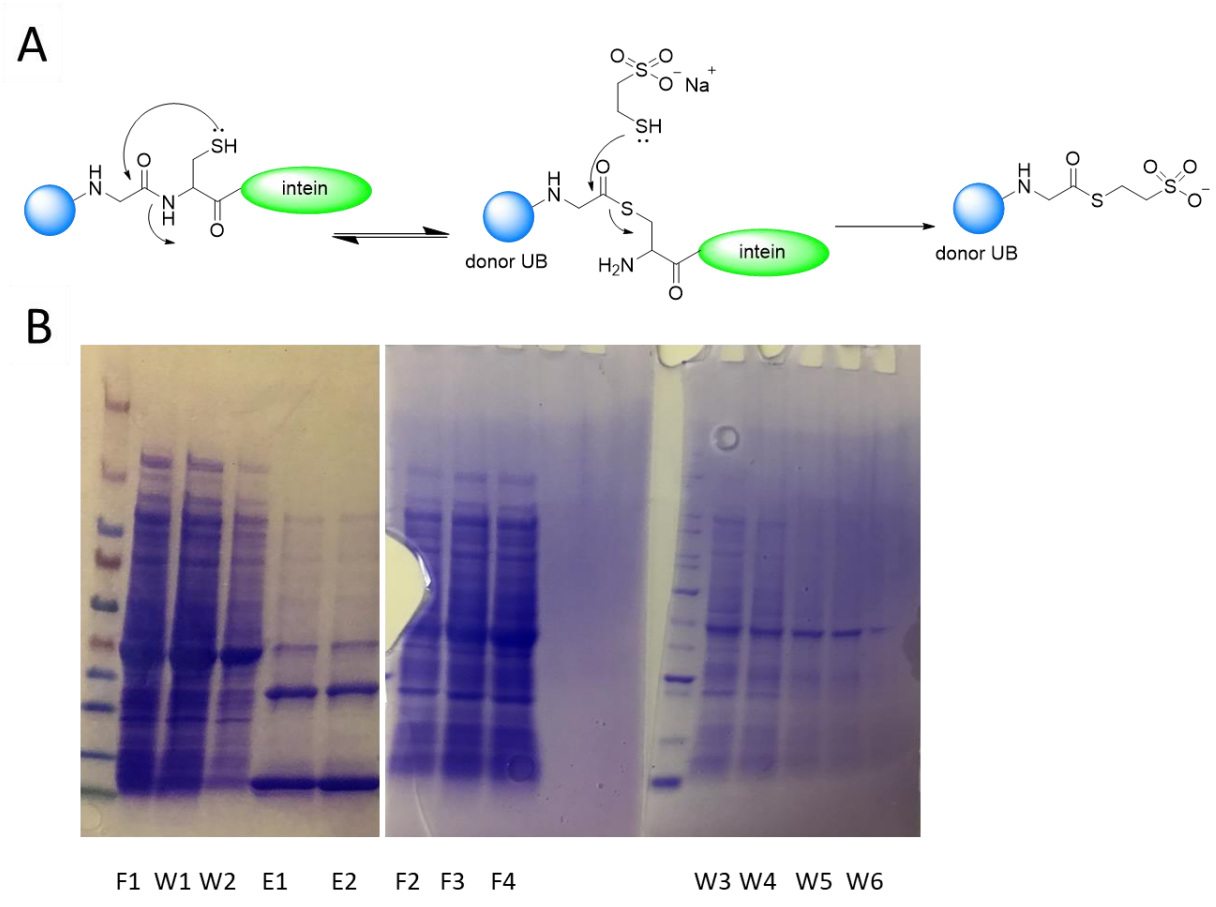


Figure 2-13 Protein modification using streamlined EPL.

Following the same strategy, tagged-free donor Ub with a thioester C-terminus was generated as such: ubiquitin-intein fusion protein was expressed in *E. Coli* and purified with chitin beads which has strong affinity to intein tag. After washing out unbound impurities in cell lysate, the ubiquitin was eluted out with 200 mM MesNa to generate Ub~MesNa, a C-terminus activated ubiquitin to facilitate native chemical ligation. Elution with DTT yields ubiquitin with a native C-terminus. The detail protocol derived from a reported streamlined EPL procedure by Muir lab[159] is descried in Chapter 2.6. However, due to the low binding capacity of chitin beads compared to the scale of protein expression (2L), remaining UB-Intein fusion protein was found in flow-through, wash1 and wash2 buffers for the chitin affinity purification; as a result, another round of



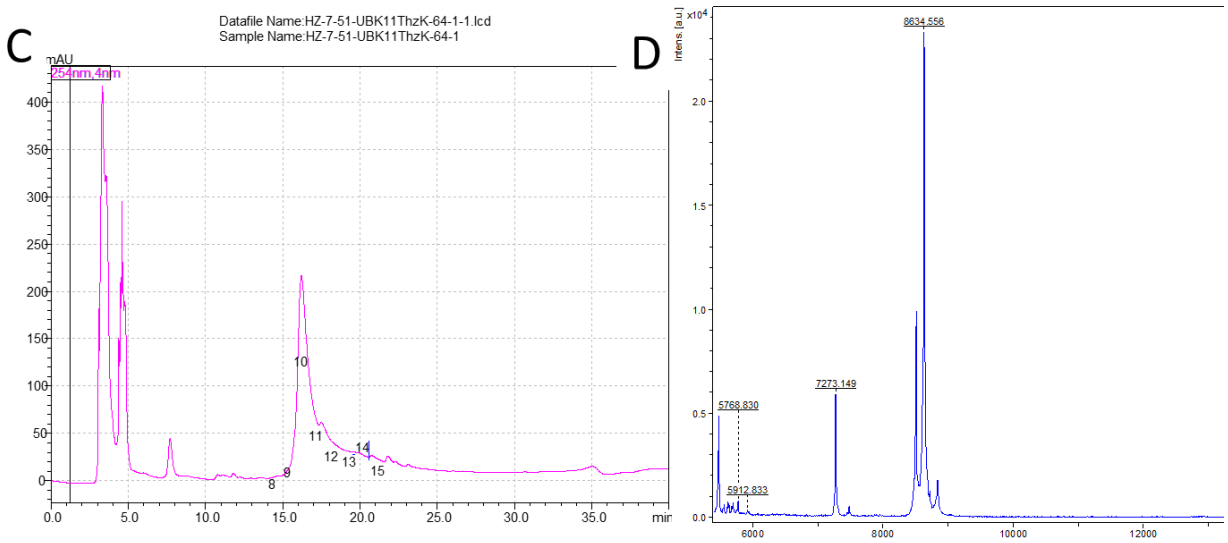


Figure-2-14 Donor UB~MesNa formation

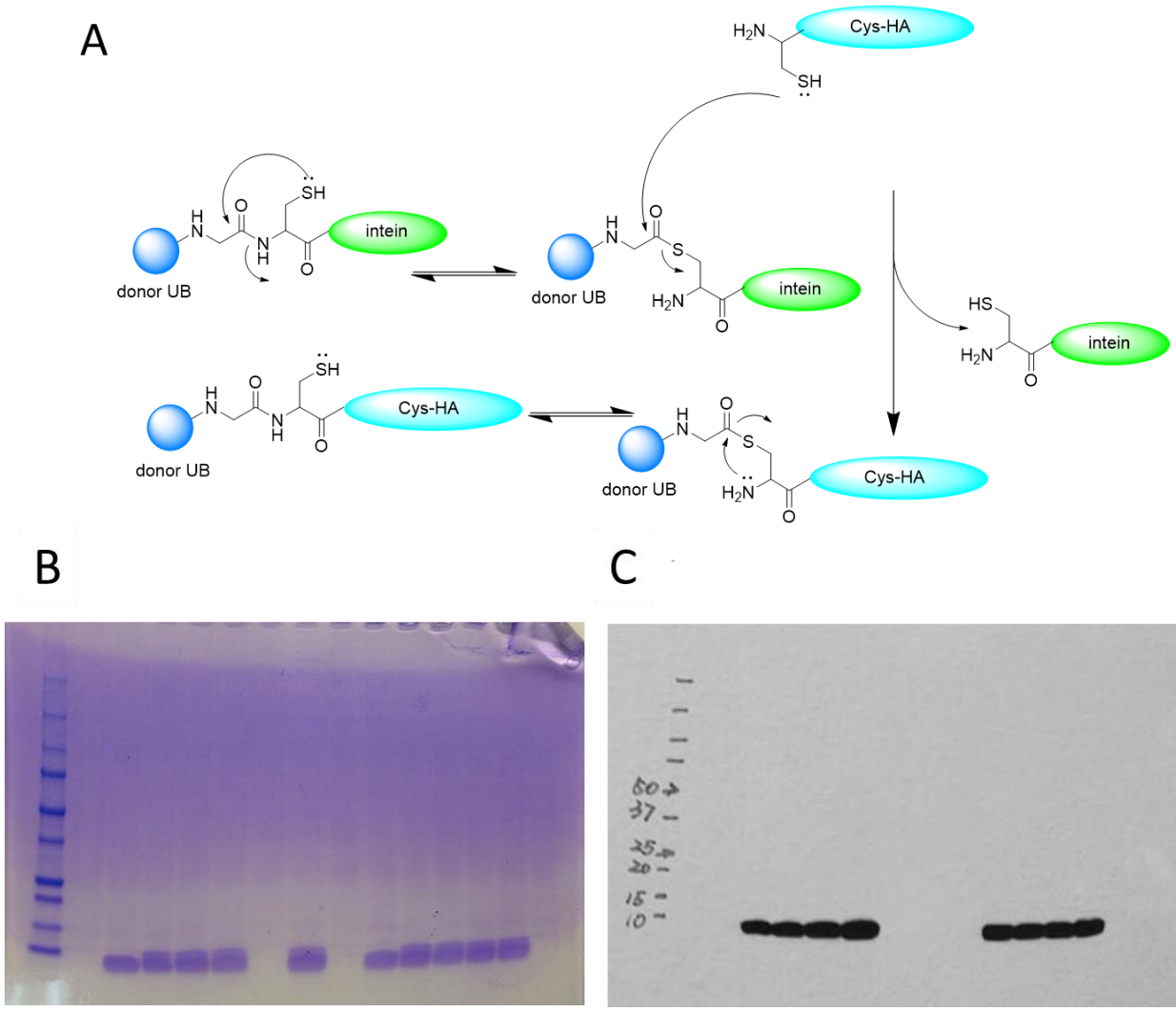
(A) Mechanism for UB~MesNa formation mediated by intein. (B) Coomassie gel for chitin affinity purification: F1,W1,W2,E1,E2 Flow through, Wash buffer washes and elutions for the first-round binding F2,F3,F4,W3,W4,W5,W6 Flow through, Wash buffer washes for the second-round binding (C) HPLC chromatogram for purifying UB~MesNa with retention time: 16-17 min (D) MALDI-LP shows the molecular weight of UB~MesNa 8634.38 (calculated molecular weight 8632.0)

purification was performed to immobilize all UB-Intein fusion protein upon MesNa cleavage. (Figure 2-14)

2.5.3 Build up ubiquitin chains: native chemical ligation

Native chemical ligation reaction (NCL) involves reacting a C-terminal peptide thioester with an N-terminal cysteinyl peptide to produce a native peptide bond between the two fragments.[160, 161] This reaction has considerably extended the size of polypeptides and proteins that can be produced by total synthesis and has also numerous applications in bioconjugation, polymer synthesis, material science, and micro- and nanotechnology research.[162, 163] Notably, an N,S-acyl shift is the key note in NCL featuring an intramolecular rearrangement. (Scheme 2-1)

The targeted linkage-specific di-UB can be conjugated together with native chemical ligation with both acceptor ubiquitin and donor ubiquitin available. Firstly, a pilot reaction between Ub~MesNa and a synthetic HA peptide featuring an N-terminal Cysteine was performed. (Figure 2-15A) If the native chemical ligation works, the donor UB can be labelled with HA tag therefore can be detected on western blot blotting on HA antibody. As a result, in a panel of ligation reactions increasing the input of the Cys-HA peptide, UB~MesNa could be labeled successfully showing on Coomassie and Western blot (Immunoblotting: HA) (Figure 2-15B and C). Nevertheless, the molecular weight of labelled ubiquitin with a C-terminal HA tag could be detect by MALDI-LP, which corresponded to the calculated one (Figure 2-15D).



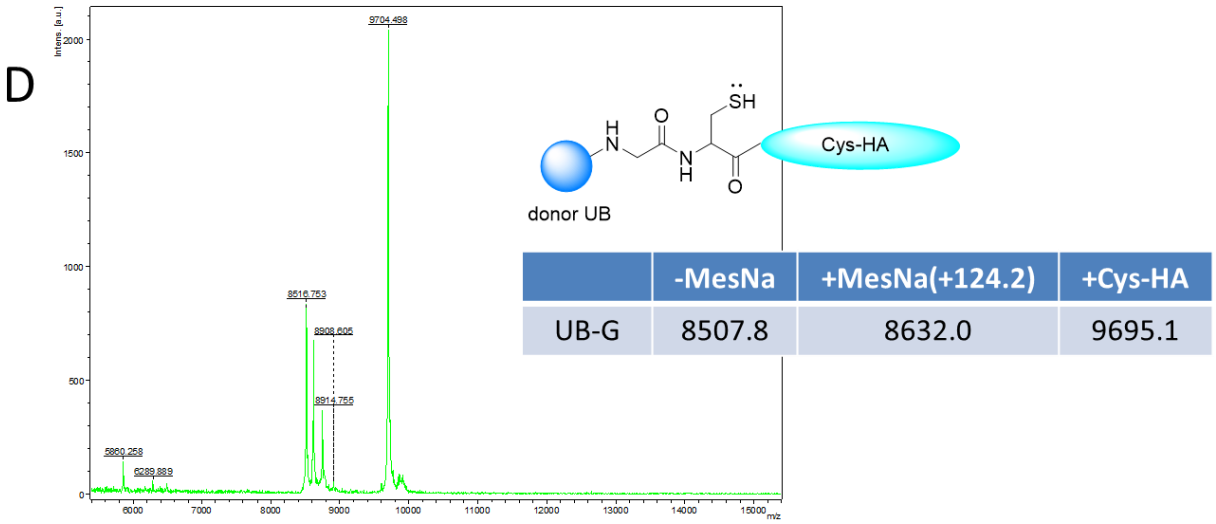
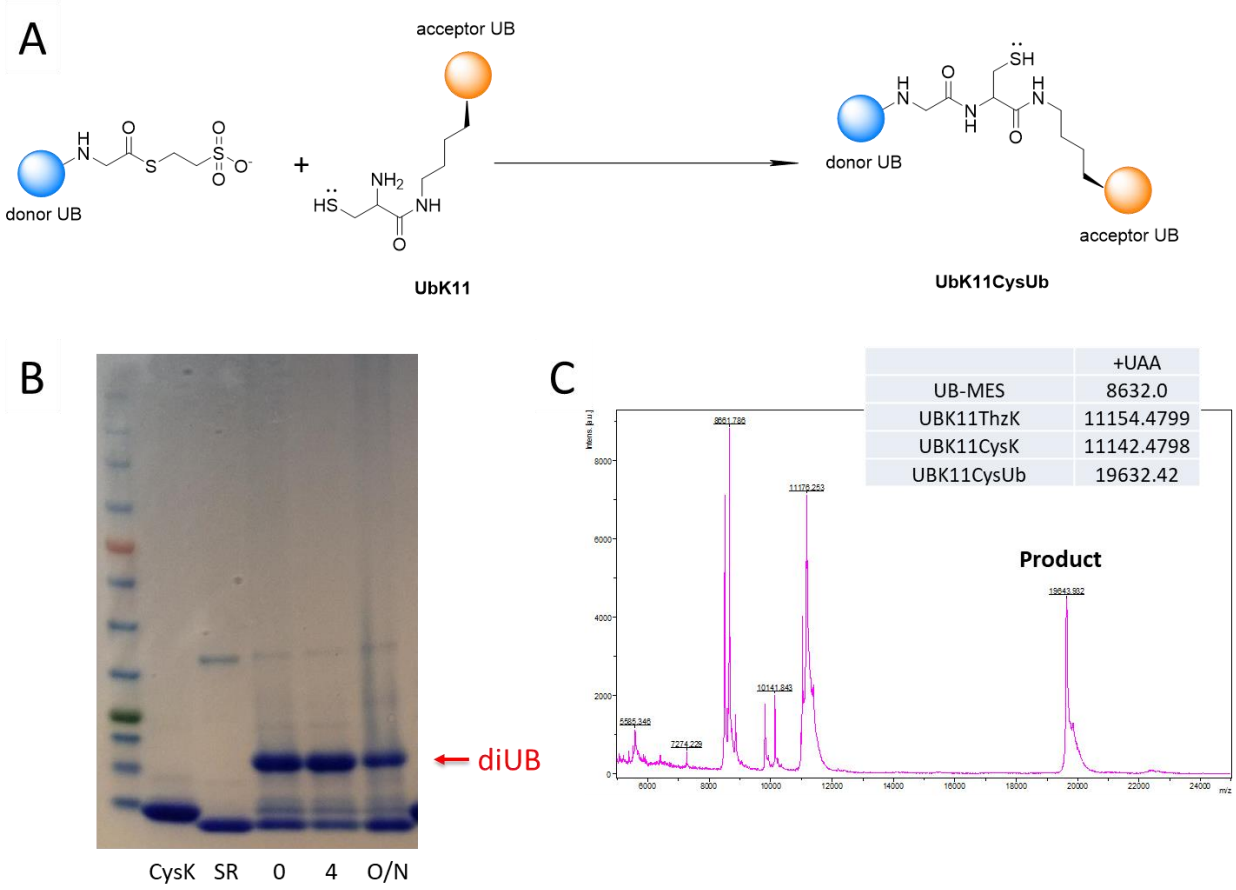


Figure 2-15 NCL between UB~MesNa and Cys-HA peptide

(A) Reaction between UB~MesNa and Cys-HA peptide yields C-termial HA-tagged UB after N, S-acyl shift.
(B) (C) Coomassie and Western blot (Immunoblotting HA) shows the reaction results: 3A1 Cys-HA peptide only, 3B1 UB~MesNa only, 3C1 to 3 E1 1,1.5,2 equivalency of Cys-HA peptide added; the right panel is a duplicate.
(D) MALDI-LP shows the correct molecular weight of UB-HA: 9704. (calculated molecular weight: 9695)

Now that the full conversion from UBThzK to UBCysK and the reactivity of Ub~MesNa towards NCL were both confirmed by two aforementioned pilot reactions, the reaction between UBCysK and UB~MesNa was investigated. (Scheme 2-1 and Figure 2-16A) To optimize the conditions for NCL, the reaction was monitored in a time-course manner and the product diUB was detected in the Coomassie gel. Additionally, the acceptor UB prepared from GCE features a His×6 tag followed by an HA tag, can be detected with Western blot immunoblotting on HA, whereas donor UB, made from intein-mediated EPL, is tagged-free. Interestingly, the NCL reaction between engineered donor UB and acceptor UB is very fast (Figure 2-16B): at 0 hour, some diUB could be detected while at 2-hour time point, the reaction was complete. The fast reaction manner can be explained by the high concentration of both components (1-2 mM) to facilitate an effective nucleophilic attack by the thiol on UBCysK to the C-terminus of Ub~MesNa,

which is the rate-determine step in NCL; the intramolecular rearrangement of N, S- shit is spontaneous and fast. Nevertheless, to fully-consume acceptor UBCysK and drive the reaction to complete, another equivalent of UB~MesNa was added at 2-hour time point. After 4 hours, the total depletion of acceptor UB was observed on Coomassie gel (Figure 2-16B); the trace amount of acceptor UB band on gel could be due to the misincorporation of UAA with a wide-type lysine incorporated and no reactivity towards NCL. Lastly, a couple of mass spectrometry approach were used to characterize prepared diUB. The diUB in the ligation reaction mixture was detect by MALDI-LP to show the correct total molecular weight; whereas the linkage specificity was implicated by an in-gel digestion followed by LC-MS analysis (Figure 2-16C and 16D). After trypsin digestion, a branched peptide TLTG**K**(CG)TITLEVEPSDTIENVK should be expected



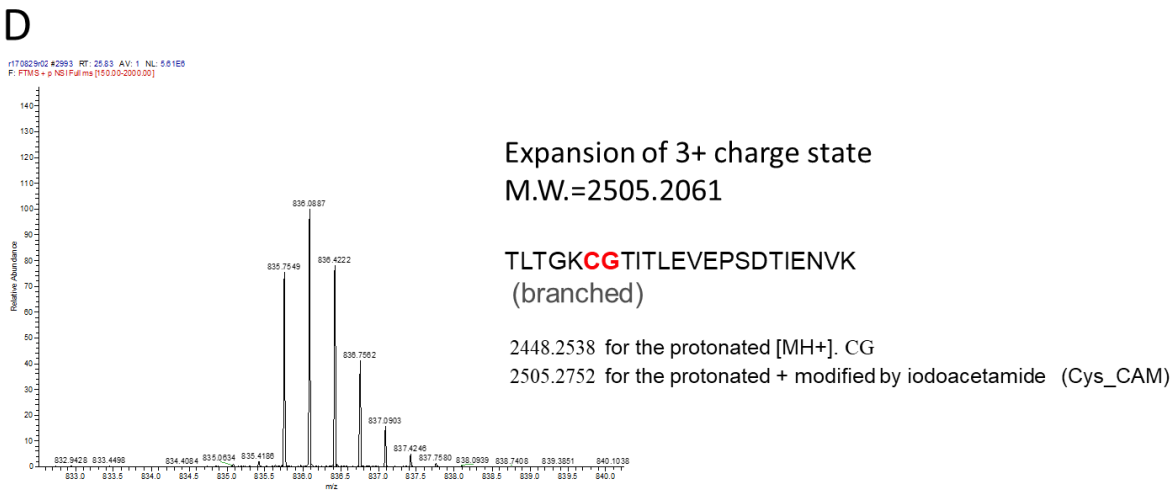


Figure 2-16 NCL between acceptor UBCysK and donor UB~MesNa

(A) Reaction between UBCysK and UB~MesNa yields chain-specific diUB (K11-diUB as an example)
(B) Coomassie gel of NCL reaction in a time-course manner; CysK: acceptor UBK11CysK (11.5kD) SR: donor B~MesNa (8.5kD) 0, 2, 4, o/n: reaction mixture at 0, 2, 4 hour and overnight. Formation of diUB (19.5KD)
(C) MALDI-LP shows the correct molecular weight of diUB 19643 (calculated molecular weight 19632)
(D) LC-MS shows the correct peptide molecular weight after trypsin digest. The molecular weight of a branched peptide with Lys11 conjugated to a GlyCys motif through isopeptide bond is detected as 2505.2061; the calculated molecular weight of this peptide is 2505.2752.

where K represents the Lys11 and (CG) stands for the isopeptide modification on Lys11 position which indicated K11-linkage specificity.

Notably, excess amount (2.0 to 3.0 equivalent) of donor UB over acceptor UB was used. A couple of reasons explain the optimized stoichiometry of this reaction. Acceptor UB prepared through GCE requires several steps of UAA synthesis, protein expression and purification; thus, consuming acceptor UB by adding more donor UB is overall efficient in large amount of diUB production. In addition, the excess amount of tag-free donor UB can be easily removed through Ni-NTA affinity purification (Figure 2-17) Since acceptor UB has an N-terminal His and HA tag, diUB with little unreactive mono acceptor can be enriched in the elute of Ni-NTA purification and prone to further purification.

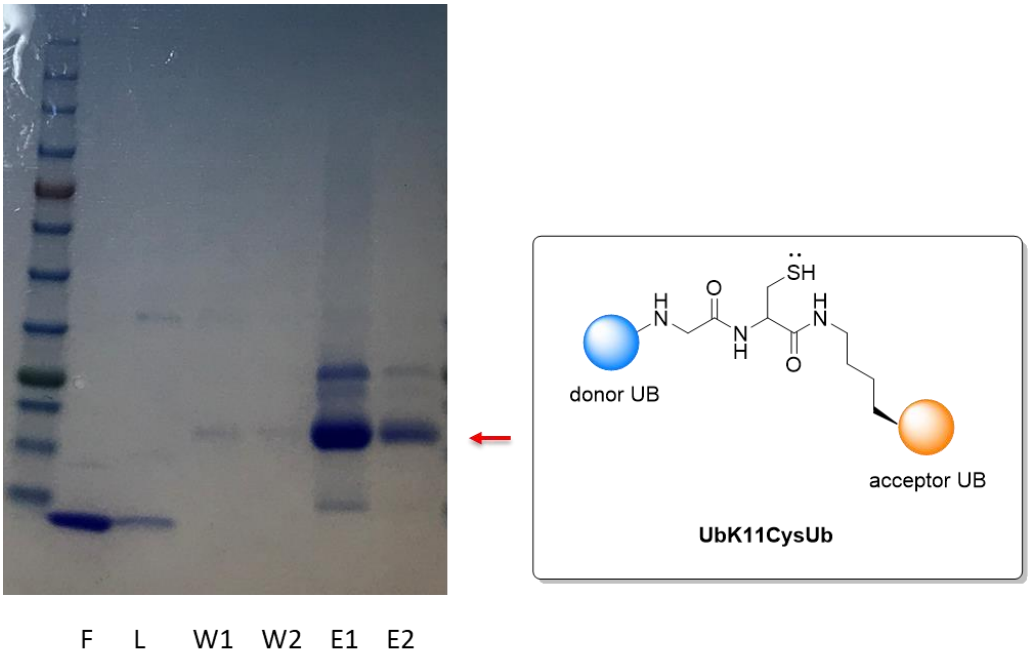
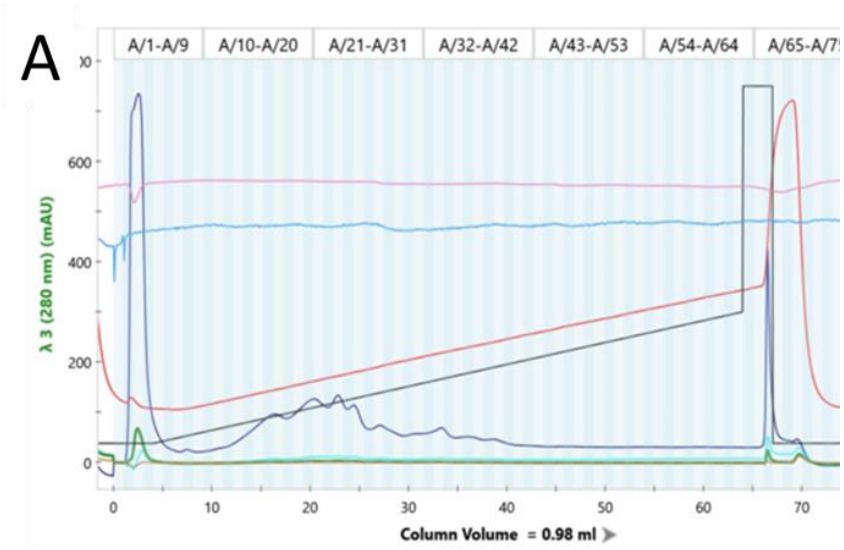
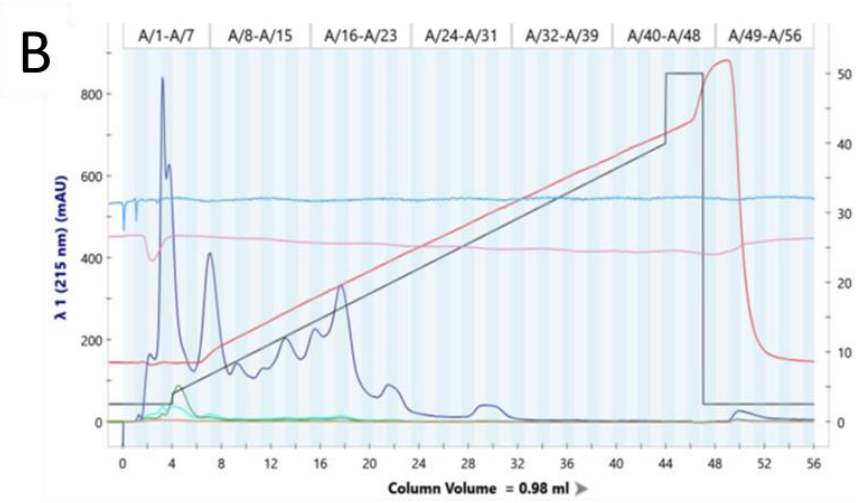


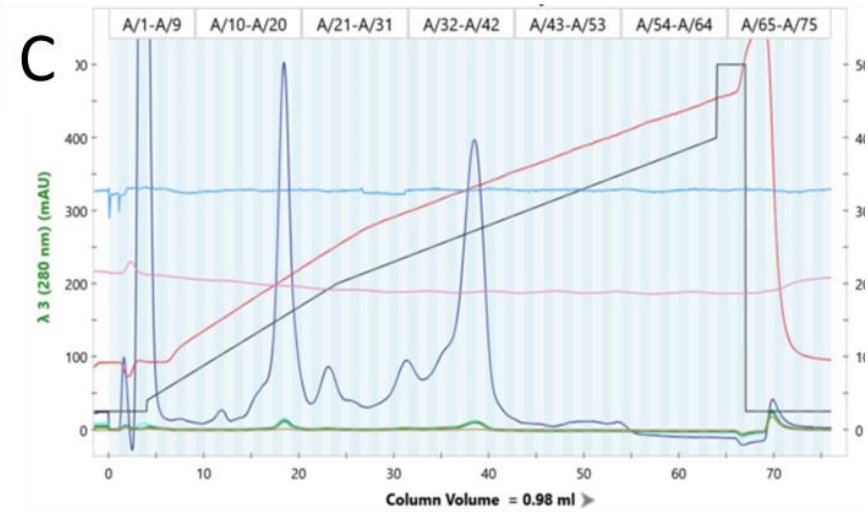
Figure 2-17 Ni-NTA affinity purification of diUB

F: Flow-through, monomeric donor UB in excess
L: Lysis buffer wash
W1, W2: Wash buffer wash
E1, E2 Elution buffer wash, enriched diUB with little monomeric acceptor UB and a higher-molecular impurity band

After Ni-NTA affinity purification to remove donor UB in excess, additional step is needed to further polish the purity of diUB. At this point, size-exclusion, cation/anion-exchange chromatography were all investigated for this purpose. Ubiquitin is a very small protein; dimeric UB is only 9 kD bigger than monomeric UB. As a result, size-exclusion chromatography was not very effective to separate them (Results are not shown). Nevertheless, cation-exchange prevails anion-exchange in separating diUB from other impurities because ubiquitin featuring several positively charged amino acid residues such as lysine or arginine charges better in cation forms (Figure 2-18A). In addition, a pH-scouting approach is used to further optimize the condition suitable for diUB separation (Figure 2-18B). In conclusion, diUB can be purified most with anion-exchange chromatography with a 10-75% salt gradient at pH 5.0 (Figure 2-18C and D).


A
A/1-A/9
A/10-A/20
A/21-A/31
A/32-A/42
A/43-A/53
A/54-A/64
λ 3 (280 nm) (mAU)
Column Volume = 0.98 ml


B
A/1-A/7
A/8-A/15
A/16-A/23
A/24-A/31
A/32-A/39
A/40-A/48
A/49-A/56
λ 1 (215 nm) (mAU)
Column Volume = 0.98 ml


C
A/1-A/9
A/10-A/20
A/21-A/31
A/32-A/42
A/43-A/53
A/54-A/64
A/65-A/75
λ 3 (280 nm) (mAU)
Column Volume = 0.98 ml

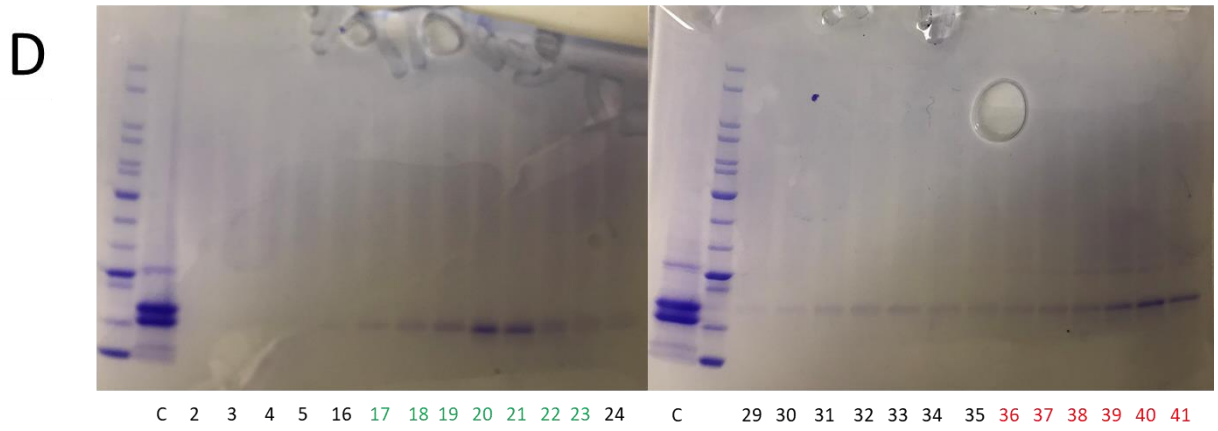


Figure 2-18 Ion-exchange chromatography for diUB purification

(A) Anion-exchange chromatography (pH=8.0) gives little separation of diUB from other impurities.
(B) Cation-exchange chromatography (pH=6.0) gives little separation of diUB from other impurities.
(C) Cation-exchange chromatography (pH=5.0) gives good separation of diUB, the peak with retention time 37 min was the correct one containing diUB. Gradient: 8-40% in 20 min following by 40-80% in 40 min
(D) Coomassie gel staining showing the separation of cation-exchange chromatography (C), which suggested the fraction 36-41 were the right ones containing diUB.

2.6 Conclusion

To summarize this chapter, DiUB with defined linkage was prepared through native chemical ligation between an acceptor UB incorporating the 1-amino-2-thiol functionality and a donor UB featuring a reactive thiolester at C-terminus. To this end, acceptor UB and donor UB were made in parallel: acceptor UB was generated with genetic code expansion approach to site-specifically incorporate thiazolidine lysine containing a protected 1-amino-2-thiol moiety whereas donor UB was produced with intein-mediated expressed protein ligation.

In a long synthetic route shown in Scheme 2-1, yield in each step is very crucial to maintain a streamline of diUB production for whatever purpose. Thereby, tremendous endeavor had been made to improve the yield for each step and eventually methylated Thiazolidine lysine (ThzK-OMe) can be obtained and purified with HPLC in gram-level. This facilitate the methylated version of UAA incorporation with a yield of 40-60 mg/1L media culture. After deprotection, HPLC

purification and lyophilization, 15-20 mg of lyophilized UBCysK was collected from the same prep subject to native chemical ligation. Nevertheless, two rounds of binding and elution with chitin beads ensures a massive production of UB~MesNa after HPLC and lyophilization. Lastly, 15-20 mg of UBCysK will generate around 15 mg diUB after affinity purification, thus around 3-5 mg diUB after cation-exchange. This streamline of diUB production renders it possible for large diUB probe synthesis for activity-base protein profiling and for diUB-E2/E3 conjugates formation to resolve their structures.

2.7 Material and Methods

2.7.1 Unnatural amino acids synthesis

Synthesis of methyl N^6-((benzyloxy)carbonyl)-N^2-(tert-butoxycarbonyl)-L-Lysinate (**10**)

Synthesis of ThzKOH and ThzKOMe was based upon previous reported protocols with some modifications[125, 164, 165]. To a 250 mL round-bottom flask, Boc-Lys(Cbz)-OH (3, 7.23g, 19.0 mmol) was added then dissolved with DMF 40 mL, followed by potassium carbonate (5.25g, 38.0 mmol). The solution was cooled to 0o C and iodomethane (1.05 ml, 24.7 mmol) was added dropwise. The temperature was allowed to raise to room temperature and the reaction was stirred under N_2 for overnight. The reaction mixture was analyzed with thin layer chromatography (Hexane: EtOAc = 1:1, Rf = 0.6), then dissolved in ethyl acetate (50 ml) and extracted twice using water (100 mL) to remove excess DMF. The organic layer was then extracted using brine (100 mL) and the combined organic layers were concentrated to dryness and purified using column chromatography (Hexane: EtOAc = 1:1). The resulting oil is 6.5 g with a yield of 87%. NMR analysis is in agreement with previous reports. ^{1}H (400 MHz, CDCl3): δ = 1.37-1.59 (13H, m, includes Boc, Hγ, Hδ), 1.62-1.90 (2H, m,Hβ), 3.17-3.21 (2H, q, J = 6.3, CH2NHCbz), 3.63 (3H,

s, CO2Me), 4.09-4.14 (1H, m, Hα). 5.02(2H, s, CH2Ph), 5.09-5.14 (2H, m, NHBoc and NHCbz), 7.26-7.36 (5H, m, CHAr). MS: m/z (ESI+): 295.16 $[M+H\text{-}Boc]^+$, calculated MS: 295.17.

Synthesis of methyl (tert-butoxycarbonyl)-L-Lysinate (**11**)

Methyl N6-((benzyloxy)carbonyl)-N2-(tert-butoxycarbonyl)-L-Lysinate (**10**, 1.0 eq) was dissolved in ethyl acetate (40 mL) and ethanol (10 ml) to which palladium on carbon (10%, 0.39 mmol) was added and hydrogen was bubbled through the solution three consecutive times to promote efficient reduction. Completion of the reaction was monitored using TLC analysis (Hexane: EtOAc = 1:1) showing the disappearance of starting material. A final ninhydrin stain (Dichloromethane: Methanol = 9:1) confirmed reduction of the carboxybenzyl group to a free amine. The palladium was removed by filtration through a pad of Celite and the product was concentrated to dryness and the resulting crude oil 4.2g, 98% (crude yield) was used for the next step without further purification. ^{1}H (400 MHz, CDCl3) : δ = 1.46-1.84 (15H, m, contains Boc, Hγ, Hδ, and NH2), 1.57-1.89 (2H, m, Hβ), 2.66 (2H, t, J = 6.7, CH2NH2), 3.79 (3H, s, CO2Me), 4.35-4.89 (1H, m, Hα), 5.63 (1H, d, J = 7.6, NHBoc). MS: m/z (ESI+): 261.06 $[M+H]^+$, calculated MS: 261.18.

Synthesis of methyl N^{ε}-Boc-L-thiazolidine-N-Boc-L-Lysinate (**13a**)

N-boc-L-thiazolidine-4-carboxylic acid (1.0 eq.) was dissolved in dichloromethane (40ml). Next, 4-Dimethylaminopyridine (DMAP, 0.5 eq.) was added followed by subsequent addition of 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC, 1.2 eq.). After an hour of stirring at 0^o C, Boc-L-Lysine Methylester (previously dissolved in 5 mL DCM) was added dropwise. Completion of the reaction was monitored using TLC analysis (Hexane: EtOAc = 1:1, iodine stain, Rf = 0.4). The mixture was diluted with dichloromethane (50 mL) and extracted with 1M HCl (50 ml) and brine (50 ml). The organic fractions were dried with magnesium sulfate and concentrated to

dryness to give the crude product. The product was purified using column chromatography hexane: ethyl acetate = 3:1 to yield white foam 3.7 g, 48%. ^{1}H (400 MHz, CDCl3): δ = 4.67 (1H, t, J=6.7, αCH), 4.85 (1H, d, J=10.2, CHaHb), 4.33 (1H, d, J=10.2, CHaHb), 3.93 (1H, t, J=6.7, αCH), 3.54-3.24 (1H, m, CHa'Hb'), 3.19-3.15 (3H, m, CH2 and CHaHb'), 1.76-1.70 (2H, m, CH2), 1.45-1.42 (2H, m, CH2), 1.42-1.25 (20H, m, including Boc and CH2). MS: m/z (ESI+): 498.20 $[M+Na]^{+}$, calculated MS: 498.22.

Synthesis of methyl N$^{\varepsilon}$-Boc-D-thiazolidine-N-Boc-L-Lysinate (**13b**)

Same as **13a** except starting with N-boc-D- thiazolidine-4-carboxylic acid. Yield: 3.4g, 44%.^{1}H (400 MHz, CDCl3): δ = 4.66 (1H, t, J=6.7, αCH), 4.75 (1H, d, J=10.3, CHaHb), 4.34 (1H, d, J=10.3, CHaHb), 3.91 (1H, t, J=6.7, αCH), 3.42-3.20 (1H, m, CHa'Hb'), 3.21-3.11 (3H, m, CH2 and CHaHb'), 1.64-1.60 (2H, m, CH2), 1.45-1.42 (2H, m, CH2), 1.41-1.21 (20H, m, including Boc and CH2). m/z (ESI+): 498.20 $[M+Na]^{+}$. calculated MS: 498.22.

Synthesis of Nε-Boc-L-thiazolidine-N-Boc-L-Lysine (**14a**)

Methyl N-Boc-L-thiazolidine-N-Boc-L-Lysinate **13a** (10.5 g, 22.0 mmol, 1 eq) was dissolved in THF:H2O (3:1, 150 mL); $LiOH{\cdot}H_2O$ (1.9 g, 44.0 mmol, 2 eq) was added in room temperature and the reaction was stirred overnight. Completion of the reaction was monitored using TLC analysis (Hexane: EtOAc = 1:1, iodine stain, Rf = 0.2). The reaction mixture was then diluted with EtOAc (50 mL) and water (50 mL). The aqueous layer was acidified to pH 4 using 5% aqueous citric acid solution and then extracted with EtOAc (~ 50 mL). The combined organic fractions were washed with brine, dried (Na_2SO4) and filtered. The filtrate was concentrated under reduced pressure, to give N$^{\varepsilon}$-Boc-thiazolidine-N-Boc-L-Lysine as a colourless gum 9.27g, 91%. The crude product was used for deprotection without further purification. m/z (ESI+): 462.30 $[M+H]^{+}$, calculated MS: 462.23.

Synthesis of Nε-Boc-D-thiazolidine-N-Boc-L-Lysine (**14b**)

Same as **14a** except starting with N$^{\varepsilon}$-Boc-D-thiazolidine-N-Boc-L-Lysinate (**13b**). Yield: 8.71g, 86%. m/z (ESI+): 498.08 $[M+Na]^+$, calculated MS: 498.22.

Synthesis of N$^{\varepsilon}$-L-thiazolidine-L-Lysine (**7a**)

N$^{\varepsilon}$-Boc-L-thiazolidine-N-Boc-L-Lysine (**14a**, 1 eq.) was dissolved in 1,2-dioxane and stirred at room temperature. 4M hydrogen chloride in dioxane was added dropwise and the reaction was allowed to stir overnight, which resulted in formation of a white precipitate. Excess solvent was removed through evaporation and the crude was further purified using ether precipitation (3700 rpm, 10 min, 4^{o} C). The crude product was further purified with RP-HPLC (Phenomenex Preparative Jupiter 10□ Proteo 90Å column, 10 mL/min) 0-30% (H2O/Acetonitrile) in 30 min. Retention time: 7-8 min. The fractions containing product were collected and lyophilized to dryness. 2.13 g, 40.6%. ^{1}H (400 MHz, D2O) δ = 4.55 (1H, t, J=6.9, αCH), 4.39 (1H, d, J=10.3, CHaHb), 4.29 (1H, d, J=10.3, CHaHb), 3.91 (1H, t, J=6.0, αCH), 3.58-3.26 (1H, m, CHa'Hb'), 3.21-3.13 (3H, m, CH2 and CHaHb'), 1.86-1.79 (2H, m, CH2), 1.49-1.44 (2H, m, CH2), 1.44-1.25 (2H, m, CH2) m/z (ESI+): 262.10 [M+H]+, calculated MS: 262.12(Figures S1).

Synthesis of N$^{\varepsilon}$-D-thiazolidine-L-Lysine (**8a**)

Same as **7a** except starting with N$^{\varepsilon}$-Boc-L-thiazolidine-N-Boc-L-Lysine (**14b**). Yield: 1.84g, 37.2%. ^{1}H (400 MHz, D2O) δ= 4.53 (1H, t, J=6.9, αCH), 4.62 (1H, d, J=10.1, CHaHb), 4.38(1H, d, J=10.1, CHaHb), 3.76 (1H, t, J=6.0, αCH), 3.58-3.26 (1H, m, CHa'Hb'), 3.20-3.14 (3H, m, CH2 and CHaHb'), 1.76-1.65 (2H, m, CH2), 1.54-1.33 (2H, m, CH2), 1.33-1.25 (2H, m, CH2) m/z (ESI+): 262.10 $[M+H]^+$, calculated MS: 262.12. (Figures S2).

Synthesis of methyl Nε-L-thiazolidine-L-Lysinate (**7b**)

Same as **7a** except starting with N$^{\varepsilon}$-Boc-L-thiazolidine-N-Boc-L-Lysinate (**14a**). Yield: 1.21 g, 57%. ^{1}H (400 MHz, D2O) δ= 4.51 (1H, t, J=6.9, αCH), 4.65 (1H, d, J=10.3, CHaHb), 4.40 (1H, d, J=10.3, CHaHb), 3.84 (1H, t, J=6.0, αCH), 3.77 (3H, s, OCH3), 3.60-3.28 (1H, m, CHa'Hb'), 3.22-3.15 (3H, m, CH2 and CHaHb'), 1.82-1.64 (2H, m, CH2), 1.60-1.43 (2H, m, CH2), 1.43-1.31 (2H, m, CH2) m/z (ESI+): 276.0 [M+H] +, calculated MS: 276.14. (Figures S3).

Synthesis of methyl N$^{\varepsilon}$-D-thiazolidine-L-Lysinate (**8b**)

Same as **7a** except starting with Nε-Boc-D-thiazolidine-N-Boc-L-Lysinate (**14b**). Yield: 1.45 g, 74%. δ= H (400 MHz, D2O) 4.53 (1H, t, J=6.9, αCH), 4.57 (1H, d, J=10.3, CHaHb), 4.47 (1H, d, J=10.3, CHaHb), 3.75 (3H, s, OCH3), 3.54 (1H, t, J=6.0, αCH), 3.67-3.33 (1H, m, CHa'Hb'), 3.33-3.20 (3H, m, CH2 and CHaHb'), 1.79-1.64 (2H, m, CH2), 1.61-1.42 (2H, m, CH2), 1.42-1.29 (2H, m, CH2) m/z (ESI+): 276.0 $[M+H]^{+}$, calculated MS: 276.14. (Figures S4).

2.7.2 Protein expression and purification of Ubiquitin K11ThzK

Protein expression was based upon reported protocols with some modifications[125, 166].10 mL overnight culture of *E.coli* DH10B cells carrying either pBK-ThzKRS or pBk-MmPylRS and pMyoUbK11TAG pylT was inoculated into 2L 2XYT containing 100 μg/mL Ampcillin and 20 μg/mL tetracycline. The cells were incubated at 37 ºC in 220 RPM shaker. 1 mM of unnatural amino acid (**7a, 7b, 8a or 8b**) was dissolve in water and pH of the solution was adjusted by 2M NaOH to 8. When the OD of the culture reached 0.6, the UAA solution was added and the culture was allowed to grow at 37 °C for another half to one hour until the OD reached approximately 0.8. Then arabinose (0.02%, final concentration) and IPTG (1 mM final concentration, optional) were added and the culture was incubated in 30 ºC shaker (200 RPM) for overnight.

The cells were harvested with centrifugation (5,500 RPM, 30 min) and the pullets were resuspended with 50 mL NTA lysis buffer (50 mM phosphates, 300 mM NaCl, 10 mM imidazole,

pH=8.0) in which 1 mg/mL lysozyme, 100 μg/mL DNaseI, 1 mM PMSF, Roche protease inhibitor, 5 mM DTT were added. The resuspended cells were incubated on ice for at least half an hour before being sonicated at 4 ºC for 30 min. The extract was clarified by centrifugation (30 min, 12,000 RPM, 4 ºC). 2 mL of Ni-NTA beads (Qiagen) were added to the extract and the mixture was incubated with agitation for overnight at 4 ºC in a cold box. The beads were transferred to a gravity column and flow through (F) was collected. Then the beads were resuspended in 30 mL NTA lysis buffer (L), followed by 20 mL, 10 mL NTA wash buffer (50 mM phosphates, 300 mM NaCl, 20 mM imidazole, pH=8.0, W1 and W2). Finally, protein was eluted from the beads with NTA elution buffer (50 mM phosphates, 300 mM NaCl, 250 mM imidazole, pH=8.0, 5 mL each time for five times and collected as E1, E2, E3, E4 and E5). The protein concentration was measured by BSA assay (Bio-rad) and the yield of protein expression with each UAA incorporated was determined with three repeats of the same protocol and materials. The fractions in protein purification collected as F, L, W1, W2, E1, E2, E3, E4, E5 was analyzed by 4-15% SDD-PAGE and visualized with Coomassie staining.

2.7.3 SfGFP Assay of UAA incorporation

pPylsfGFPY151TAG was transformed into electrocompetent DH10B cells with pBkThzkRS, Mm pBkBocLys, and Mb pBkBocLys respectively. The cells were plated onto LB agar supplemented with 15 μg/ml tetracyclin and 50 μg/ml kanamycin (cells containing the pBkThzKRS plasmid was plated onto 100 μg/ml ampicillin instead of kanamycin). After overnight incubation at 37 ºC, a colony from each plate was inoculated into 5 mL LB supplemented with its respective antibiotics (Tet15 with either Kan50 or Amp100) and let grow overnight in 37 ºC shaker. 500 μL of the starter culture was inoculated into 20 mL of LB supplemented with its respective antibiotics and arabinose with a final concentration of 0.2 %w/v. The cultures were

allowed to grow to OD600 = 0.5, at which point 2.5 mL of each culture was aliquoted into culture tubes. A final concentration of 2 mM L-ThzK-OH, L-ThzK-OMe, D-ThzK-OH, D-ThzK-OMe, L-BocLys-OH, L-BocLys-OMe were added to its respective cultures in triplicates. The cultures were incubated in the 37 ºC shaker overnight to allow for protein expression. The cultures were harvested by spinning at 5,000 rpm for 10 minutes and washed twice with 2 mL PBS. The cells were pelleted again and resuspended in 2 mL PBS and 100 µL of each culture was spotted onto a Greiner Bio-One 96 well clear bottom plate. Florescence of the sfGFP were taken by excitation at 485 nm and emission at 510 nm using a SpectraMax M5 plate reader.

2.7.4 Transforming UBThzK into UBCysK

After protein expression and NTA affinity purification, UBThzK was dialyzed into dialysis buffer (25mM Tris, 50mM NaCl, 1mM DTT, pH=8.0) then lyophilized. Then lyophilized protein was rebuffered in 3-5 mL "rebuffer buffer" (6M Guanidium chloride, 200 mM phosphate buffer, pH=7.2). 400 mM methoxyamine was then added and pH was adjusted to 4.0; the resulting solution was incubated at 37ºC for 4-5h. The protein is then purified with HPLC (Semi-preparative C12 column at flow rate 4 mL/min, 15-80% Acetonitrile/water + 0.1% TFA). The fractions were analyzed by MALDI-LP or ESI-POS while the correct fractions were combined and lyophilized.

2.7.5 Intein-mediated donor UB preparation

E.Coli BL21(DE3) transformed with pTBX1-Ub(ΔG76)-Intein was grown in 1L LB containing 100 µg/mL ampicillin at 37°C until OD600 reached to 0.6[159]. Protein expression was induced by addition of 1 mM IPTG for overnight at 20°C. After harvesting the cells by centrifugation (8000g, 30 min), the cell pellets were resuspended by Chitin lysis buffer (20mM sodium phosphate, 200 mM NaCl, 1mM EDTA, 1mM TCEP, 20 µM PMSF, pH=7.2) and

sonicated on ice for 30 min to lysis the cells. The crude cell lysate was collected after centrifugation (12000RPM, 1 hour) to remove cell debris. Then regenerated chitin beads were added to the crude cell lysate and incubate at 4°C cold box overnight. The flow-through was collect next day and the chitin beads were washed twice with Chitin wash buffer (20mM sodium phosphate, 200 mM NaCl, 1mM EDTA,1mM TCEP, pH=7.2) until the Chitin elution buffer (20mM sodium phosphate, 200 mM NaCl, 1mM EDTA,10 mM TCEP, 200 mM MesNa, pH=7.2) was treated to the resin. Chitin beads were incubated the elution buffer for 2 days at room temperature on a rotating platform before the elution was collected. The resin was treated with Chitin elution buffer again to make sure all the UB-intein fusion protein was eluted out. Flow-through (F) and Chitin wash buffer washes (W1, W2) and Chitin elution buffer (E1, E2) were analyzed by SDS-PAGE coupled with Coomassie staining. The remaining UB-intein fusion protein was subjected to rebinding with new-regenerated chitin beads for a second round of elution. All the elutions were analyzed by MALDI-LP or ESI-POS and then HPLC purified HPLC (Semi-preparative C12 column at flow rate 4 mL/min, 10-70% Acetonitrile/water + 0.1% TFA). The correct fractions were characterized by mass spectrometry again and lyophilized to dryness.

2.7.6 Native chemical ligation to generate diUB

UBCysK (3.8 mg, 225 nmol) was dissolved in 100 µL ligation buffer (200 mM sodium phosphate pH 7.1, 6 M GdnCl, 100 mM mercaptophenylacetic acid (MPAA), 60 mM TCEP)[104]. In parallel, UB~MesNa (9.8 mg, 453 nmol) was dissolved in 200 µL ligation buffer and 100 µL was added to the UBCysK solution at 0 hour. The reaction was stirred at room temperature under N_2 for 2 hours before the second portion (100 µL) of UB~MesNa was added. The resulting reaction was stirred for another 2 hours and UBCysK, UB~MesNa, reaction mixture at 0, 2, 4 hours were analyzed by SDS-PAGE coupled with Coomassie staining. The reaction was quenched with 1M

TCEP (pH=8.0) and then diluted into 3 mL high salt buffer (200 mM sodium phosphate pH 7.1, 6 M GdnCl.) The diluted solution was dialyzed into 1×PBS buffer supplemented with 1 mM DTT. The next day, the dialyzed solution was treated with Ni-NTA resin and incubated on a rotating platform at 4°C cold box overnight. The unreactive UB~MesNa or hydrolyzed donor UB was enriched in the flow-through (F), NTA lysis buffer wash (L), NTA wash buffer washes (W1,W2) before the desired diUB with little unreactive mono aceeptor UB was collected in NTA elution buffer (E1,E2,E3,…). The fractions during the NTA affinity purification were analyzed by SDS-PAGE coupled with Coomassie staining. The elutions containing diUB product were combined and dialyzed into cation-exchange buffer A (50 mM ammonium acetate, pH=5.0, 1 mM 2-mercaptoethanol) and further purified by Bio-Rad GNC in a 15-80% gradient of cation-exchange buffer B (50 mM ammonium acetate, pH=5.0, 500 mM sodium chloride, 1 mM 2-mercaptoethanol). The fractions were analyzed by SDS-PAGE gel and then diUB product was collected and dialyzed into NTA dialysis buffer (25 mM Tris, 50 mM NaCl, 1 mM DTT).

3 Targeting the unknown ubiquitin chain regulators:

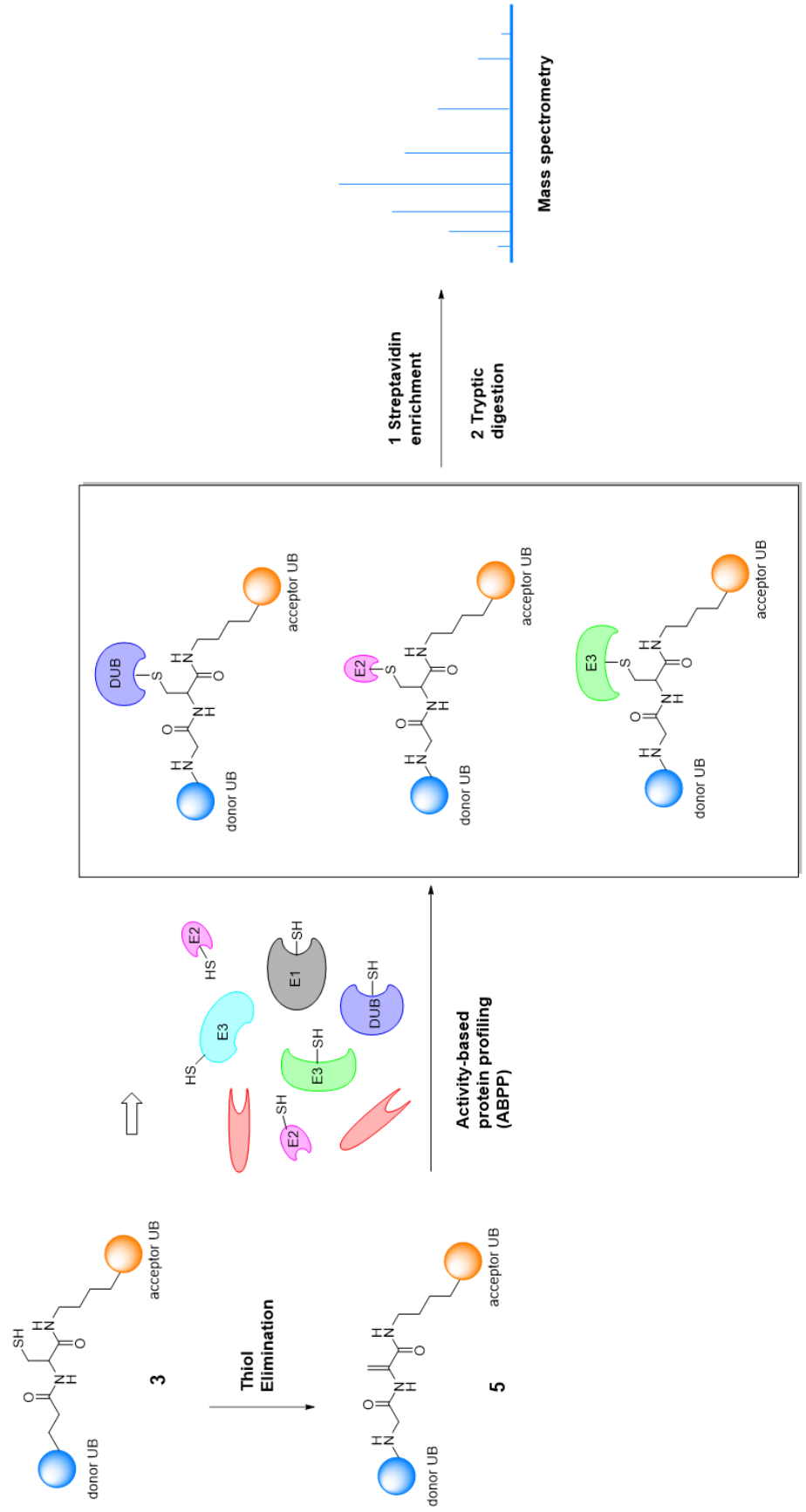
Activity-based protein profiling by Di-Ubiquitin probes

Activity-based protein profiling (ABPP) utilizes interdisciplinary approaches including organic synthesis, protein engineering, sometimes isotope labeling coupling with cutting-edge mass spectrometry analysis to design active-site-directed covalent probes to interrogate specific subsets or families of enzymes in complex proteomes and to provide the basis for a quantitative readout of the functional state of individual enyzmes.[167, 168] Compared to *in vitro* protein activity assay, ABPP probes subsets of enzymes with more accurate assessment of their function states in cells and tissues to have a more clear understanding in terms of protein expression and post-translation modifications. Initially, the functions of serine hydrolases[169, 170], cysteine proteases[171, 172] and kinase[173] were elegantly probed and investigated with ABPP, revealing the distinct reactive mechanisms and substrate specificity. Most recently, ABPP was used widely for enzyme activity profiling including DUBs [174], enzymes associated with cancer or microbe. [175-177]

3.1 Generating thiol-reactive Di-Ubiquitin probes through dethiolation

DiUB synthesized by UAA incorporation and native chemical ligation bears a cysteine near the linkage site of acceptor and donor UB. Notably, this very cysteine serves at first as the nucleophile that initiates the native chemical ligation and it is not redundant or doesn't require desulfurization given that it can be utilized as a reactive warhead to probe the chain specificity of

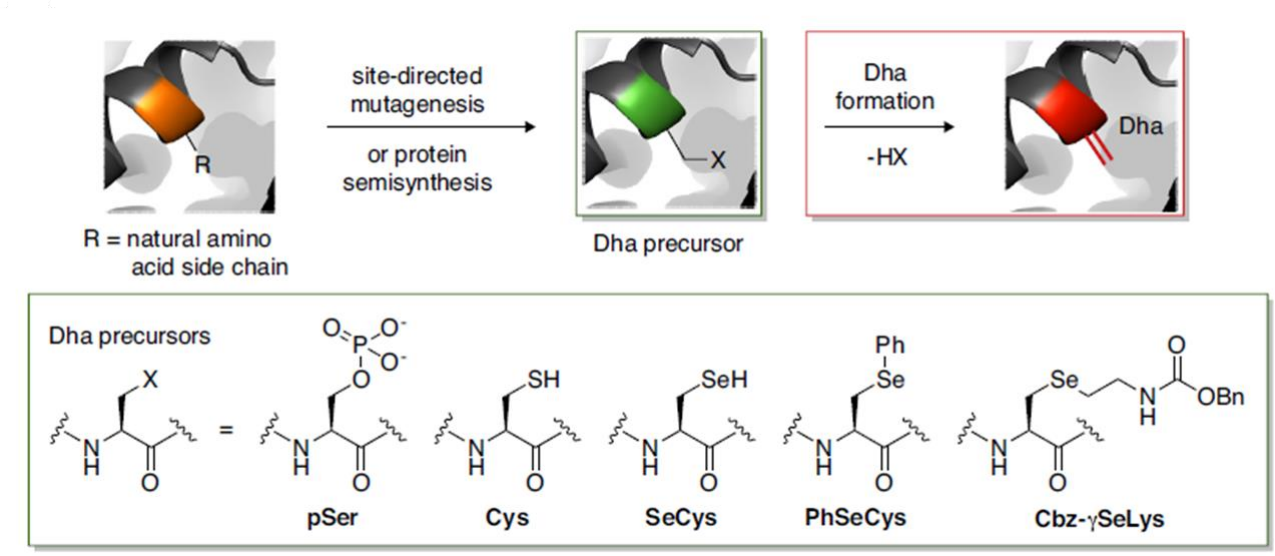
known E2, E3, DUBs through forming tri-protein complex followed by structural analysis or unknown ubiquitin regulators through activity-based drug profiling design. (Scheme 3-1)

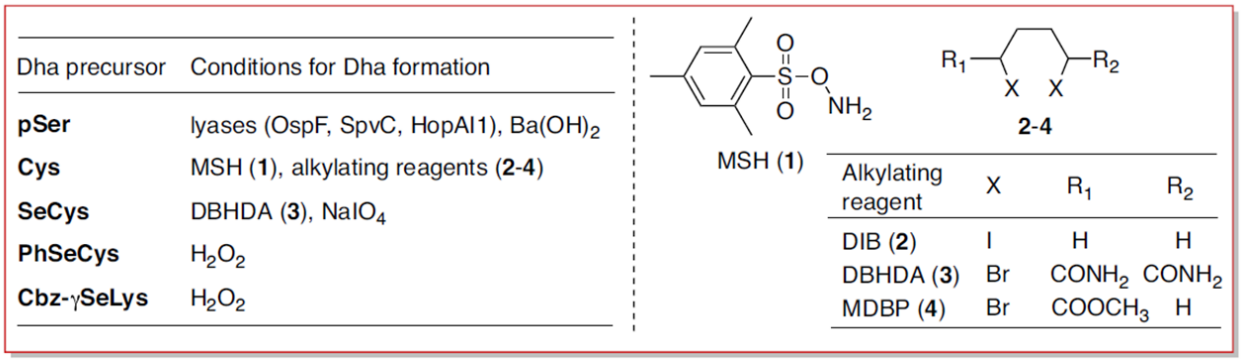


Scheme 3-1 Activity-based protein profiling with diUB-Dha

In this design, diUB-Cys (**3**) prepared is transformation into diUB-Dha (**5**) featuring a 1,4-Michael acceptor, which can covalently capture any E2, E3 or DUBs that can recognize and react with diUB with pre-defined linkage. After treating diUB-Dha (**5**) with lysed proteasome, the linkage-specific ubiquitin chain regulators will be linked to the probe through a stable thiolether bond. In addition, since acceptor UB is His×6 tagged and HA tagged, diUB probe along with conjugated UB chain modifiers can be enriched and immobilized with streptavidin beads. Upon tryptic digestion, the signature peptides can be detected by LC-MS and the intact protein can be identified.

Converting cysteine embedded in a peptide or protein into an electrophilic trap, dehydroalaine, proves to be an effective way for orthogonal protein labeling[178], protein conjugation[179, 180] and activity-based protein profiling[117, 181]. Several bio-compatible approaches were reported to fulfill this task.[165] In general, the thiol residue, inserted at the position of interest, is first transformed into leaving groups, which upon elimination, yield Dha. (Figure 3-1)





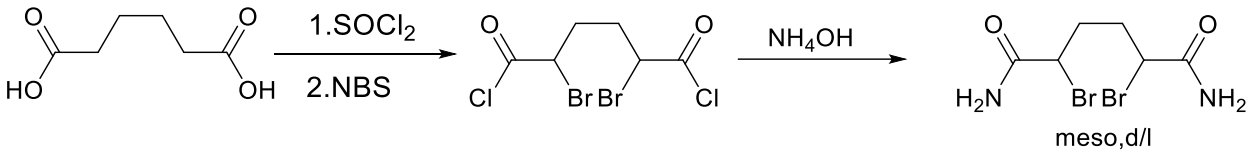
Dha precursor	Conditions for Dha formation
pSer	lyases (OspF, SpvC, HopAI1), $Ba(OH)_2$
Cys	MSH (**1**), alkylating reagents (**2-4**)
SeCys	DBHDA (**3**), $NaIO_4$
PhSeCys	H_2O_2
Cbz-γSeLys	H_2O_2

Alkylating reagent	X	R_1	R_2
DIB (**2**)	I	H	H
DBHDA (**3**)	Br	$CONH_2$	$CONH_2$
MDBP (**4**)	Br	$COOCH_3$	H

Figure 3-1 Methods to introduce dehydroalanine (Dha)

Methods to introduce dehydroalanine (Dha) as a 'tag' to proteins. The position of interest is typically activated by conversion to a Dha precursor followed by its elimination to Dha. Protein taken from PDB: 1N2E

To this end, 2,5-dibromohexandiamide or α,α'-Dibromoadipyl(bis)amide, DBHDA or DBAA, was used to fulfill the conversion due to the reported convenient dethiolation approach mediated by it.[117, 165, 182, 183] First of all, DBAA was synthesized following reported protocol[165].(Scheme 3-2) Gram-level of final product was obtained with full characterization by ^{1}H-NMR and ESI-POS.



Scheme 3-2 Synthesis of α, α'-Dibromoadipyl(bis)amide (DBAA)

With DBAA available, the dethiolation was carried out according to reported protocol. Like previously described, DBAA is prone to be nucleophilic attacked by the thiol on diUB-Cys (**3**) yielding in DBAA adduct on diUB (**3a**). (Figure 3-2) This adduct can undergo another intramolecular nucleophilic attack to form a five-member ring containing sulfur cation

intermediate (**3b**). Now with a good leaving group generated, **3b** can yield to dehydroalanine product (**5**) via β-elimination.

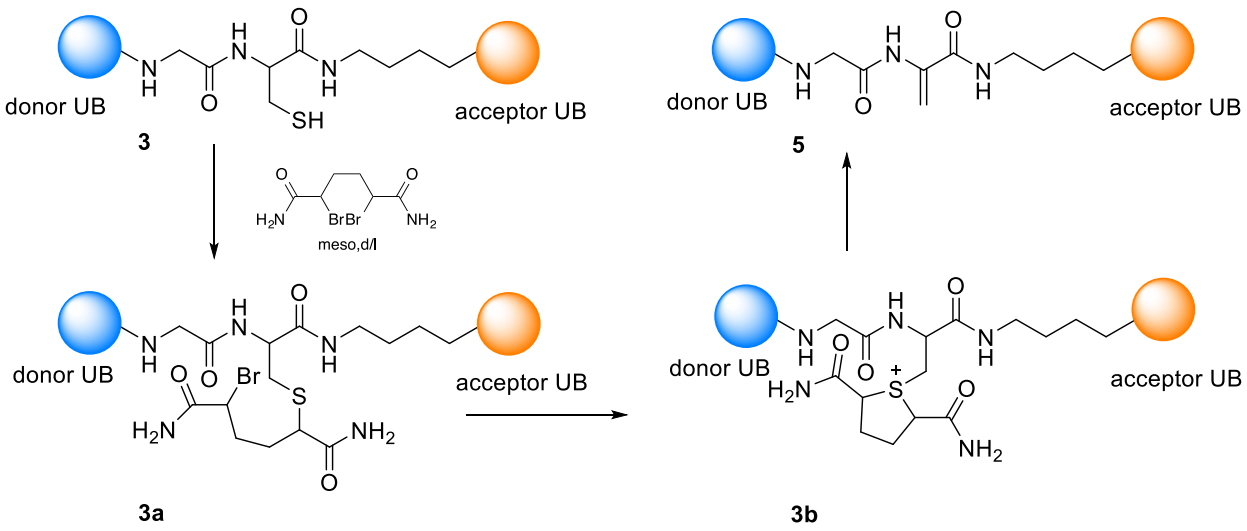


Figure 3-2 DiUB dethiolation mechanism via DBAA

The reaction was analyzed by offline LC-MS. The diUB-Cys (**3**) has a molecular weight of 20485, which corresponded to the calculated molecular weight 20484. (Figure 3-3) After the incubate diUB-Cys with DBAA at 37°C for 4 hours, the reaction mixture was analyzed by LC-MS yielding 20450, which also corresponded to the calculated molecular weight 20450. Notably, the mass loss of 34 after the dethiolation was detected by the LC-MS before and after the reaction, indicating the success of conversion.

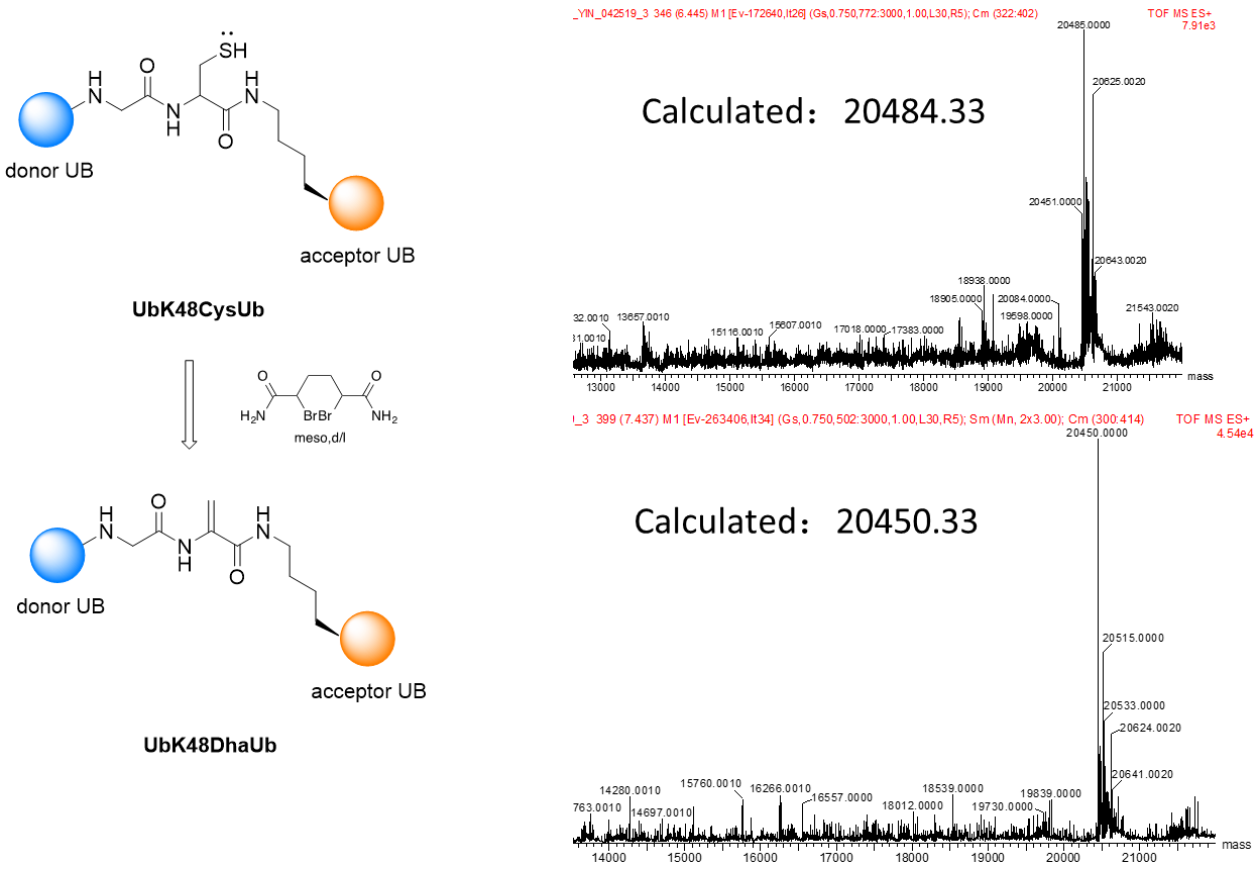


Figure 3-3 ESI-POS shows the correct molecular weight of diUB before and after dethiolation.

A biotin labeling assay was performed to validate the formation of Michael acceptor. If the conversion from cysteine to dehydroalanine is successful, the diUB-Dha probe should react with Biotin bearing a terminal thiol (Figure 3-4A). Indeed, the diUB-Dha was able to be labeled by Biotin forming diUB-Biotin that can be detected on Western blot blotting on Streptavidin. (Figure 3-4B) Notably, diUB-Dha itself in Lane A was free of Biotin and Biotin-PEG-SH itself in Lane C only had small molecular (less than 10 kDa) bands, which both serve as negative control for this assay.

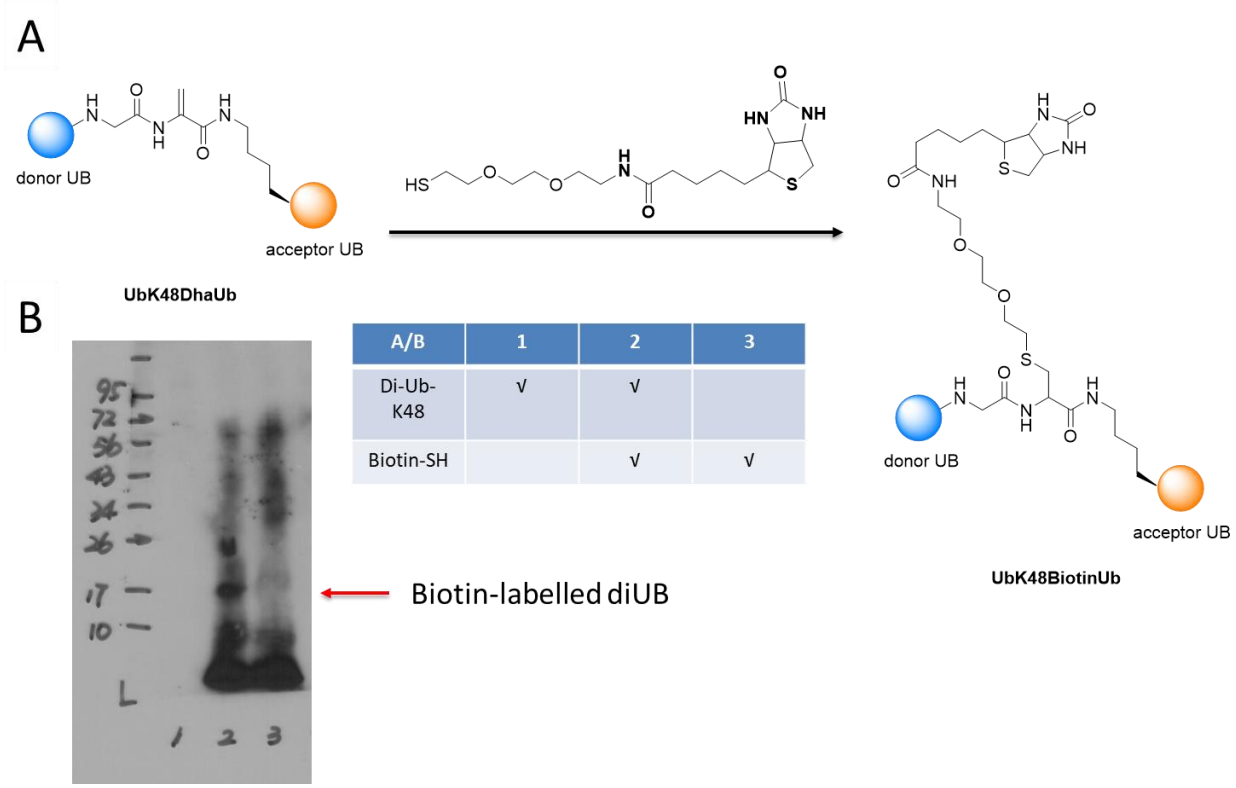


Figure 3-4 Biotin-PEG-SH labelling assay for diUB-Dha

(A) Biotin-PEG-SH reacts with diUB-Dha to form an adduct onto diUB chain.
(B) Western blot blotting on Streptavidin to show the Biotin-labelled diUB with a correct MW band around 20 kDa.

3.2 Verifying the reactivity of Di-Ubiquitin probes with known regulators

To confirm the conversion from diUB-Cys (**3**) to diUB-Dha **(5**) and its reactivity towards various linkage-specific UB regulators, K48-diUB-Dha was synthesized at first given that there are several known E2, E3, or DUB that can target specifically to K48 chains or at least reported to mediate K48 chains. After the reaction and characterization (Figure 3-3), K48-diUB-Dha was dialyzed into low salt buffer to remove small molecular reagents. The labelling reaction carried out at pH 7.6 which was around physiological pH to prevent site reactions from other nucleophiles on protein, such as lysine side chain. Additionally, EDTA (0.1mM) was added to remove any trace amount of metal ions that are predicted to oxidize thiols into disulfide bond. The reaction can be

monitored by both Coomassie gel and Western blot blotting on HA given that acceptor UB is equipped with an N-terminal His-HA tag.

For this purpose, two E2s (Ubch5b, Ubch7) and two HECT E3s (HUWE1, NEDD4) were selected and reacted with K48-diUB-Dha. Ubch5b belongs to Ube2D subfamily that is reported to build up promiscuous UB chains whereas Ubch7 (Ube2L3) only transfer charged UB to a cysteine residue on HECT or RBR E3s without ubiquitin chain elongation activity. HUWE1 is known to mediate K6, K48, K63 chains on substrate while NEDD4 extends K63 chains. Surprisingly, K48-diUB-Dha was not reactive to Ubch5b, Ubch7, HUWE1 or NEDD4 based upon Coomassie gel and WB results. (Figure 3-5)

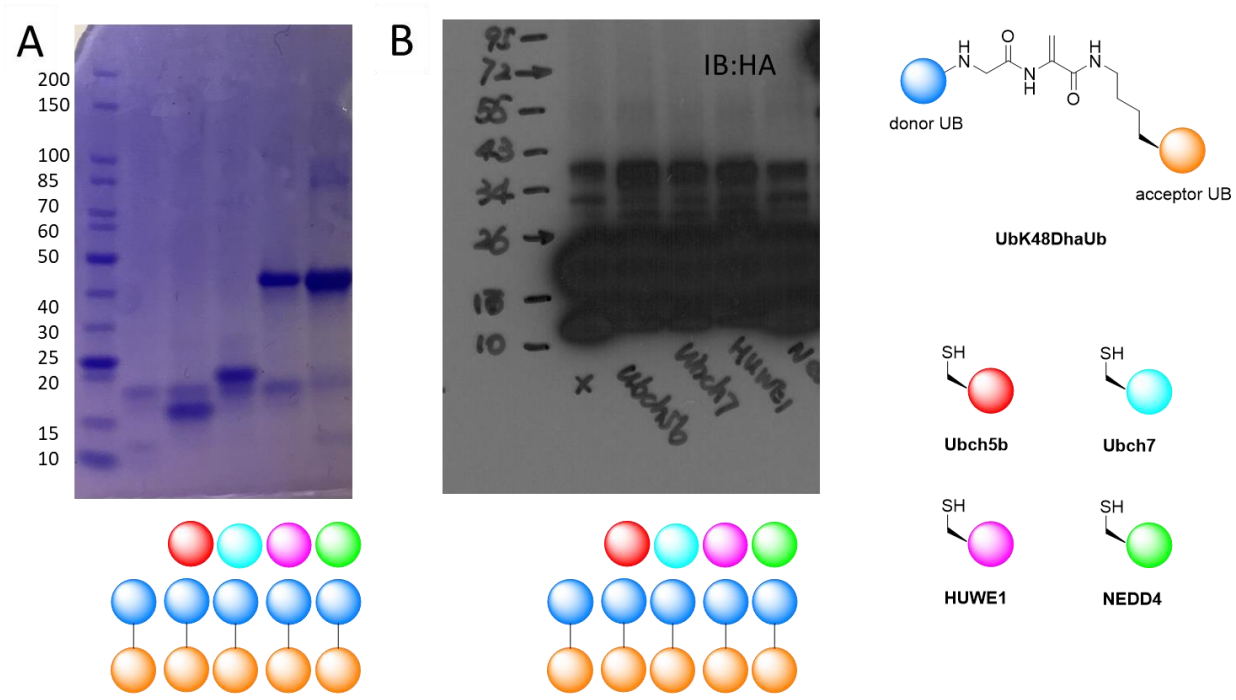


Figure 3-5 K48-DiUB-Dha reacts with E2s and E3s

(A) SDS-PAGE with Coomassie staining shows the reactions between K48-diUB-Dha and various colored-coded E2s or E3s.
(B) Western blot (IB: HA) shows the reactions between K48-diUB-Dha and various colored-coded E2s or E3s.

While Ubch7 and NEDD4 inert to K48-diUB are expected, Ubch5b may need extra components, for example, E3s, to extend UB chains. In this regard, E2-25K (Ube2K) is reported to elongate K48 UB chains in the absence of an E3[184], thus it is the next subject to try with K48-diUB-Dha labeling. Another possible reason is due to the reactivity of catalytic cysteines on E2 or E3s. The nucleophilicity of cysteines on them is effective enough to facilitate transthiolation from UB~E2 or UB~E3 but may not be trapped to an 1,4-Michael acceptor at physiological pH. Nevertheless, the double bond of dehydroalanine embedded on the isopeptide backbone of diUB is next to an amide. As a result, the electron density of this double bond is increased by forming an enamine due to resonance, rendering it less electrophilic for less reactive thiols.

Considering the reactivity, DUBs with a catalytic cysteine that mediates amide bond cleavage in UB chain should be more reactive to diUB-Dha. Indeed, labeling reactions were observed when incubating K48-diUB-Dha with various DUBs that were known to cleave K48 linkages.(Figure 3-6) OTUB1 has been reported to mediate exclusively K48 chains[86] whereas OTUB2 is also known to mediate several linkage UB chains including K48 ones[185, 186]. Interestingly, as recently reported, the roles of both OTUB1 and OTUB2 are heavily implicated in cancer[187, 188]. Clinical studies have associated elevated OTUB1 expression with high grade, invasiveness and metastasis in several tumor types including lung, breast, ovarian, glioma, colon and gastric. In addition, OTUB1, as one of the most abundant DUBs in cell, displays a catalytic-independent non-canonical activity where it inhibits the transfer of UB onto protein substrates by sequestration of E2.[86] Similarly, OTUB2 was reported to promote cancer metastasis via Hippo-independent activation of YAP and TAZ, two transcriptional regulators that play important roles in tumorigenesis.

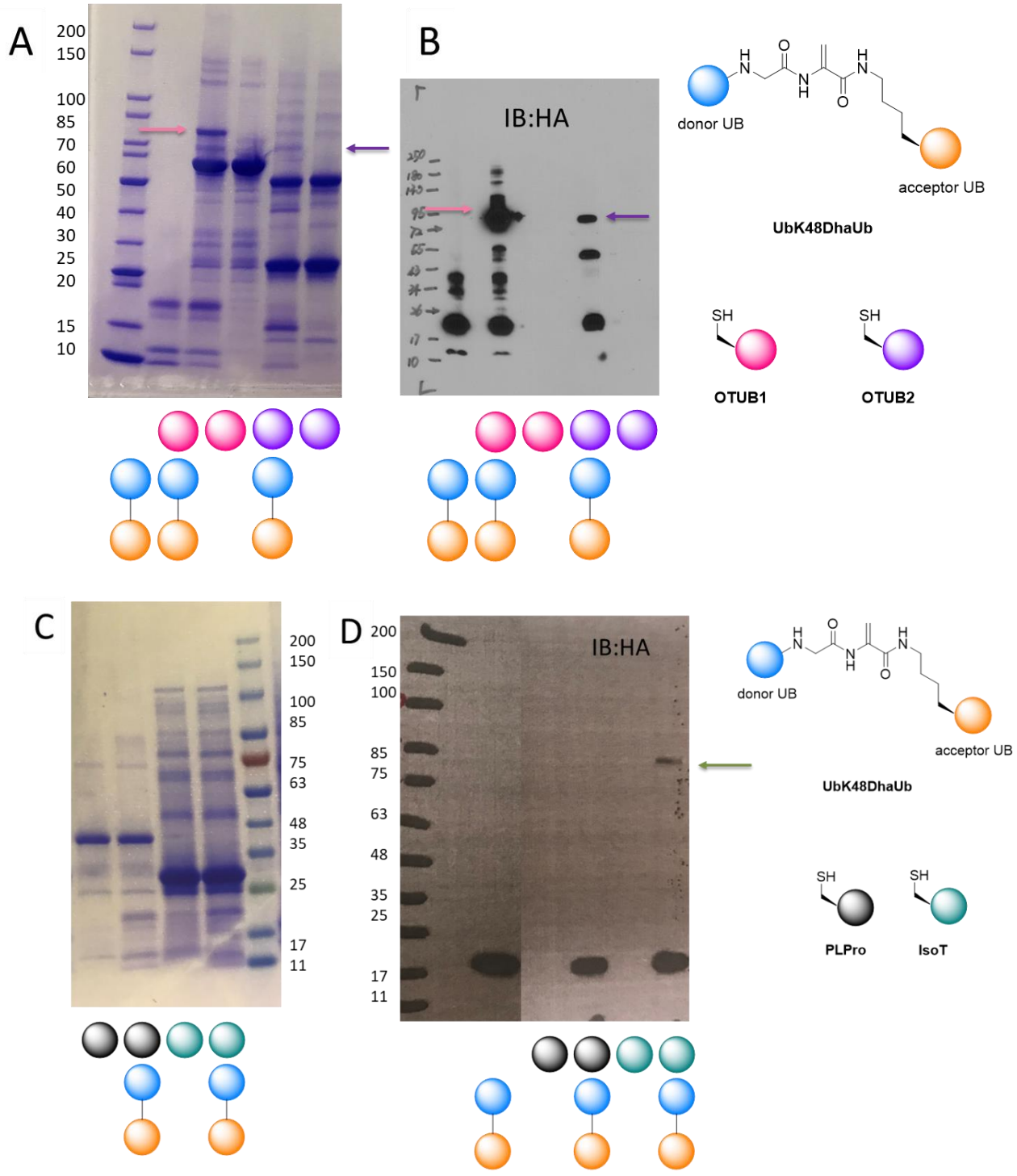


Figure 3-6 Reaction between K48-DiUB-Dha with DUBs

(A)(B) Coomassie gel and Western blot (IB: HA) results suggest K48-DiUB-Dha was able to react with OTUB1 and OTUB2.
(C)(D) Coomassie gel and Western blot (IB: HA) results suggest K48-DiUB-Dha was able to react with isoT but not PLPro.

Two more DUBs were tested against K48-DiUB-Dha, namely, IsoT and PLPro. Other than OTUB1 and OTUB2, IsoT belongs to USP superfamily that is known to mediate promiscuous UB chains due to lack of a well-defined S' site. In particular, isoT (USP5/USP14) specifically dissembles unanchored K48 and linear UB chains from the proximal end.[189, 190] In addition, a viral DUB, the severe acute respiratory syndrome (SARS) Coronavirus Papain-like Protease (PLPro)[191], that cleaves ISG15, a two-domain UB-like protein and K48 UB chains respectively was selected.

Indeed, after the reaction, the diUB-labelled OTUB1, OTUB2, isoT can be successfully detected by SDS-PAGE gel coupled with Coomassie staining and Western blot immunoblotting HA on diUB. (Figure 3-6). Notably, DUBs alone and diUB itself didn't reveal the diUB-DUB band as expected. However, PLPro was not able to react with K48-DiUB-Dha. This founding actually matches the observation reported in a paper co-published by Ovaa, Huang, Lima[191] where SARS-PLPro was treated with monoUB-ABP, distal-DiUB-ABP and in-between-DiUB-ABP made from SPPS. As a result, distal-DiUB-ABP reacted well with PLPro while monoUB-ABP reacted fairly but in-between-DiUB-ABP (the one that similar to our K48-DiUB-Dha) was inert to PLPro. They provided an explanation to this in the paper based upon the crystal structure of SARS-PLPro with several UB binding sites including S1, S1', S2, S3. They claimed that PLPro possibly binds the in-between-DiUB-ABP via S2-S1, preventing it from binding and reacting via S1-S1' interactions.

Lastly, as a negative control, DUBs were incubated with DiUB-Dha pretreated with β-Mercaptoethanol (BME). Presumably, BME featuring a terminal thiol will block the reactivity of DiUB-Dha by forming an adduct bound with thiolether. The result was as expected(Figure 3-7); no DUB could react with BME-pretreated DiUB probe, proving the reactivity of DiUB-Dha is due to the electrophilicity of Dehydroalanine motif with a Michael acceptor.

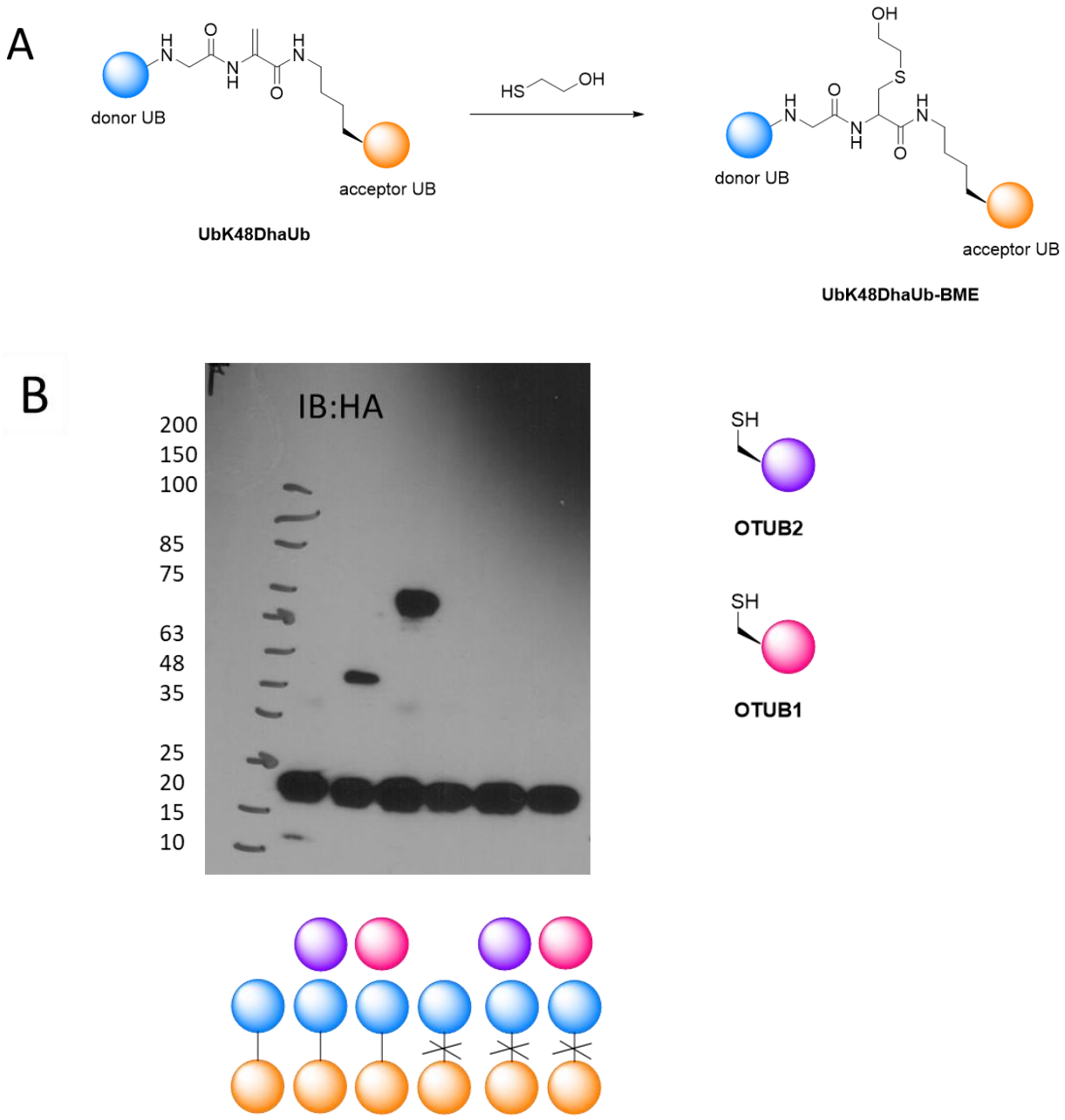


Figure 3-7 DUBs react with K48-DiUB-Dha pretreated with BME

(A) Reaction between DiUB-Dha with BME yielding an BME adduct therefore blocks its reactivity
(B) OTUB1 and OTUB2 were able to react with K48-DiUB-Dha but not BME pretreated K48-DiUB-Dha suggested by Western blot (IB: HA)

3.3 Conclusion

To summarize this chapter, DiUB-Cys prepared by unnatural amino acid incorporation and native chemical ligation was readily transformed into an activity-based probe by dethiolation. The generated DiUB-Dha was able to react with UB regulators, especially linkage-specific DUBs with different efficiency. DiUB embedded with dehydroalanine was synthesized with SPPS by Brik group and the DUBs were probed[120]; our work provided a more convenient way to circumvent time and effort consuming SPPS and enable large-scale of diUB production available for protein profiling with recombinant UB regulators or cell lysate. Notably, due to the unique position of the 1,4-Michael acceptor in the diUB backbone, the electrophilicity was moderate and prone to be reactivate to stronger thiols but not all UB regulators featuring a catalytic cysteine. This could bring an advantage as an activity-based probe that selectivity is ensured both by the linkage specificity and the reactivity of the handle. Nevertheless, for some DUBs like PLPro that cannot be properly labelled, UB chain probes with longer length or branched architecture can be a future direction. Furthermore, some regulators may require additional components to be full reactive; probing the diUB-Dha with cell lysate should be able to “fish-out” more reported and even more intriguingly unknown linkage-specific UB regulators in their “active” mode. Last but not the least, the linkage types can be expanded to atypical chains such as K6, K27, K29, K33 given that a lot of their regulators (E3s, DUBs) have not be discovered yet.

3.4 Material and Method

3.4.1 α,α'-Dibromoadipyl(bis)amide (DBAA) synthesis

DBAA was synthesized according to a reported protocol with some modifications[128, 165]. Adipic acid (5.00 g, 35.1 mmol) was added to a 500 mL round bottom flask and suspended in thionylchloride (25.0 mL, 207 mmol). The flask was equipped with a condenser and the reaction was heated to reflux (open to air, bath temp 80 ºC). After 30 minutes at reflux, all adipic acid had dissolved. The reaction was stirred for an additional 60 minutes at reflux and then cooled to room temperature. CCl_4 (20 mL) was added to the reaction followed by NBS (18.7 g, 82 mmol). The reaction was stirred vigorously and 2 drops of HBr (48% aq.) was added by pipette. The reaction was heated to reflux, again open to air. The reaction gradually turns from red to black over the course of an hour. After 2 hours at reflux, the reaction was cooled to room temperature and then to 0 ºC. The mixture was stirred at 0 ºC to ensure all succinimide had precipitated. The solid was removed by filtration. Et_2O (20 mL) was used to rinse and complete the filtration. The filtrate was concentrated under reduced pressure to give a thick, dark red liquid. In a 500 mL round bottom flask, 50 mL of NH4OH (25% aq.) was cooled to 0 ºC. The crude acid chloride was added dropwise over 20 minutes to the ammonia solution with rapid stirring. After the addition was complete, the reaction was stirred vigorously at 0 ºC for 1 hour. The bis-amide product precipitated from the reaction mixture. The dark solid was isolated by filtration and partially dried. The product was purified by triturating in MeOH/H2O: The solid was suspended in H2O (100 mL) and MeOH (100 mL) and heated to 60 ºC. The mixture was stirred rapidly at 60 ºC for 30 minutes. After this time, the mixture was cooled to room temperature. The resulting white solid (a mixture of meso and d/l diastereomers) was isolated by filtration and washed with MeOH (200 mL). The product (22) was dried under high vacuum (3.6 g, 37%). 1H NMR (400 MHz, DMSO): δ = 1.75-2.08 (4H, m,

CH2CH2), 4.28-4.36 (2H, m, 2 Å~ CHBr), 7.30 (2H, s), 7.69 (2H, s) (2 Å~ NH2). 13C NMR (100 MHz, DMSO): (both diastereomers reported), δ = 32.5, 32.6 (CH2CHBr), 48.2,48.5 (CH2CHBr), 169.87, 169.92 (C=O). LRMS m/z (ESI+): 324.9 [M+Na]+.

3.4.2 DUB expression

OTUB1, OTUB2, IsoT(USP5) and PLPro plasmid were purchased from Addgene. Upon transformation into *E.Coli* BL21(DE3) cell, OTUB1,OTUB2 and PLPro were expressed according to a standard His tag protein expression protocol. [84, 191, 192] IsoT was expressed based upon a standardized GST protein expression protocol except for an additional "heat shock" step to activate the molecular chaperones to correctly fold the protein. [193, 194]

3.4.3 Dethiolation with DBAA

Dethiolation was carried out following reported protocol with some modifications.[165] cAb-DiUB-Cys was first reduced with DTT to remove any contaminant disulfide: DTT (1.2 mg, 7.8 umol) was added to 500 uL of DiUB-Cys (c = 1 mg/mL in 10 mM PBS, pH 8.0) and shaken at room temperature for 15 minutes. After this time, the protein solution was passed through a PD minitrap (GE Healthcare), previously equilibrated with 10 mM PBS (pH 8.0), eluting with 1 mL of the same buffer. The reduced protein was stored on ice until needed. DBAA (1.2 mg, 3.3 μmol) was added to a 1.5 mL plastic tube as a solid. An 875 μL aliquot of the reduced DiUB-Cys prepared above (c = 0.5 mg/mL in 10 mM PBS, pH 8.0) was added to the same tube. No DMF was used in this reaction. The reaction was shaken at room temperature for 1 hour, and then 37 °C for 4 hours. The reaction was then cooled to room temperature and precipitated DBAA was removed by centrifugation (1 min, 16K g).

3.4.4 Labeling reaction with DUBs

DiUB-Dha was dialyzed into 50 mM Tris, pH=7.5, 25 mM NaCl containing 0.1 mM EDTA; DUB was dialyzed into 50 mM Tris, pH=7.5, 25 mM NaCl containing 0.1 mM EDTA and then pretreated with 1mM TCEP upon reaction. 15 μM DiUB probe was incubated with 20 μM DUB at 37 ºC for 5 hours and the reaction was analyzed with SDS-PAGE coupled with Coomassie staining and western blot immunoblotting on HA.

4 DEFINING THE TOPOLOGIES OF UBIQUITIN CHAINS: THE LINKAGE SPECIFICITY OF E2S, E3S OR DUBS REVEALED BY DI-UBIQUITIN PROBES

Protein-protein interactions (PPIs) are critical regulatory events in physiology and pathology, and they represent an important target space for pharmacological intervention. However, protein-protein interactions are normally hard to capture due to the weak binding affinity between their interfaces and dynamic topologies during protein trafficking. [195] To this end, Chemical cross-linking, along with yeast two-hybrid, fluorescence resonant energy transfer (FRET), and co-immunoprecipitation have become important tools for detection and characterization of protein-protein interactions.[196, 197] In particular, chemical crosslinking helped to determine the interfaces between two entangled proteins by increasing the binding affinity and stabilizing transient protein complex. [129, 198] So far, mechanism-based design of a tri-component protein complex proves to be novel approach to decipher weak and transient protein-protein interactions.

4.1 E2 at work: Probing UBE2S chain specificity with K11 Di-Ubiquitin

It has been known a decant ago that K11 UB chains are critical cell-cycle regulators in human cells.[44] These K11-linked chains are synthesized during mitosis by the E3 anaphase-promoting complex (APC/C) and its E2s Ube2C and Ube2S.[35, 199, 200] Working together, these enzymes ubiquitinate mitotic regulators, such as cyclin B, securing, HURP as a degradative signal.[201] In this regards, inhibition K11 polyubiquitination blocks mitotic progression in *Xenopus*, *Drosophila* and humans as a result of inaccurate cell division and tumorigenesis. However, the

molecular basis of generating K11 chains by APC/C and its E2s was largely unclear when we tried to stress this question by diUB probe.

Previously, Rape group had proposed a model of the interactions between donor UB and Ube2S [202]; Ube2S was known to mediate K11 chains on mitotic substrates. According to NMR mapping and docking, a noncovalent hydrophobic binding of donor UB is required for Ube2S activity (Figure 4-1A). Therefore, some key residues were identified through site-mutagenesis. However, authors also claim that the acceptor binding was too transient to be detected by NMR, independently of whether the donor ubiquitin had been linked to the Ube2S activate site or not.[202] Later on, to validate the donor UB~Ube2S interaction, a stable linkage between them was made through disulfide bond to replace the original unstable thioester bond; the structure of the crosslinked heterodimer was solved and reported (Figure 4-1B)[203]. As a result, the structure contains an interface between Ube2S and UB in trans that resembles the earlier model in general

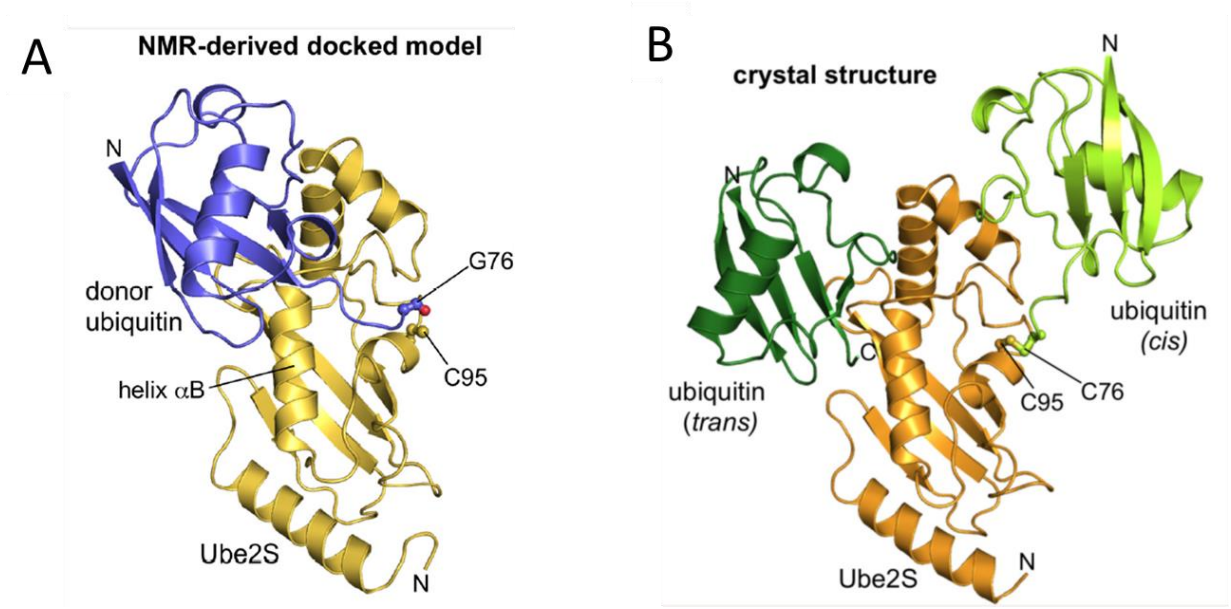


Figure 4-1 Donor UB and Ube2S interaction

(A) NMR-derived docked model. Donor UB in blue and Ube2S in orange.
(B) Crystal structure (PDB: IZDN) of donor UB and Ube2S linked with disulfide. Donor UB (dark green) in trans and a backside binding UB (light green) in cis. Ube2S in orange.

terms but differs in detail in that the crystallographic interface is more hydrophobic. Notably, UB in NMR-derived docked model was claimed to be donor UB but in crystal structure, was a backside binding UB in trans. The disulfide bound UB was in cis, which later on was reported as a autoinhibition component for E2s.[204] Still, no acceptor UB binding was discussed in the paper.

As discussed in Chapter 2.1, to correctly place both donor UB and acceptor UB on Ube2S with direct evidence on structure, it was hypothesized that if the transition state that donor UB and acceptor UB bind simultaneously to Ube2S (**2**) can be captured by stable protein complex (**1**), the detailed intact interfaces should be resolved. (Figure 4-2)

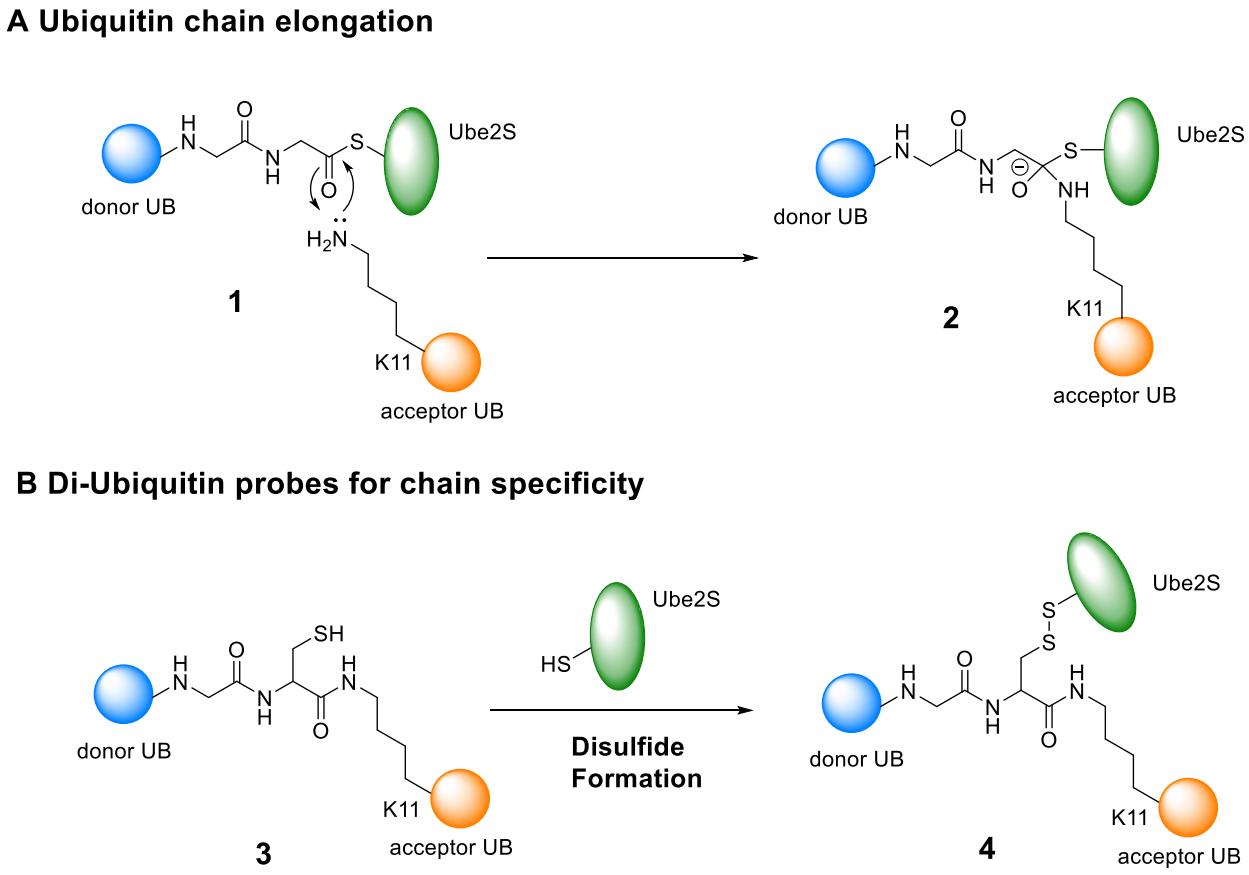


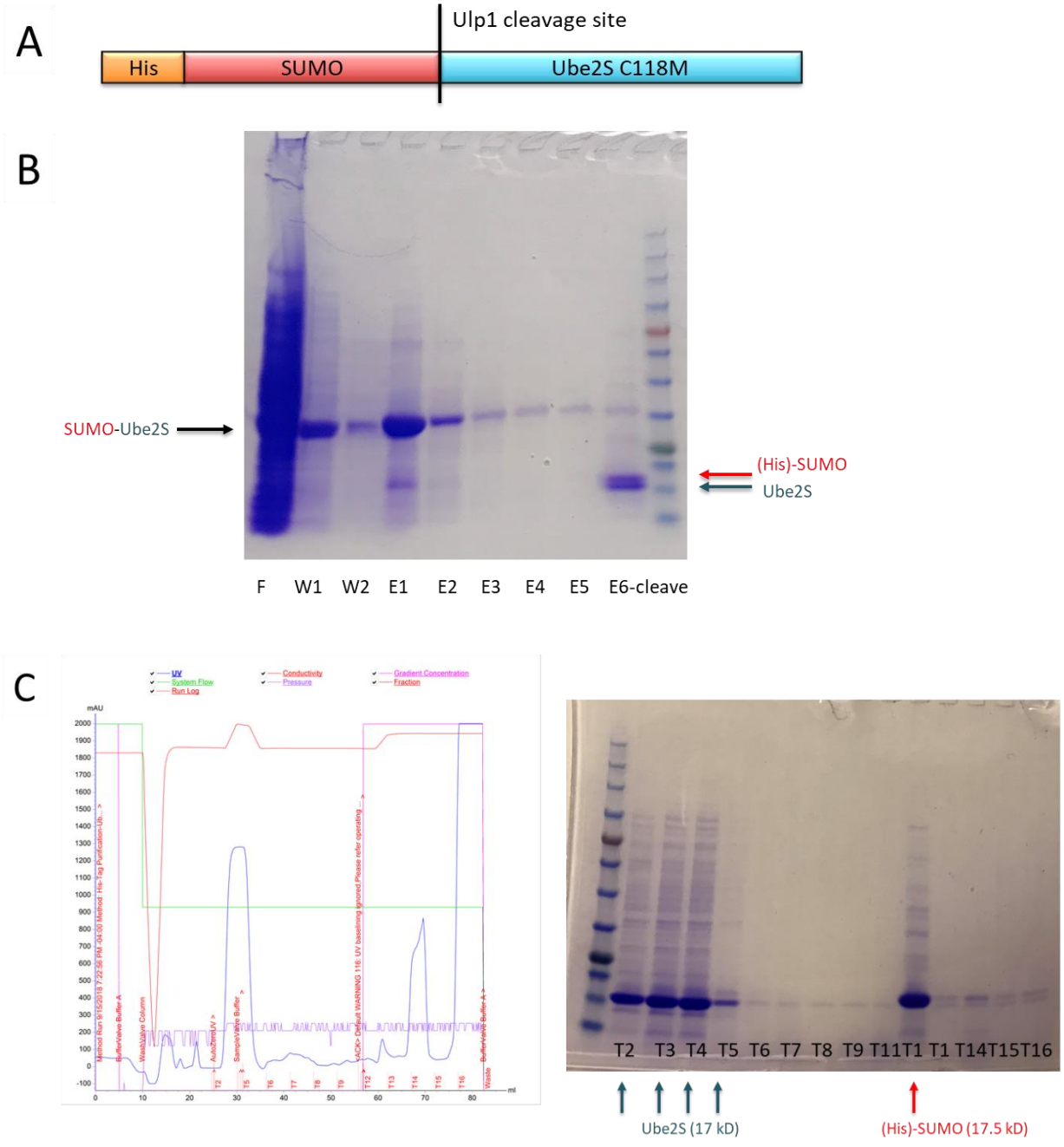
Figure 4-2 Tri-protein complex design to mimic transition state of UB chain elongation

Forming disulfide bond between two proteins or inside one protein is naturally occurring[205, 206] and also frequently used for biochemist. To generate disulfide bond, any reducing reagents are prohibited. Firstly, $CuCl_2$ was used as an oxidative reagent. However, the resulting products were a mixture of UB, Ube2S homodimer and UB-S-S-Ube2S.And the reaction took over two days to reach equilibrium. (Data not shown) Thereby, 5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB (Ellman's reagent)[207] was chosen to generate heterodimeric protein complex. Thiols react with this compound, cleaving the disulfide bond to give 2-nitro-5-thiobenzoate (TNB^-), which ionizes to the TNB^{2-} dianion in water at neutral and alkaline pH. This TNB^{2-} ion has a yellow color.(Figure 4-3) In addition, this reaction is rapid and stoichiometric, with the addition of one mole of thiol releasing one mole of TNB. The TNB2− is quantified in a spectrophotometer by measuring the absorbance of visible light at 412 nm, using an extinction coefficient of 14,150 M-1 cm−1 for dilute buffer solutions, and a coefficient of 13,700 M−1 cm−1 for high salt concentrations, such as 6 M GdnCl or 8 M urea.[208, 209]

Figure 4-3 Mechanism of Ellman's reagents to form disulfide bond

To probe Ube2S linkage specificity, recombinant Ube2S (UBC domain, residue 1-156) was first prepared in large scale; this excludes a flexible 66-residue C-terminal extension which cannot be crystalized.[203] Additionally there are two cysteine (C98 and C118) in Ube2S UBC domain, thus the non-catalytic cysteine was mutated to methionine (C118M). Then Ube2S (UBCdomain) (17 kD) was expressed with an N-terminal His-SUMO tag (17.5 kD); after affinity

purification, the tag was cleaved by SUMO deubiquitiase Ulp1 [210, 211]. The cleaved Ube2S is subjected to affinity purification aided by ATKA Fast Protein liquid chromatography and further polished by anion-exchange chromatography to yield homogeneous Ube2S (C118M) UBC domain. (Figure 4-4)



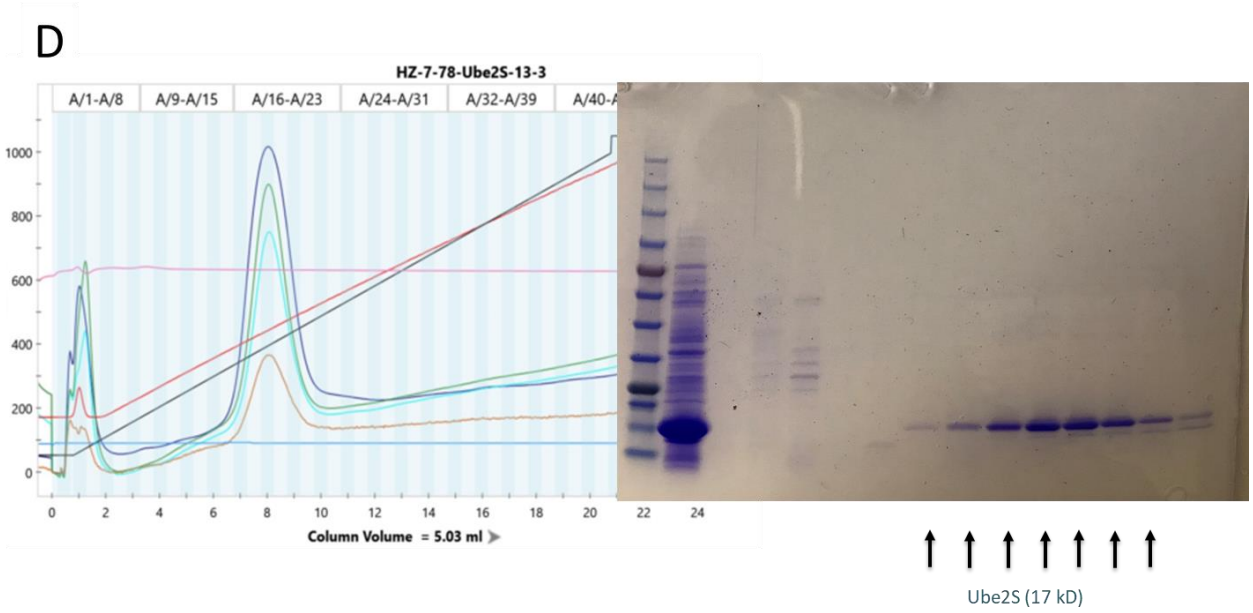


Figure 4-4 Expression and Purication of Ube2S (C118M) UBC domain

(A) Ube2S construct: His-SUMO-Ube2S (C118M) fusion protein upon cleavage by Ulp1
(B) Expression of Ube2S in BL21(DE3) cells and affinity purification of fusion protein. (F: Flow-through; W1, W2: NTA-wash buffer wash; E1, E2: NTA-Elution buffer elution
(C) NTA-Affinity purification to separate cut Ube2S and His-SUMO tag by AKTA FPLC.
(D) Anion-exchange purification to purify Ube2S as a tag-free protein.

Next, the disulfide bond formation was carried out to conjugate Ube2S UBC domain to diUB probe bearing a thiol at the isopeptide site. DiUB-Cys features an His-HA tag on acceptor UB that can facilitate affinity purification; thereby untagged Ube2S was added in excess to consume diUB-Cys to drive the reaction to completion; nevertheless, diUB-Cys probe was hard to prepare whereas Ube2S was obtained in large quantity. Ratio of diUB-Cys and Ube2S-TNB was investigated (results not shown) and finally 1:4 ratio was the best to consume all potential diUB based on Coomassie staining and Western blot results. (Figure 4-5) Because both diUB-Cys and Ube2S-TNB are around 17 kDa, it was hard to observe the completion the reaction on Coomassie staining. Fortunately, on Immunoblotting diUB-Cys band almost disappeared at 1:4 ratio, indicating the full conversion from diUB-Cys to diUB-S-S-Ube2S.

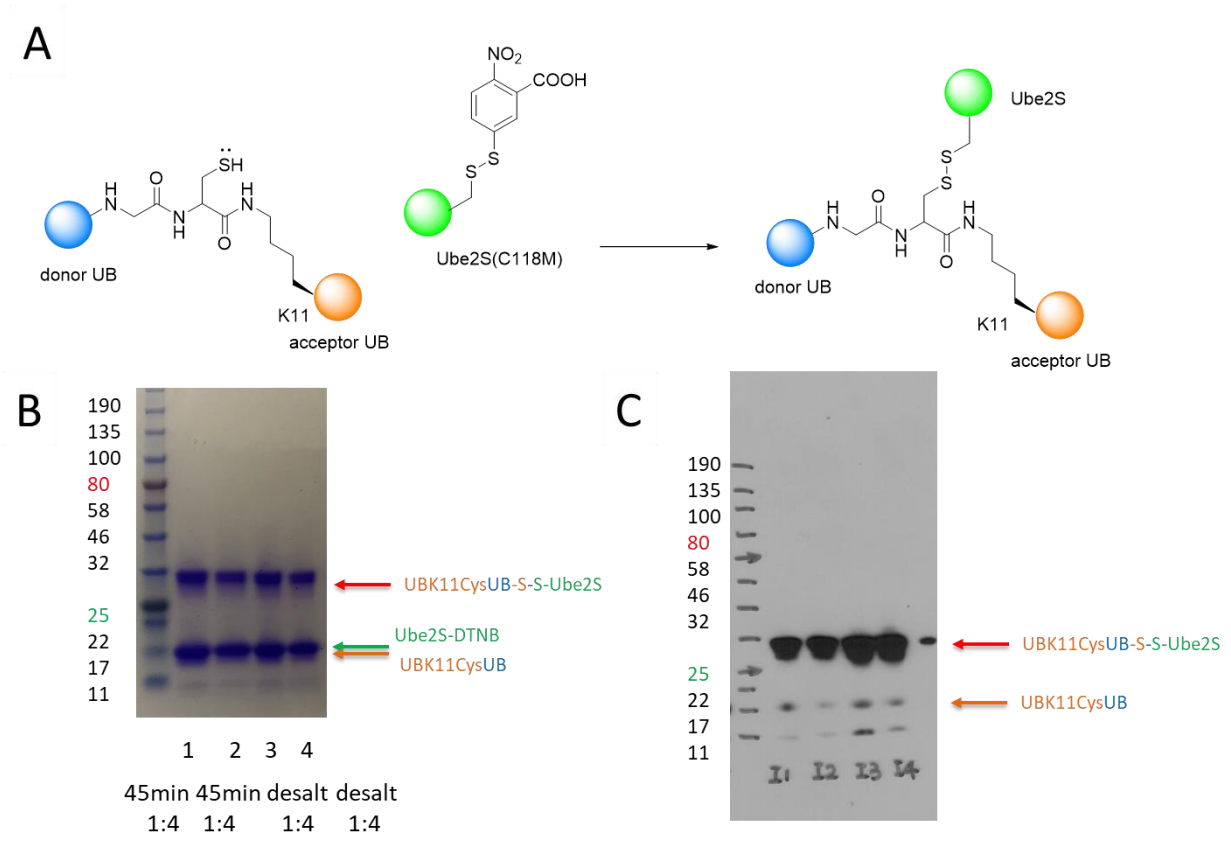


Figure 4-5 DiUB and Ube2S conjugation through disulfide bond formation

(A) Conjugation reaction between UBK11CysUb and Ube2S-DTNB.
(B) Coomassie gel from the reaction mixture (Lane 1 and Lane 2) and after desalting to remove excess TNB.
(C) Western blot (IB: HA) shows the reaction mixture (Lane 1 and Lane 2) and after desalting to remove excess TNB.

A tandem purification approach was utilized to purify the tri-protein complex. After desalting, the crude mixture including mostly diUB-S-S-Ube2S and unreactive Ube2S, was treated with Ni-NTA resin upon affinity purification. The beads were applied to a column and the fluid was collected with gravity filtration. (Figure 4-6A) Notably, an affinity purification method utilizing AKTA chromatography was investigated; however, the result was not as good. During the wash, imidazole concentration was gradually increased from 0, 5, 20, 100, 250 mM. Under this gradient, unbound proteins, mainly Ube2S in excess, were washed out at 0-5 mM imidazole

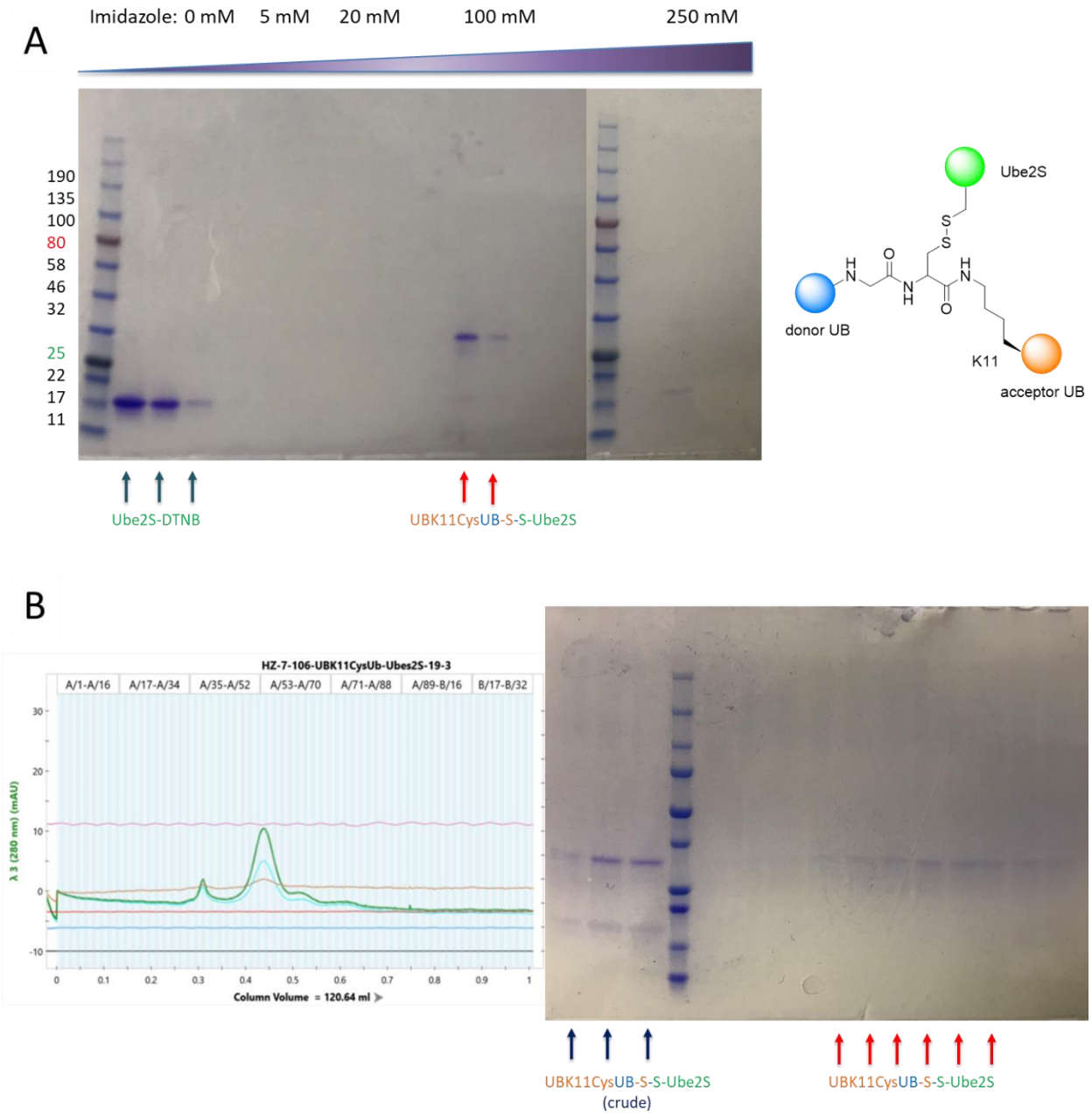


Figure 4-6 Tandem purification of UBK11CysUB-S-S-Ube2S

(A) NTA-Affinity purification of UBK11CysUB-S-S-Ube2S via gravity filtration. The unbound Ube2S and the desire product were separated by increasing imidazole concentration.
(B) Size exclusion to polish the purity of UBK11CysUB-S-S-Ube2S

whereas desired protein complex was collected in buffer containing 100 mM imidazole. The beads then were fully stripped with 250 mM imidazole for cleaning purpose. Lastly, the complex was

further purified by size exclusion. (Figure 4-6B)

The final product going through steps of conjugations and tandem purifications was not decent in amount. When concentrated down, the protein complex concentration can only reach 5 mg/mL due to the limited amount. Using the automatic crystal tray set-up platform, two standard sparse matrix screening was performed, namely JCSG-plus from Molecular Dimensions and Index from Hampton Research; unfortunately, no protein crystal was obtained in all conditions. The yield of the tri-protein complex was not very promising; scaling-up is needed to try more conditions with a higher concentration of stock solution. The follow-up experiment was described in Chapter 4.3.1.

4.2 Probing HECT E3 chain specificity with Di-Ubiquitin

HECT type ubiquitin ligases (E3s) features the same catalytic cysteine in the UB transfer process as E2s and mediates UB chains with different linkage preferences. (Figure 1-10) All HECT E3s present the catalytic HECT domain, composed of a bulkier N-terminal lobe (N-lobe) that contains the E2 binding domain, and a C-terminal lobe (C-lobe) carrying the catalytic cysteine. Additionally, the two lobes are connected by a flexible hinge region that allows the C-lobe to move around in order to facilitate UB transfer.[212-214] Nevertheless, the ability to build linkage-specific poly-UB chains seems to be an intrinsic characteristic of HECT E3s. For example, E6AP is K48 specific[215] whereas NEDD4 family members are known to mediate K63 chains[216]. A structural determination of the complexes formed between HECT domains, donor and acceptor UB is challenging, due to the low affinities of the underlying interactions (high micromolar to millimolar affinity range *in vitro*). To stress this question, chemical mimics of donor-HECT domain complexes[214, 217, 218] reveal that a conserved hydrophobic interface of donor UB with HECT C-

lobe is heavily implicated in its binding; however, no structures of HECT domains with UB in an acceptor position facing the active site have been solved. Here diUB probe conjugated with HECT domain E3 has been proposed to capture both the donor UB and acceptor UB binding on HECT domain.

4.2.1 HUWE1

Recent studies suggest HECT-type E3 HUWE1 regulates the stability of various cellular substrates thus mediates several key physiological processes, including DNA replication and damage repair, cell proliferation and differentiation, apoptosis.[219-221] Interestingly, HUWE1 can target oncoproteins, namely N-MYC, C-MYC, MCL1[222, 223] but also tumor suppressor proteins like p53[224] for degradation. Importantly, HUWE1 was reported to generate K6-, K11-, K48- poly UB chains[225] and it also uses ubiquitinated substrate to seed chain extension by adding K48-linked UB onto K63-linked UB chains formed by other E3s.[226] HUWE1 crystal structure had been solved[227, 228] (Figure 4-7) However, how HUWE1 binds to donor UB and more intriguingly how

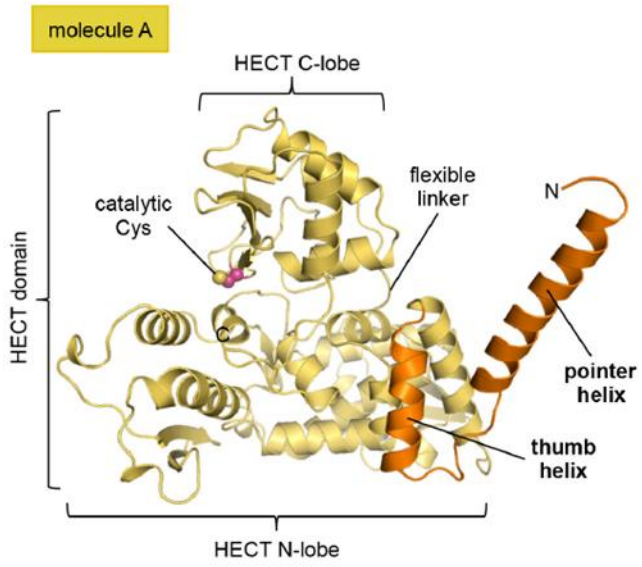


Figure 4-7 Crystal structure of HUWE1 (HECT domain) with an N-terminal extension of dimerization region. (PDB 5LPB, Ref.228)

HUWE1 orientate acceptor UB to achieve chain specificity were largely unknow. To this end, the structure of UBK48CysUb conjugated with HUWE1 HECT will be very interesting and informative to determine given that it can provide direct evidence of HUWE1 chain elongation mechanism.

To maintain the homogeneity of disulfide product, other than the catalytic cysteine (C4341), two extra cysteines that lies outside the enzyme was mutated to alanine. Recombinant HUWE1 (HECT domain) with an N-terminal His tag followed by an HRV-3C protease cleavage site was expressed with BL21(DE3) cells and purified with standard Ni-NTA affinity purification protocol. His-tag was then removed by 3C protease, yielding a tag-free version of HUWE1 (HECT domain). (Figure 4-8)

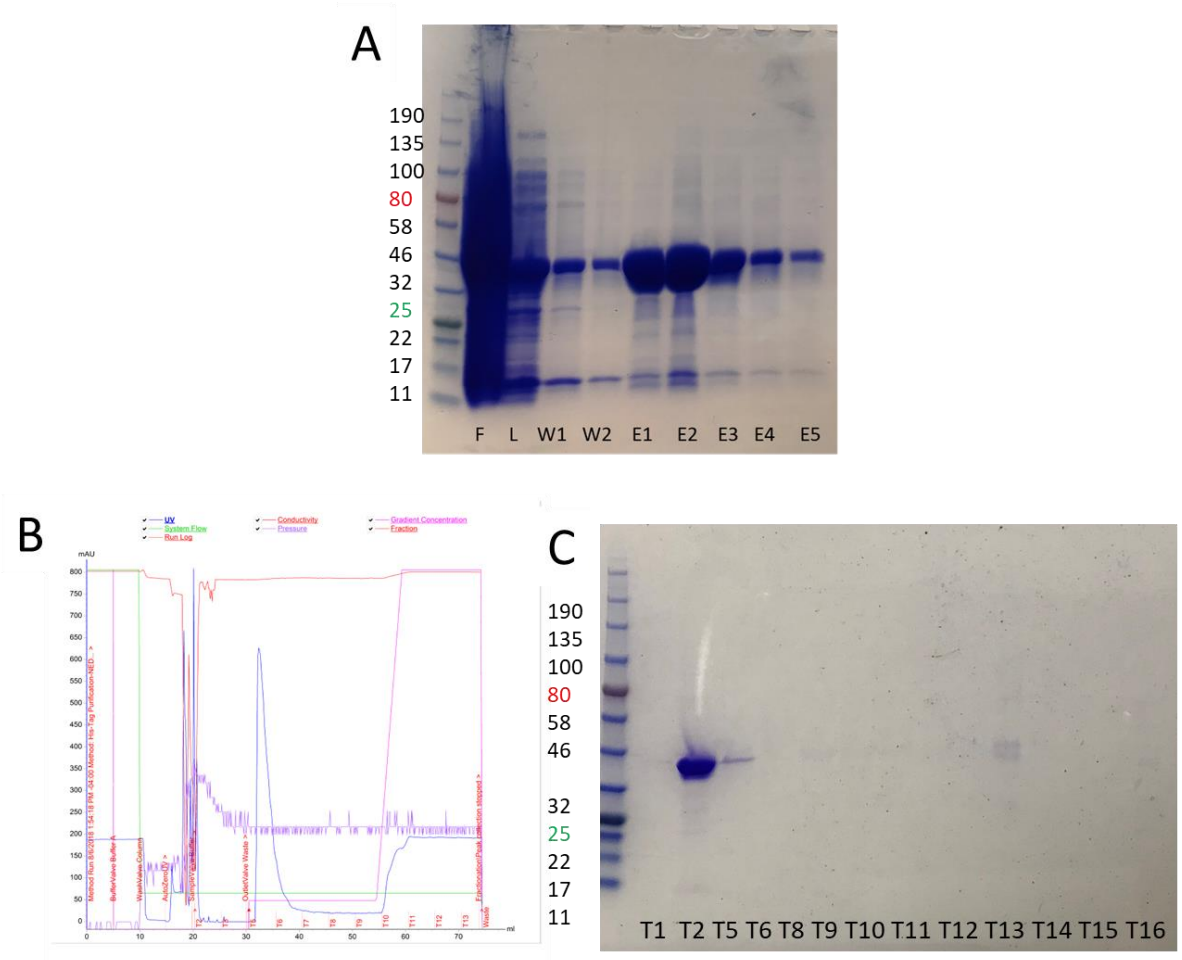
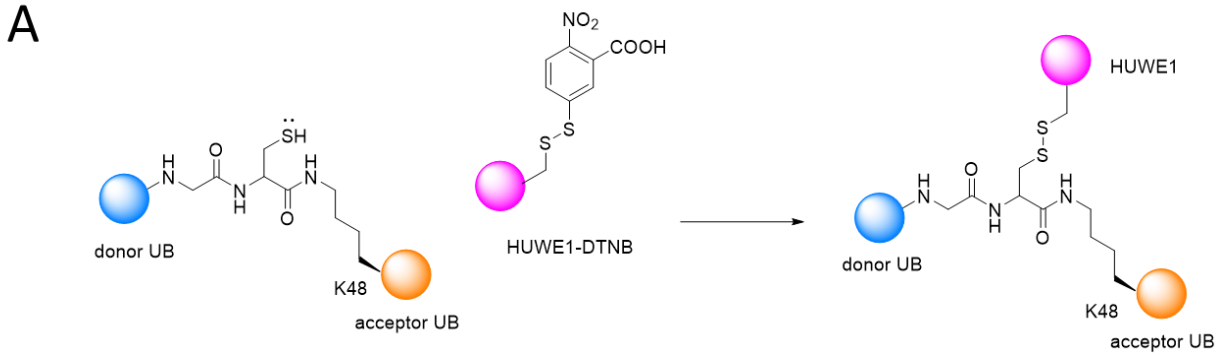


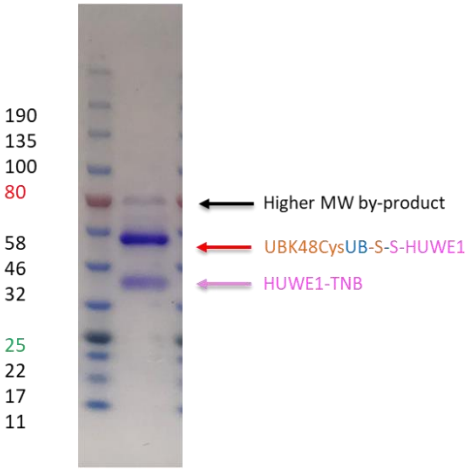
Figure 4-8 Purification of HUWE1 HECT domain

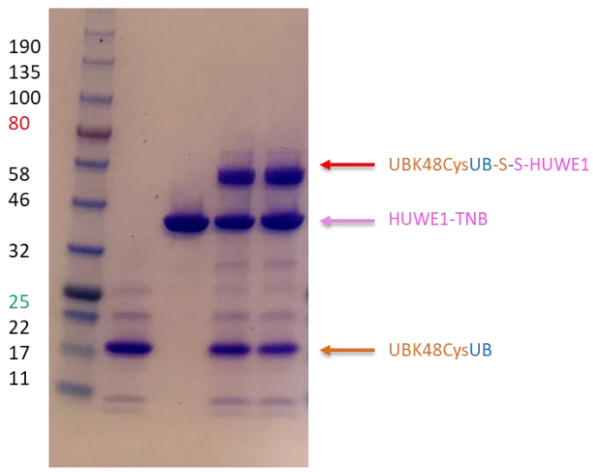
(A) HUWE1 HECT domain expression and purification. F: Flow-through; L: NTA-Lysis buffer wash; W1, W2: NTA-Wash buffer wash; E1-E5: NTA-elution buffer wash.
(B) AKTA chromatography of affinity purification after His-Tag remove with 3C Protease.
(C) Coomassie gel shows the peak in (B) is the correct unbound HUWE1 HECT domain without His-tag

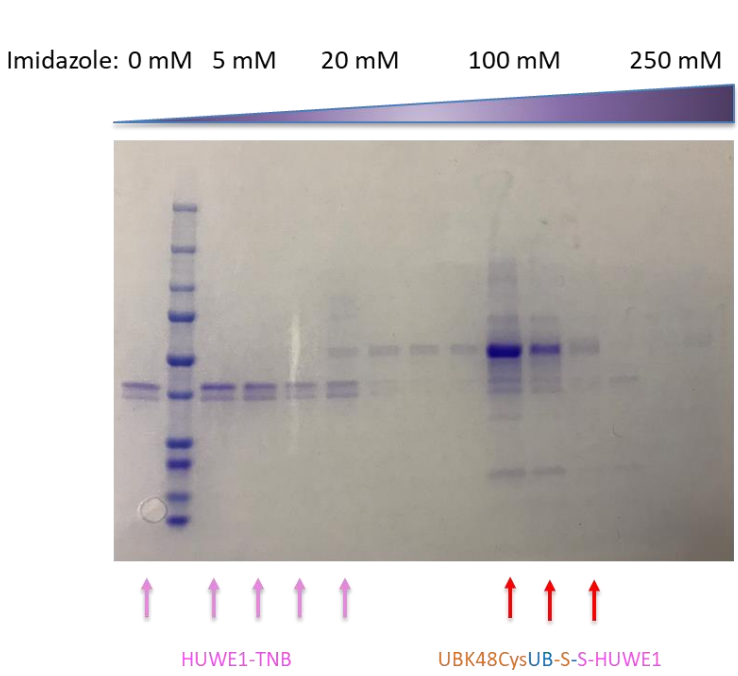
With tag-free HUWE1 in hand, similarly to disulfide formation between UBK11CysUB and Ube2S, DTNB pre-activation approach was utilized to generate UBK48CysUB-HUWE1. With a 2:1 ratio of HUWE1-TNB versus diUB-Cys, the desire tri-protein complex was readily formed in 45 min. However, there were additional higher molecular bands detected on Coomassie gel of the reaction mixture (Figure 4-9A, B), which suggested the additional conjugation could happen between diUB-Cys with two internal non-catalytic cysteines. Consequently, the rest two internal cysteines are mutated to alanine yielding single cysteine HUWE1 HECT domain with which no additional higher molecular weight bands were detected when conjugated with diUB-Cys. (Figure 4-9C) Following the same purification protocol as UBK11CysUB-S-S-Ube2S, the tandem purification delivered UBK48CysUB-S-S-HUWE1. (Figure 4-9D, E, F)

Generally, to consume 6mg purified diUB, over 20 mg tag-free HUWE1 will be needed and thus generate over 8 mg UBK48CysUB-S-S-HUWE1 after affinity purification. In the final step of size exclusion chromatography, 2-3 mg desired product was obtained with relatively pure quality. This approach has been investigated over 10 repeats with yield and purity consistent in between different diUB preps. This suggests that a streamline production, all the way from UAA synthesis and incorporation, acceptor UB deprotection and conjugation with donor UB, diUB coupling with HUWE1 HECT and tandem purification, has been achieved. This should provide enough material for crystallization condition screening and structure determination.


A
donor UB
K48
acceptor UB
HUWE1-DTNB
HUWE1
donor UB
K48
acceptor UB


B
190
135
100
80
58
46
32
25
22
17
11
Higher MW by-product
UBK48CysUB-S-S-HUWE1
HUWE1-TNB


C
190
135
100
80
58
46
32
25
22
17
11
UBK48CysUB-S-S-HUWE1
HUWE1-TNB
UBK48CysUB


D
Imidazole: 0 mM 5 mM 20 mM 100 mM 250 mM
HUWE1-TNB
UBK48CysUB-S-S-HUWE1

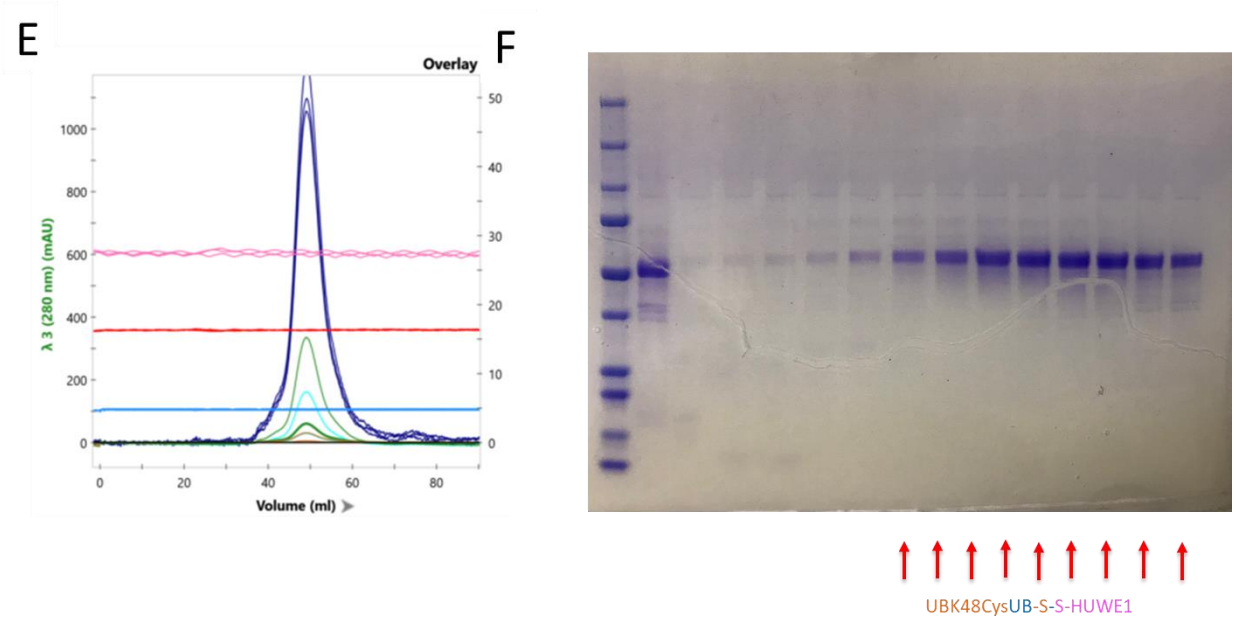


Figure 4-9 Reaction and purification of UBK48CysUB-S-S-HUWE1

(A) Reaction between UBK48CysUB and DTNB pre-activated HUWE1
(B) Coomassie gel shows the disulfide formation of UBK48CysUB and HUWE1 with two extra non-catalytic cysteines. Note that there was extra higher molecular weight by-product.
(C) Coomassie gel shows the disulfide formation of UBK48CysUB and HUWE1 with only catalytic cysteine
(D) Affinity purification of UBK48CysUB-S-S-HUWE1
(E) (F) Diagram and Coomassie gel of size exclusion to further purify UBK48CysUB-S-S-HUWE1

For initial screening for crystallization conditions, the purified protein complex was firstly concentrated as much as possible. (Figure 4-10A) According to the concentration curve in terms of volume. The final concentration of UBK48CysUB-S-S-HUWE1 complex could reach over 10 mg/mL, which was used as stock solution for screening. Several standard sparse matrix screenings were investigated: JCSG-plus from Molecular Dimensions, Index from Hampton Research, Classics I from Qiagen; the temperature of incubating these plates was varied: 20 ºC or 4 ºC. Unfortunately, no crystal was observed under these conditions, except for a couple of salt crystals. (Figure 4-10B, C, D)

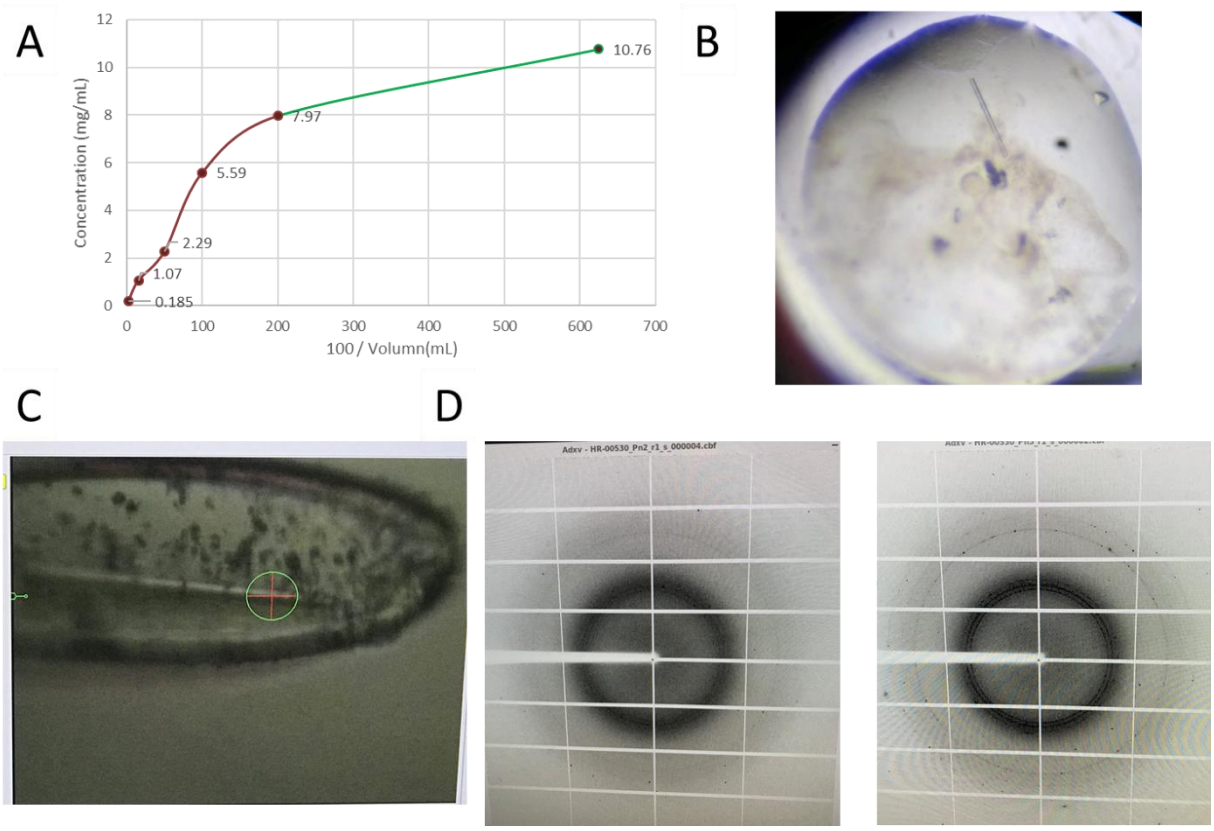


Figure 4-10 Preliminary screening for crystallization conditions for UBK48CysUB-S-S--HUWE1

(A) The concentration of UBK48CysUB-S-S-HUWE1 can reach over 10 mg/mL.
(B) A (salt) crystal observed from JCSG-plus plate
(C) The crystal in (B) was mounted and subjected to X-ray diffraction.
(D) The X-ray diffraction pattern for crystal in (B): it turned out to be salt crystals.

As a future direction, the purity of tri-protein complex can be increased by adding two more steps of anion-exchange chromatography to purify HECT E3 and diUB after affinity purification respectively. In addition, the N-terminal His-HA tag on the acceptor UB could perturb the integrity of protein complex by prohibiting potential interactions between acceptor UB and HUWE1 or introducing non-native interferences with any of the other proteins. Therefore, a TEV cleavage site followed by a linker sequence (GSGGS) was introduced in between the His-HA tag and acceptor UB sequence. After the affinity purification of UBK48CysUB-S-S-HUWE1 and before the final size exclusion step, TEV was added to cleave the tag off so that all three

components are free of any tag. As a matter of fact, the amount of TEV needed to cleave the N-terminal tag off the acceptor UB was a lot more than common equivalency, maybe suggesting that an additional binding of His-HA tag with the protein complex may exist and block the access of TEV to bind to the cleavage site.

4.2.2 NEDD4-1

NEDD4-1 is one of the most studied HECT-type E3 ligase, which belongs to NEDD4 subfamily HECT E3s. Overexpression of NEDD4 has been reported in several cancer types and downregulation of NEDD4 appears to reduce proliferation, migration and invasion of cancer cells.[229] Importantly, tumor suppressor PTEN is discovered to be the substrate of NEDD4.[230] Different from E6AP or HUWE1, NEDD4 mainly generate K63-linked UB chains. [216] The mechanism of NEDD chain specificity was partially explained by the crystal structure of donor UB-S-S-NEED4 (Figure 4-11): the donor UB, transferred from the E2 (UbcH5B), is bound to the

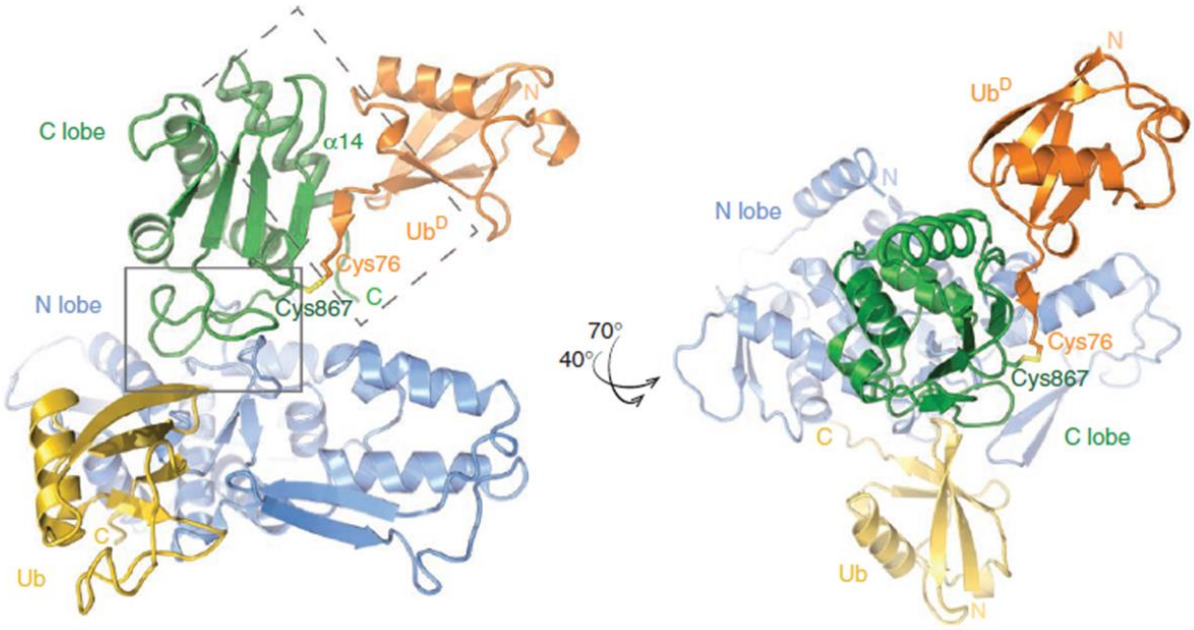
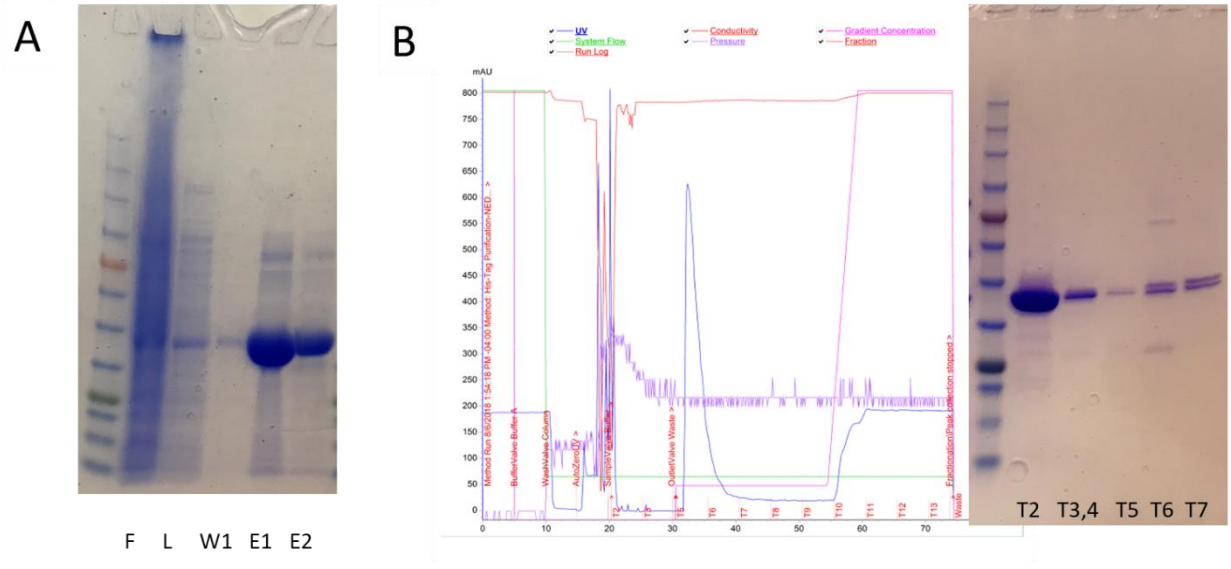


Figure 4-11 Crystal strcuture of Donor UB-S-S-NEDD4 (HECT) with another non-covalent bound UB (PDB:4BBN)

Nedd4 C lobe with its C-terminal tail locked in an extended conformation, primed for catalysis; such stretched conformation of donor UB appears to be mutual for a priming state of E2 or E3s.[231, 232] Although in this crystal structure, a non-covalent UB bound to UBD of NEED4-1 was observed this UB didn't present the acceptor UB carrying the attacking lysine. Nevertheless, the authors concluded the last 60 amino acids of HECT domain are responsible for both donor and acceptor UB binding to determine chain specificity, however, direct evidence is missing.

There are three cysteines in NEDD4-1 with two non-catalytic cysteine and one catalytic cysteine. With the other two cysteine mutated into alanine, NEDD4-1 was expressed with BL21(DE3) cells with decent yield (Figure 4-12A) and a tag-free NEDD4-1 HECT domain was obtained after a similar cleavage with 3C protease like HUWE1. (Figure 4-12B) The reaction between UBK63CysUB and NEDD4 was successful with a ratio 1:2. (Figure 4-12C, D) Notably, NEDD4 formed a homodimer during the disulfide formation as a by-product, indicating NEDD4 features a more reactive cysteine. So far, it has been shown that UBK63CysUB-S-S-NEDD4 can be purified through affinity purification. (Figure 4-12E)



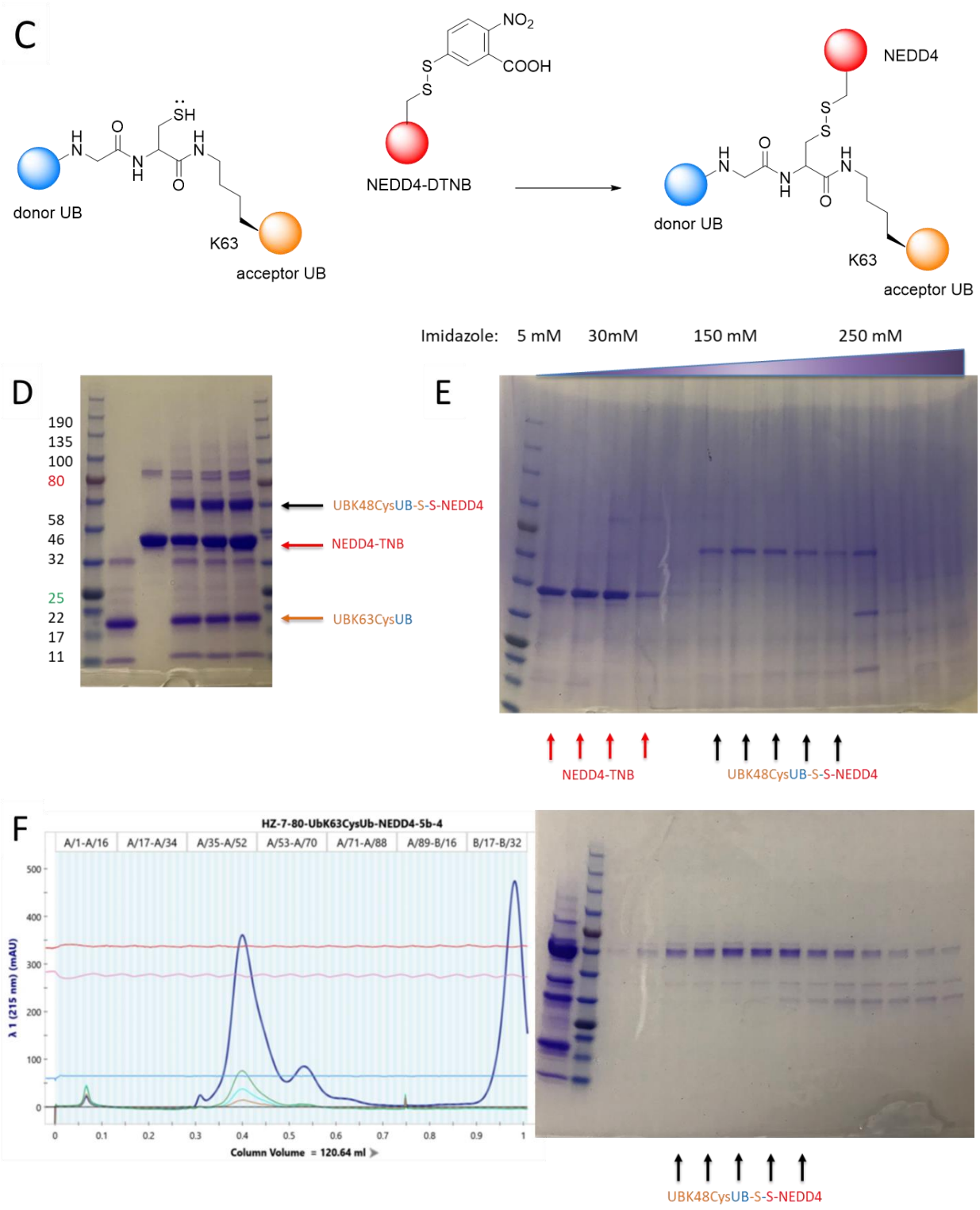


Figure 4-12 Synthesis and purification of UBK63CysUB-S-S-NEDD4

(A) Protein expression and purification of NEDD4 HECT domain. (B) AKTA affinity purification of tag-free NEED4 HECT domain after treated with 3C Protease, following by Coomassie staining. (C) (D) Reaction between UBK63CysUB and NEDD4. (E) Affinity purification UBK63CysUB-S-S-NEDD4 (F) Size exclusion of UBK63CysUB-S-S-NEDD4.

4.3 Probing E2 chain specificity with Di-Ubiquitin, RING E3s and substrates

RING E3s don't carry a catalytic cysteine that can receive the charged UB from E2 and pass on the UB to substrate like HECT type E3s. However, RING E3 help to recruit both UB~E2 and substrate in close proximity to facilitate the substrate ubiquitination and chain elongation. Most of the E2s cannot elongate specific chains without corresponding E3s; determine the structure of diUB-S-S-E2 complex together with RING E3s and substrates should reveal the whole picture of how E2 and RING E3 arrange ubiquitinated substrate as an acceptor as well as donor UB to achieve linkage-specific UB chain elongation.

4.3.1 UBK11CysUB-Ube2S-APC/C-substrate complex

While we are making linkage-specific diUB featuring a thiol as reactive warhead, the field is moving forward fast. Recent papers[233-235] revealed the mechanism underlying the initiation of ubiquitination on APC/C substrate by Ube2C and chain elongation on ubiquitinated substrate by APC/C and Ube2S by solving the Cryo-EM structure of substrate-acceptor UB-donor UB-Ube2S/Ube2C (full length)-APC/C complex. (Figure 4-13) However, in order to generate a such complicated protein complex, a few compromises had to be made: the acceptor UB had to be replaced by a UB variant (UBv) bearing 17 mutations selected from a phage library so that the acceptor UB can be anchored onto the APC11's RING with increased affinity (1.6 μM K_D); the substrate peptides had to be fused to N-terminus or C-terminus of acceptor UB to mimic the ubiquitinated substrate; and most importantly, donor UB, acceptor UB fused with substrate peptide, and E2 (Ube2S or Ube2C) were crosslinked together with 3-headed sulfhydryl-reactive TMEA. (Figure 4-14) to mimic the transient UB transfer intermediate. As a result, catalytic core, RING-UBv and E2 portions of the maps displayed local resolution ~6-10 Å with an overall resolution 6

Å; the donor UB was not visible even at low contour in either complex, suggesting the local variability or the dynamics of the "closed" E2~UB conjugate in solution. We reasoned a better-designed acceptor UB-donor UB-Ube2S could provide a detailed mechanism of K11 chain elongation by Ube2S by zooming in the catalytic core with closer mimics. (Figure 4-14) To this end, a collaboration was set up immediately with Dr. Nicolas Brown in University of North Carolina at Chapel Hill, who published the Cryo-EM structures earlier; we will generate protein complex shown in Figure 4-13C and he will put it along with APC/C components and determine the structure by Cryo-EM.

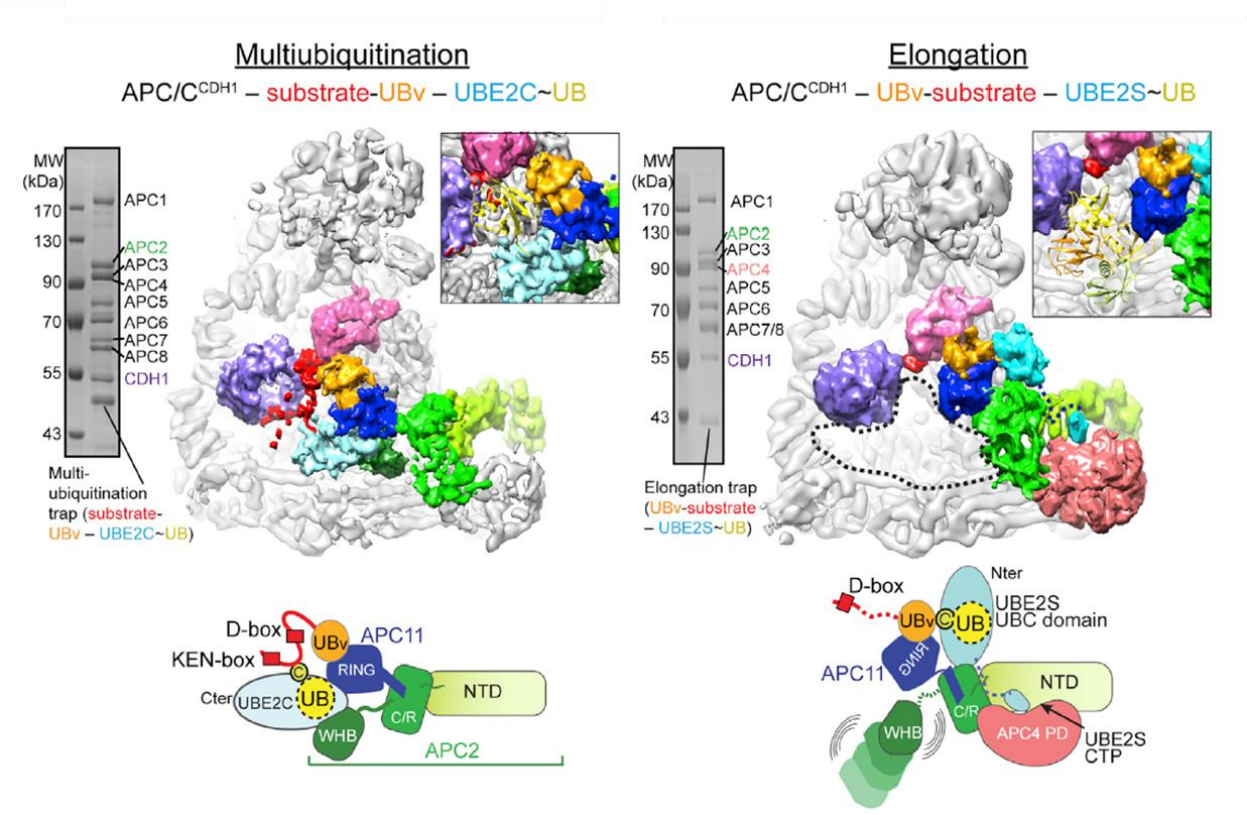
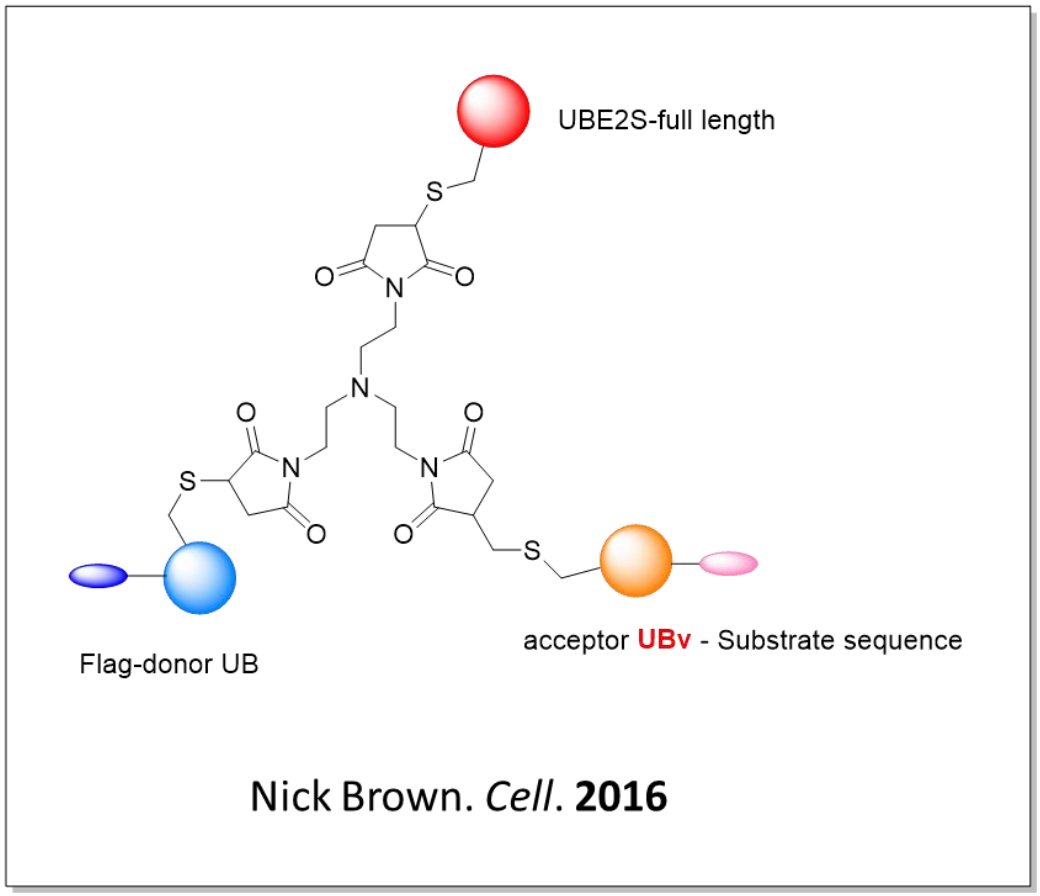
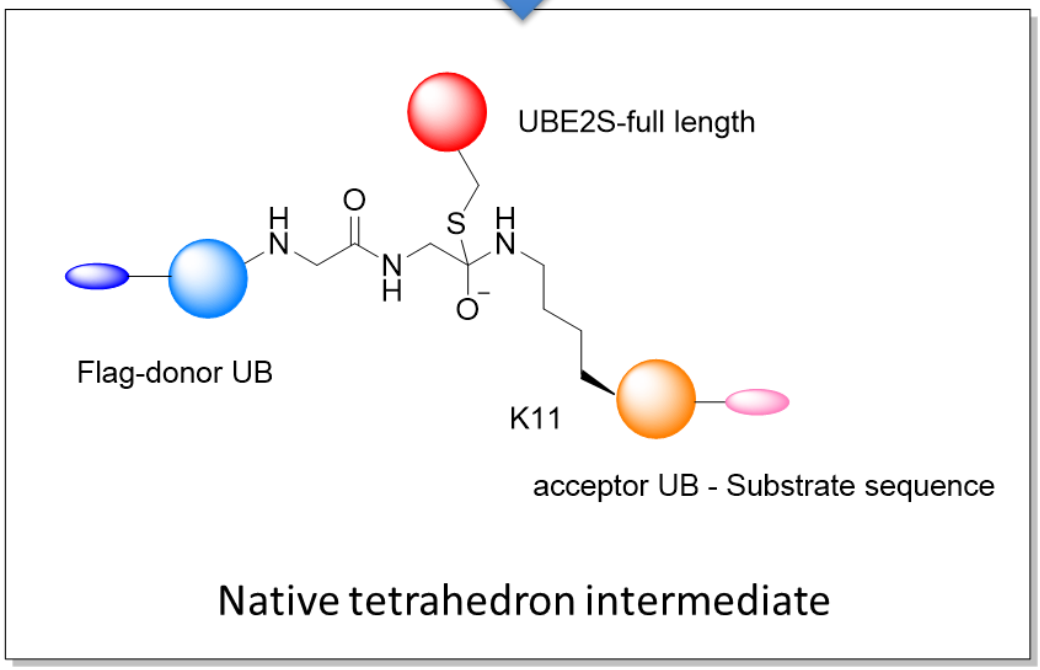


Figure 4-13 Cryo-EM structures of Ube2C/Ube2S and APC/C mediating ubiquitin initiation and elongation on the substrate.


UBE2S-full length
Flag-donor UB
acceptor UBv - Substrate sequence
Nick Brown. Cell. 2016


UBE2S-full length
Flag-donor UB
K11
acceptor UB - Substrate sequence
Native tetrahedron intermediate

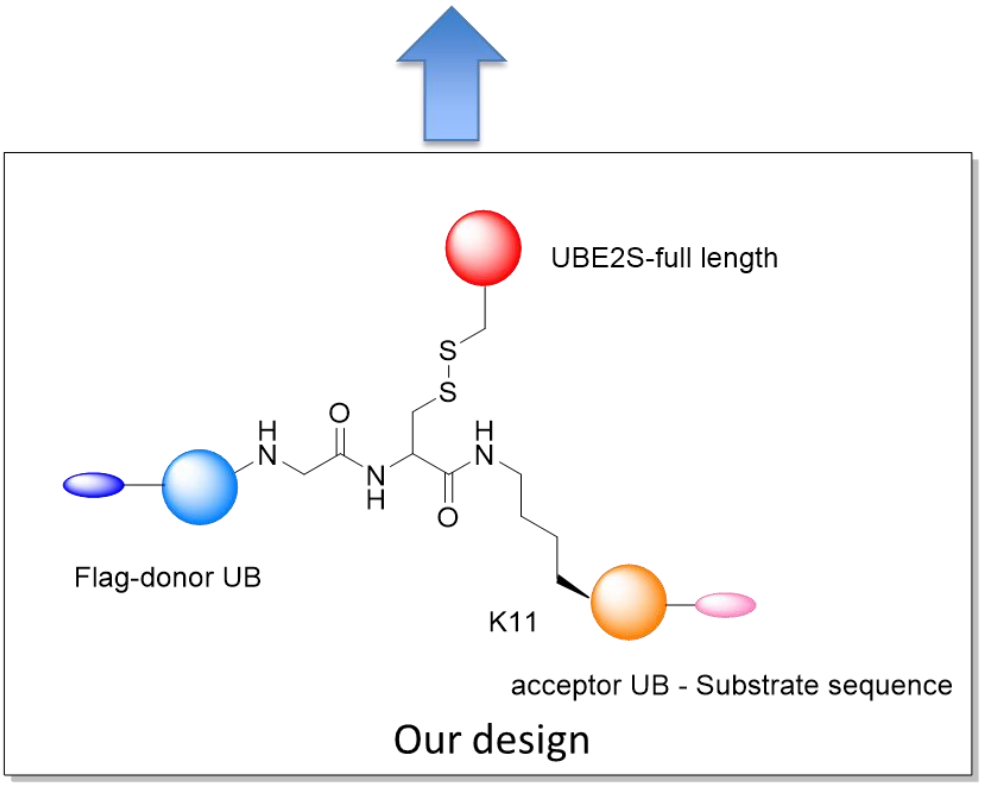


Figure 4-14 Comparison of designs to crosslink donor UB-acceptor UB-Ube2S

To generate this newly designed protein complex, a fusion protein comprised of an N-terminal substrate peptide and acceptor UB with unnatural amino acid ThzK incorporated at K11 position was made and well characterized by ESI-POS (Figure 4-15A, B). Then a Flag-tagged donor UB was made and conjugated to the acceptor UB. (Figure 4-15C, D) The next step is scaling up the diUB synthesis so that it can conjugated to full length Ube2S; the flexible 66-residue C-terminal extension of Ube2S cannot be crystalized but is very crucial for anchoring itself onto APC/C complex.[203]

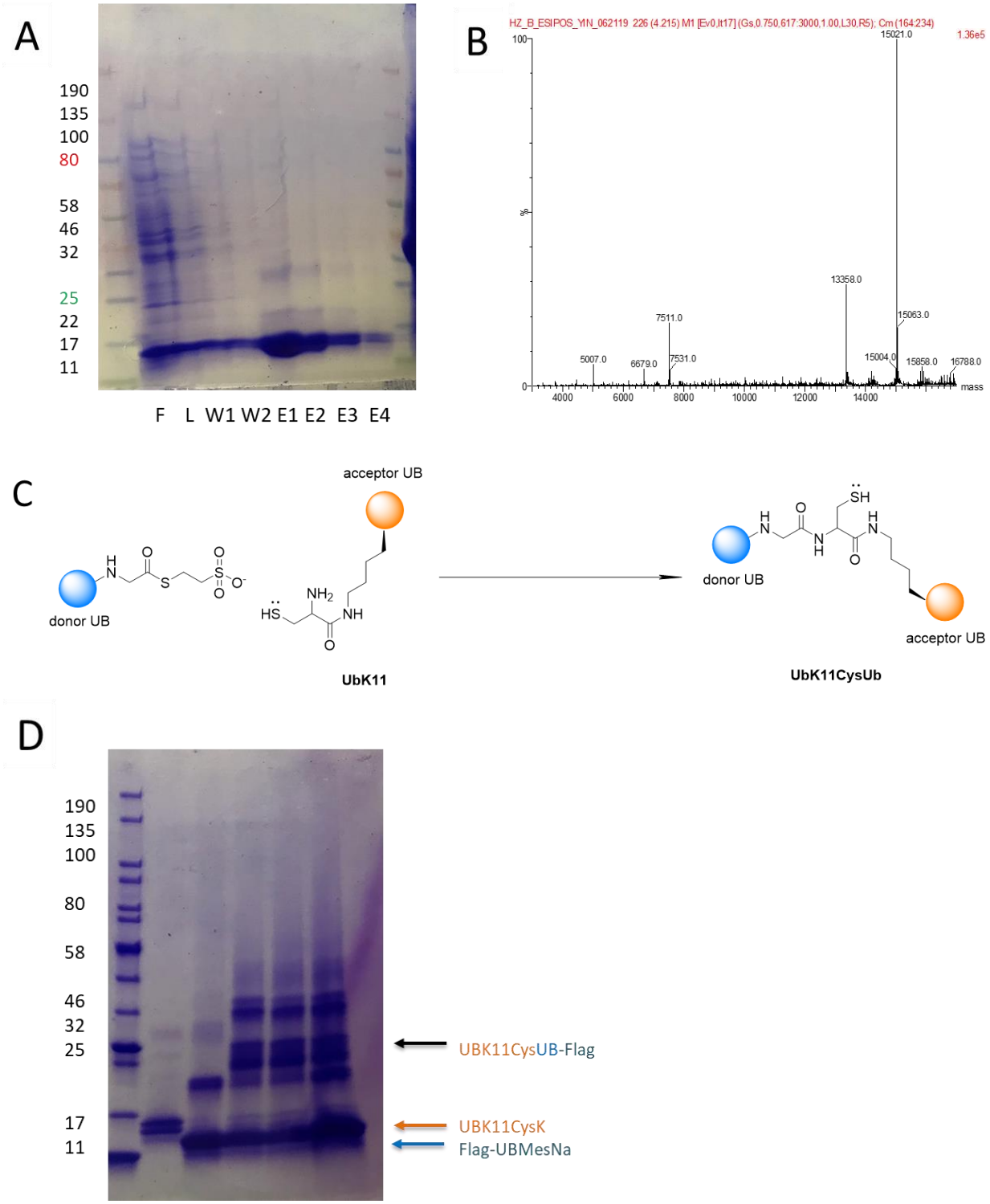


Figure 4-15 Preparation of UBK11CysUB fused with APC/C substrate peptide sequence

(A) ThzK incorporation of UBK11ThzK-HSL1
(B) ESI (+) shows the correct molecular weight of UBK11CysK-HSL1, calculated MS 15022.76
(C) Reaction between Donor UB-MesNa and acceptor UBK11CysK-HSL1
(D) The ligation reaction was analyzed by SDS-PAGE coupled with Coomassie staining.

4.3.2 UBK48CysUB-UBE2G1-Cul4-RBX1-DDB1-SV5 complex

Cullin-Ring E3 ligases[236, 237] are a superfamily of RING E3s responsible for as much as 20% of ubiquitin-dependent protein turn-over in cells.[238] The basal state of a CRL complex as being composed of a cullin, a RING finger protein (RBX1 or RBX2) that serves as the site for E2 binding and ubiquitin transfer, and an SR (substrate receptors) module, which specifically recruit a target protein. CRL relies a specific E2, CDC34 (Ube2R1), to assembly K48-linked UB chains. The chain specificity of Cdc34 associated with CRL was in part explained by kinetic studies and site-mutagenesis[69], suggesting residues near CDC34 activate site are complementary to the surface surrounding K48 in acceptor UB. In addition, a recent study of donor UB forming a disulfide with Cdc34 catalytic cysteine[239], revealing key interactions between them. However, detailed information is still missing for acceptor UB recognition; and how Cdc34 (UBE2G1), along with its CRLs, mediates the UB chain initiation and elongation on their substrates.

To stress this issue, we are collaborating with Dr. Ning Zheng from University of Washington who solved lots of structures of CRLs and their substrates.[240-244] Based upon the determined structure of UBE2G1-Cul4-RBX1-DDB1-SV5 (Figure 4-16A, unpublished data from Zheng lab), we have design the structure of UBK48CysUB-S-S-UBE2G1-SV5 (Figure 4-16B) to fill in the gap of the whole picture for SV5 ubiquitination and chain elongation. This time, instead of fusion the substrate SV5 to the N-terminus of acceptor UB. A closer mimic of triazole linkage introduced by click chemistry was designed to represent the isopeptide bond between SV5 and acceptor UB: the azide group on C-terminus of acceptor UB for this purpose was generated with intern-mediated expressed ligation; the propargyl lysine was incorporated into the substrate with genetic code expansion. (Scheme 4-1)

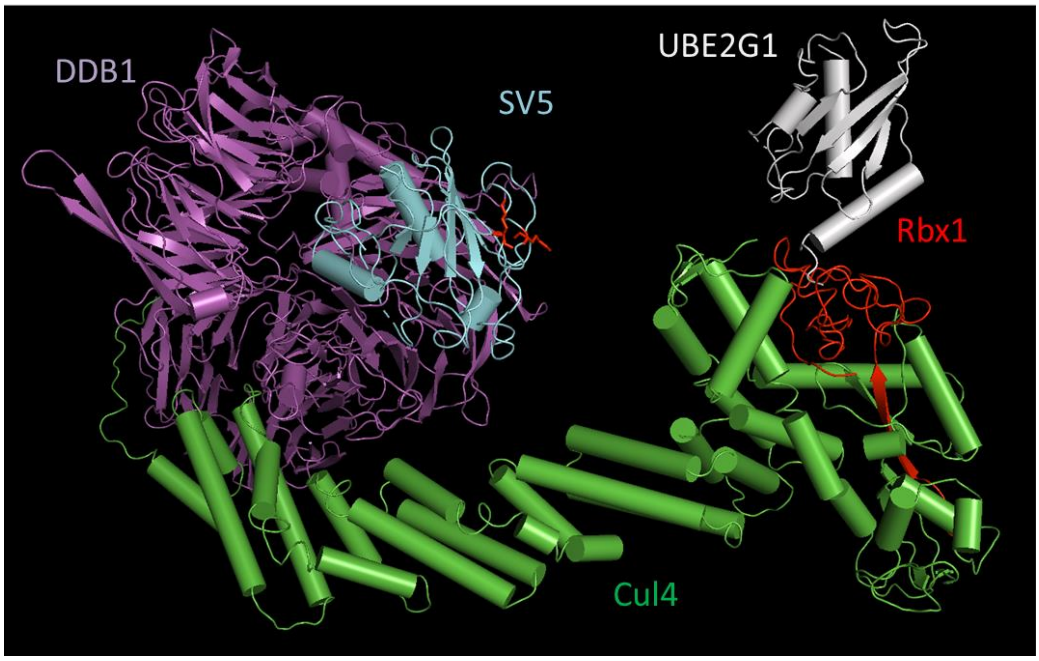


Figure 4-17 Structure of Cul4-DDB1-SV1-RBX1-UBE2G1

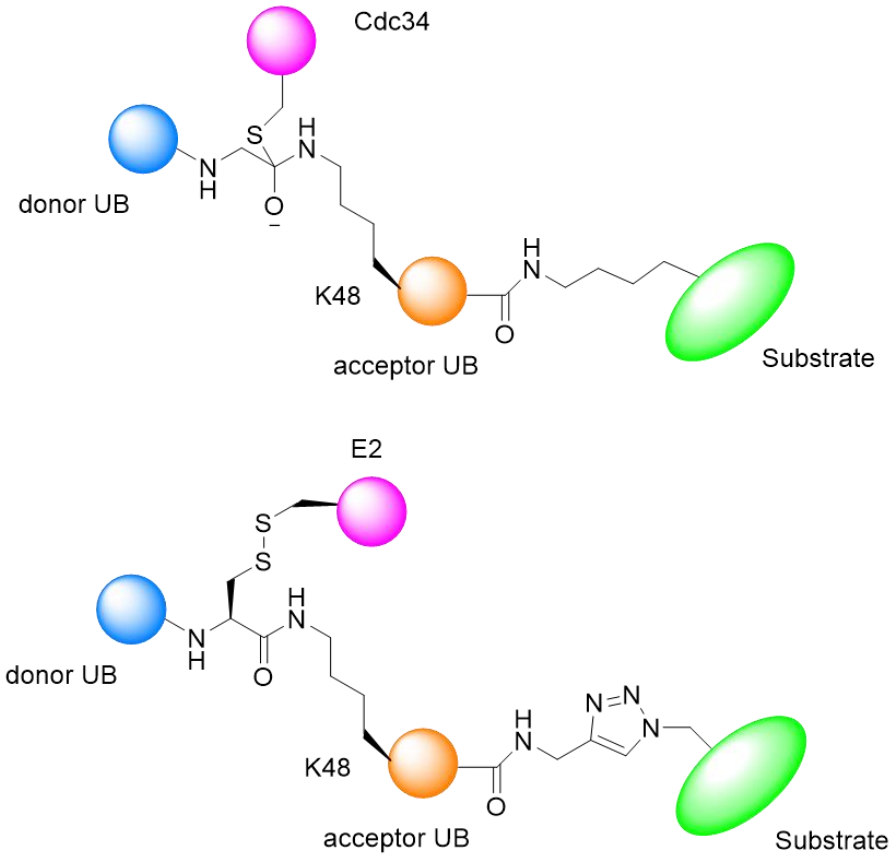
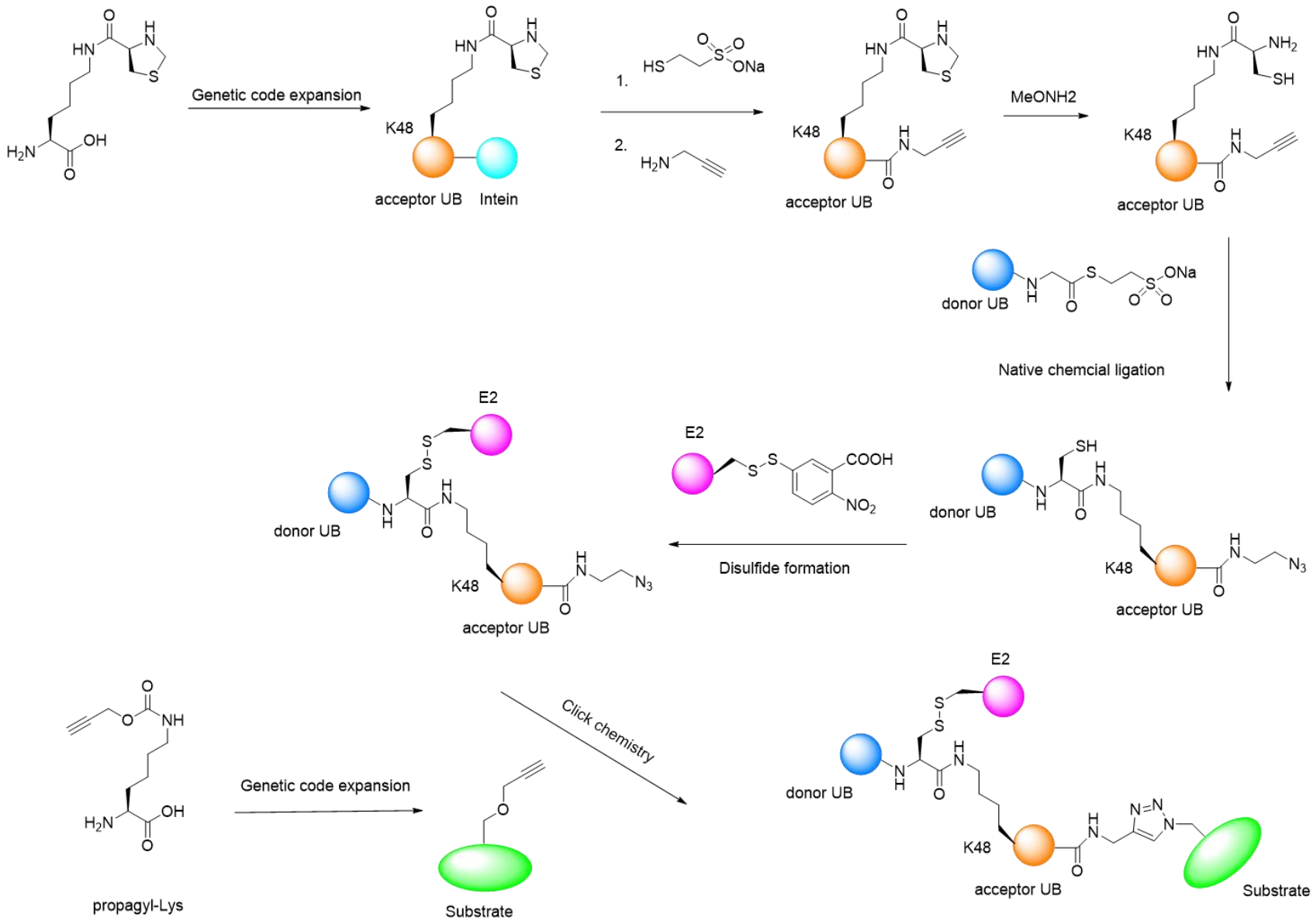
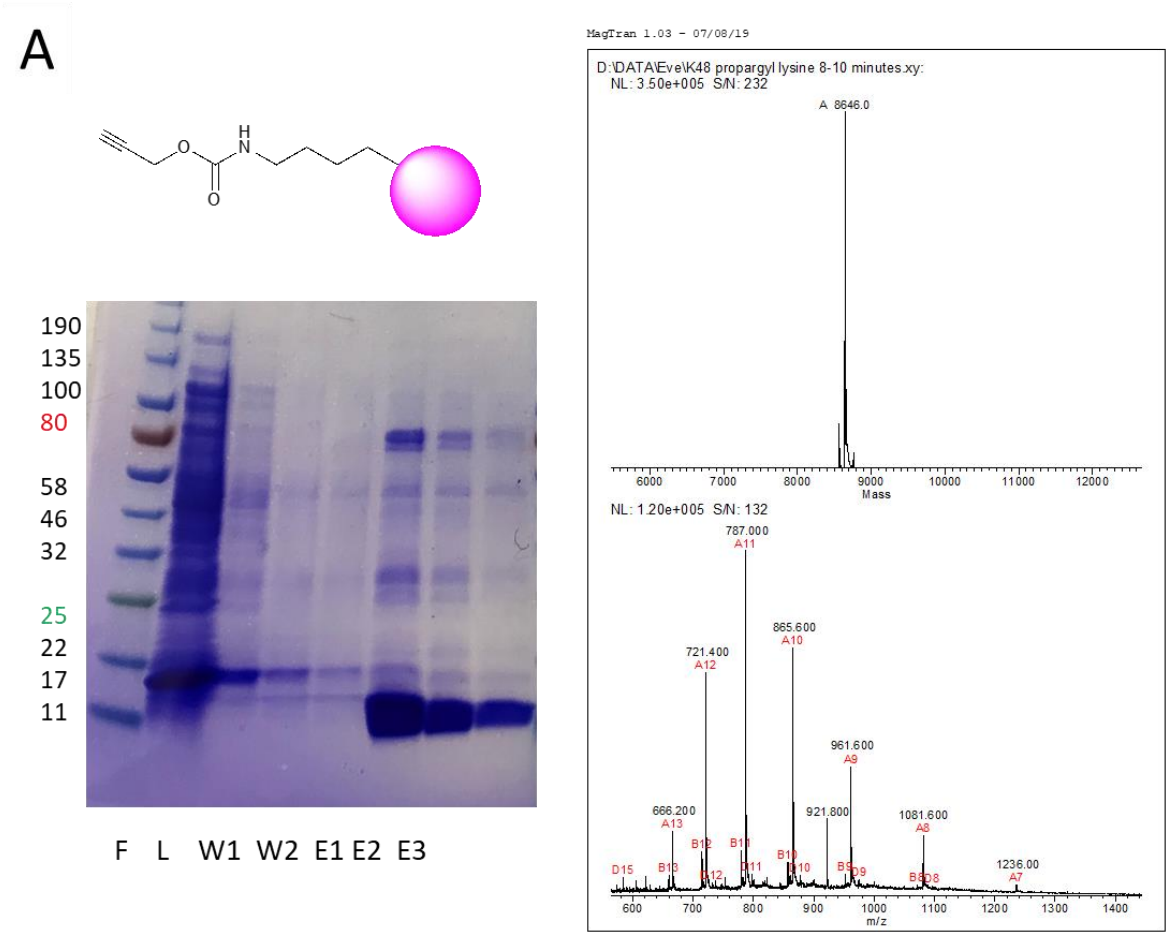


Figure 4-16 Design of donor UB- acceptor UB-Cdc34 (UBE2G1)-Substrate complex



Scheme 4-1 Synthetic route for UBK48CysUB-UBE2G1-SV5

Here in this thesis, three pilot reactions are demonstrated, which are three new yet crucial steps for the donor UB- acceptor UB-UBE2G1-SV5 complex design. First one is incorporating propargyl lysine into Ubiquitin with decent yield (Figure 4-18A); another one is generating a C-terminal azide utilizing intein-mediated expressed ligation (Figure 4-18B); and the last one is click reaction between donor UB with an azide and acceptor UB bearing an alkyne, catalyzed by Copper (I) and Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) (Figure 4-18C).The success of these experiments suggest the design is reasonable and feasible.



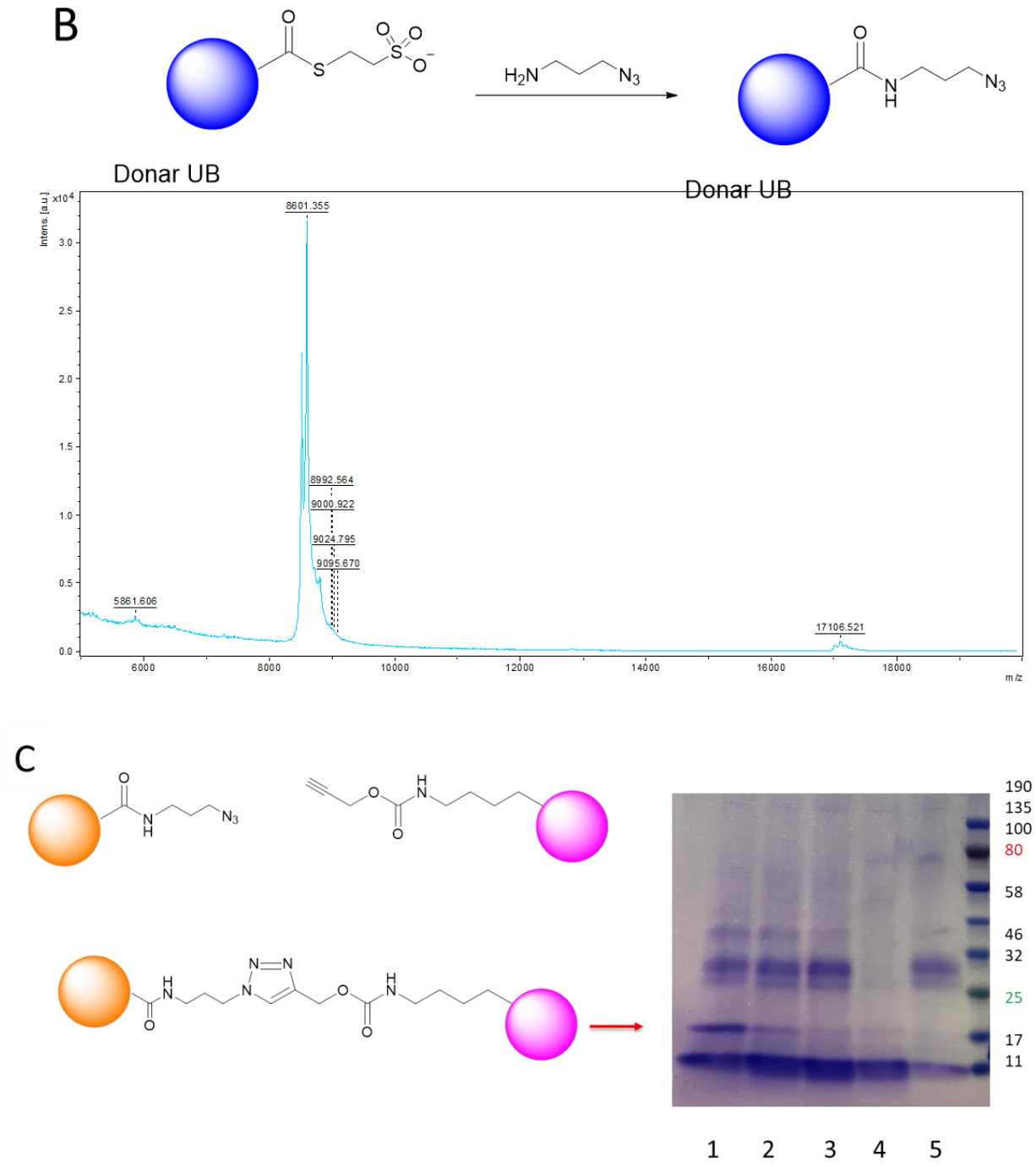


Figure 4-18 Pilot reactions for donor UB-acceptor UB-UBE2G1-SV5 design

(A) Incorporation of propargyl lysine into UB, characterized by Coomassie staining and ESI (calculated MW 8645.0)
(B) Generation of C-terminal azido UB via intein-mediated expressed ligation, characterized by MALDI-LP (calculated MW 8589.9)
(C) Click reaction between UB-N_3 and UB-alkyne, monitored by SDS-PAGE: Lane 1 Reaction with two components and Cu(I), THPTA; Lane 2, 3, 4, 5: control reactions omitting THPTA, Cu(I), UB-N_3,UB-alkyne respectively.

4.4 Conclusion

To summarize, three diUB with different linkage types were made, namely K11, K48 and K63. These diUB features a thiol as a reactive warhead embedded near the isopeptide site that can form a three-component complex with Ube2S, $HECT^{HUW1}$, $HECT^{NEDD4}$ respectively through disulfide conjugation upon DTNB pre-activation. A streamline of production of these complexes had been realized with milligram level yield, hopefully can facilitate more crystallization condition screening and eventually solve the structures. By doing this, the detailed mechanism and direct evidence of UB transfer during the chain elongation by linkage-specific E2 or HECT-type E3s are expected to be revealed.

Nevertheless, by collaborating with crystallographers, donor UB- E2s- acceptor UB- covalently crosslinked with substrate or substrate peptide sequence is rationally designed and readily made, hoping to fit in the big picture of how linkage-specific E2s, associating with giant RING E3s such as APC/C or CRLs, initiate the ubiquitin on a certain substrate and further elongate UB chains in a linkage-specific manner.

Chemical biology expands the toolbox of chemists and biologists and empower them to address complicated questions that cannot be solved with traditional biochemistry approaches. For instance, these diUB probes with customized designs and modifications, should provide access for structural biologists to map dynamic, transient, weak protein-protein interactions in UB transfer, therefore to uncover detailed mechanism underlying chain specificities of E2 or HECT E3 with direct evidences and to reconstitute full pictures of E2, E3 mediated ubiquitination on substrates and linkage-specific chain elongation.

4.5 Materials and method

4.5.1 Protein expression and purification

Ube2S (UBC domain) was expressed with BL21(DE3) cells as a His-SUMO-Ube2S fusion and purified with a standard Ni-NTA affinity purification. Then Ulp1 was selected to cleave the SUMO tag off and tag-free Ube2S was purified with AKTA HisTrap affinity purification, followed by anion-exchange chromatography.

HUWE1 and NEDD4 HECT domain were expressed with BL21(DE3) cells with an N-terminal His tag. Upon purification with Ni-NTA, 3C protease was selected to cleave the His tag off and tag-free HUWE1 or NEDD4 was purified with AKTA Histrap affinity purification, followed by anion-exchange chromatography.

4.5.2 Disulfide formation

Disulfide formation between diUB-Cys and E2 or E3s followed a reported protocol[245].

In short, diUB and E2 or E3 with a single cysteine was pretreated with 1 mM TCEP (pH=8.0) and incubated at 37 °C for half an hour to fully reduce the disulfide. Then diUB and E2 or E3 were desalted with PD-10 columns in a crosslink buffer (25mM Tris, pH=8.0, 50 mM NaCl). DiUB was stored on ice until use. E2 or E3 were then added to a freshly made “activation” buffer. (25 mM Tris, pH=8.0, 50mM NaCl, 2.5 mM 2,2’- Dipyridyldisulfide, 2.5% DMSO) and incubated at 25°C for half an hour. The resulted yellow solution was then concentrated down to proper volume and pass through PD-10 desalting column again to remove DTNB in excess. The concentration of both diUB and E2 or E3 were measured by BSA assay and the disulfide bond was formed by incubated 1 equiv. of diUB along with 2-4 equiv. of E2 or E3. After 1 hour, the reaction mixture was desalted with PD-10 column and treated with Ni-NTA resin for affinity purification.

REFERENCES

1. Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J., Jr., Protein posttranslational modifications: the chemistry of proteome diversifications. *Angewandte Chemie (International ed. in English)* **2005,** *44* (45), 7342-72.
2. Walsh, C. T., *Posttranslational modification of proteins*. Roberts and Company, Englewood: 2006.
3. Ciechanover, A.; Heller, H.; Elias, S.; Haas, A. L.; Hershko, A., ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. *Proc Natl Acad Sci U S A* **1980,** *77* (3), 1365-8.
4. Ciechanover, A.; Hod, Y.; Hershko, A., A heat-stable polypeptide component of an ATPdependent proteolytic system from reticulocytes. *Biochem. Biophys. Res. Commun.* **1978,** *81*, 1100-1105.
5. Hershko, A.; Ciechanover, A.; Rose, I. A., Resolution of the ATP-dependent proteolytic system from reticulocytes: a component that interacts with ATP. *Proc Natl Acad Sci U S A* **1979,** *76* (7), 3107-10.
6. Wilkinson, K. D.; Urban, M. K.; Haas, A. L., Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *J Biol Chem* **1980,** *255* (16), 7529-32.
7. Frescas, D.; Pagano, M., Deregulated proteolysis by the F-box proteins SKP2 and beta-TrCP: tipping the scales of cancer. *Nat Rev Cancer* **2008,** *8* (6), 438-49.
8. Hoeller, D.; Dikic, I., Targeting the ubiquitin system in cancer therapy. *Nature* **2009,** *458* (7237), 438-44.
9. Lipkowitz, S.; Weissman, A. M., RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nat Rev Cancer* **2011,** *11* (9), 629-43.
10. Nakayama, K. I.; Nakayama, K., Ubiquitin ligases: cell-cycle control and cancer. *Nat Rev Cancer* **2006,** *6* (5), 369-81.
11. Schwartz, A. L.; Ciechanover, A., Targeting proteins for destruction by the ubiquitin system: implications for human pathobiology. *Annu Rev Pharmacol Toxicol* **2009,** *49*, 73-96.
12. Deshaies, R. J.; Joazeiro, C. A., RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **2009,** *78*, 399-434.
13. Hatakeyama, S.; Nakayama, K. I., U-box proteins as a new family of ubiquitin ligases. *Biochem Biophys Res Commun* **2003,** *302* (4), 635-45.
14. Rotin, D.; Kumar, S., Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol* **2009,** *10* (6), 398-409.
15. Schulman, B. A.; Harper, J. W., Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* **2009,** *10* (5), 319-31.
16. Ye, Y.; Rape, M., Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* **2009,** *10* (11), 755-64.
17. Harper, J. W.; Tan, M. K., Understanding cullin-RING E3 biology through proteomics-based substrate identification. *Mol Cell Proteomics* **2012,** *11* (12), 1541-50.
18. Swatek, K. N.; Komander, D., Ubiquitin modifications. *Cell research* **2016,** *26* (4), 399-422.
19. Kwon, Y. T.; Ciechanover, A., The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends in biochemical sciences* **2017,** *42* (11), 873-886.
20. Komander, D.; Rape, M., The ubiquitin code. *Annual review of biochemistry* **2012,** *81*, 203-29.
21. Ohtake, F.; Saeki, Y.; Sakamoto, K.; Ohtake, K.; Nishikawa, H.; Tsuchiya, H.; Ohta, T.; Tanaka, K.; Kanno, J., Ubiquitin acetylation inhibits polyubiquitin chain elongation. *EMBO reports* **2015,** *16* (2), 192-201.

22. Koyano, F.; Okatsu, K.; Kosako, H.; Tamura, Y.; Go, E.; Kimura, M.; Kimura, Y.; Tsuchiya, H.; Yoshihara, H.; Hirokawa, T.; Endo, T.; Fon, E. A.; Trempe, J. F.; Saeki, Y.; Tanaka, K.; Matsuda, N., Ubiquitin is phosphorylated by PINK1 to activate parkin. *Nature* **2014,** *510* (7503), 162-6.
23. Sauve, V.; Sung, G.; Soya, N.; Kozlov, G.; Blaimschein, N.; Miotto, L. S.; Trempe, J. F.; Lukacs, G. L.; Gehring, K., Mechanism of parkin activation by phosphorylation. *Nature structural & molecular biology* **2018,** *25* (7), 623-630.
24. Gladkova, C.; Maslen, S. L.; Skehel, J. M.; Komander, D., Mechanism of parkin activation by PINK1. *Nature* **2018,** *559* (7714), 410-414.
25. 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.; Gygi, S. P., Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular cell* **2011,** *44* (2), 325-40.
26. Wagner, S. A.; Beli, P.; Weinert, B. T.; Nielsen, M. L.; Cox, J.; Mann, M.; Choudhary, C., A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Molecular & cellular proteomics : MCP* **2011,** *10* (10), M111.013284.
27. Xu, P.; Duong, D. M.; Seyfried, N. T.; Cheng, D.; Xie, Y.; Robert, J.; Rush, J.; Hochstrasser, M.; Finley, D.; Peng, J., Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell* **2009,** *137* (1), 133-45.
28. Dammer, E. B.; Na, C. H.; Xu, P.; Seyfried, N. T.; Duong, D. M.; Cheng, D.; Gearing, M.; Rees, H.; Lah, J. J.; Levey, A. I.; Rush, J.; Peng, J., Polyubiquitin linkage profiles in three models of proteolytic stress suggest the etiology of Alzheimer disease. *The Journal of biological chemistry* **2011,** *286* (12), 10457-65.
29. Ziv, I.; Matiuhin, Y.; Kirkpatrick, D. S.; Erpapazoglou, Z.; Leon, S.; Pantazopoulou, M.; Kim, W.; Gygi, S. P.; Haguenauer-Tsapis, R.; Reis, N.; Glickman, M. H.; Kleifeld, O., A perturbed ubiquitin landscape distinguishes between ubiquitin in trafficking and in proteolysis. *Molecular & cellular proteomics : MCP* **2011,** *10* (5), M111.009753.
30. Bremm, A.; Freund, S. M.; Komander, D., Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nature structural & molecular biology* **2010,** *17* (8), 939-47.
31. Cook, W. J.; Jeffrey, L. C.; Carson, M.; Chen, Z.; Pickart, C. M., Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *The Journal of biological chemistry* **1992,** *267* (23), 16467-71.
32. Hospenthal, M. K.; Freund, S. M.; Komander, D., Assembly, analysis and architecture of atypical ubiquitin chains. *Nature structural & molecular biology* **2013,** *20* (5), 555-65.
33. Kristariyanto, Y. A.; Abdul Rehman, S. A.; Campbell, D. G.; Morrice, N. A.; Johnson, C.; Toth, R.; Kulathu, Y., K29-selective ubiquitin binding domain reveals structural basis of specificity and heterotypic nature of k29 polyubiquitin. *Molecular cell* **2015,** *58* (1), 83-94.
34. Kristariyanto, Y. A.; Choi, S. Y.; Rehman, S. A.; Ritorto, M. S.; Campbell, D. G.; Morrice, N. A.; Toth, R.; Kulathu, Y., Assembly and structure of Lys33-linked polyubiquitin reveals distinct conformations. *Biochem J* **2015,** *467* (2), 345-52.
35. 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.; Dixit, V. M., K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. *Molecular cell* **2010,** *39* (3), 477-84.
36. Michel, M. A.; Elliott, P. R.; Swatek, K. N.; Simicek, M.; Pruneda, J. N.; Wagstaff, J. L.; Freund, S. M.; Komander, D., Assembly and specific recognition of k29- and k33-linked polyubiquitin. *Molecular cell* **2015,** *58* (1), 95-109.
37. Komander, D.; Reyes-Turcu, F.; Licchesi, J. D.; Odenwaelder, P.; Wilkinson, K. D.; Barford, D., Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO reports* **2009,** *10* (5), 466-73.

38. Kanarek, N.; Ben-Neriah, Y., Regulation of NF-kappaB by ubiquitination and degradation of the IkappaBs. *Immunol Rev* **2012,** *246* (1), 77-94.
39. Lawrence, T., The nuclear factor NF-kappaB pathway in inflammation. *Cold Spring Harb Perspect Biol* **2009,** *1* (6), a001651.
40. Kirisako, T.; Kamei, K.; Murata, S.; Kato, M.; Fukumoto, H.; Kanie, M.; Sano, S.; Tokunaga, F.; Tanaka, K.; Iwai, K., A ubiquitin ligase complex assembles linear polyubiquitin chains. *The EMBO journal* **2006,** *25* (20), 4877-87.
41. Gerlach, B.; Cordier, S. M.; Schmukle, A. C.; Emmerich, C. H.; Rieser, E.; Haas, T. L.; Webb, A. I.; Rickard, J. A.; Anderton, H.; Wong, W. W.; Nachbur, U.; Gangoda, L.; Warnken, U.; Purcell, A. W.; Silke, J.; Walczak, H., Linear ubiquitination prevents inflammation and regulates immune signalling. *Nature* **2011,** *471* (7340), 591-6.
42. Ikeda, F.; Deribe, Y. L.; Skanland, S. S.; Stieglitz, B.; Grabbe, C.; Franz-Wachtel, M.; van Wijk, S. J.; Goswami, P.; Nagy, V.; Terzic, J.; Tokunaga, F.; Androulidaki, A.; Nakagawa, T.; Pasparakis, M.; Iwai, K.; Sundberg, J. P.; Schaefer, L.; Rittinger, K.; Macek, B.; Dikic, I., SHARPIN forms a linear ubiquitin ligase complex regulating NF-kappaB activity and apoptosis. *Nature* **2011,** *471* (7340), 637-41.
43. Rahighi, S.; Ikeda, F.; Kawasaki, M.; Akutsu, M.; Suzuki, N.; Kato, R.; Kensche, T.; Uejima, T.; Bloor, S.; Komander, D.; Randow, F.; Wakatsuki, S.; Dikic, I., Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. *Cell* **2009,** *136* (6), 1098-109.
44. Jin, L.; Williamson, A.; Banerjee, S.; Philipp, I.; Rape, M., Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell* **2008,** *133* (4), 653-65.
45. Meyer, H. J.; Rape, M., Enhanced protein degradation by branched ubiquitin chains. *Cell* **2014,** *157* (4), 910-21.
46. Thrower, J. S.; Hoffman, L.; Rechsteiner, M.; Pickart, C. M., Recognition of the polyubiquitin proteolytic signal. *The EMBO journal* **2000,** *19* (1), 94-102.
47. Kirkpatrick, D. S.; Hathaway, N. A.; Hanna, J.; Elsasser, S.; Rush, J.; Finley, D.; King, R. W.; Gygi, S. P., Quantitative analysis of in vitro ubiquitinated cyclin B1 reveals complex chain topology. *Nat Cell Biol* **2006,** *8* (7), 700-10.
48. Flick, K.; Ouni, I.; Wohlschlegel, J. A.; Capati, C.; McDonald, W. H.; Yates, J. R.; Kaiser, P., Proteolysis-independent regulation of the transcription factor Met4 by a single Lys 48-linked ubiquitin chain. *Nat Cell Biol* **2004,** *6* (7), 634-41.
49. Flick, K.; Raasi, S.; Zhang, H.; Yen, J. L.; Kaiser, P., A ubiquitin-interacting motif protects polyubiquitinated Met4 from degradation by the 26S proteasome. *Nat Cell Biol* **2006,** *8* (5), 509-15.
50. Kanayama, A.; Seth, R. B.; Sun, L.; Ea, C. K.; Hong, M.; Shaito, A.; Chiu, Y. H.; Deng, L.; Chen, Z. J., TAB2 and TAB3 activate the NF-kappaB pathway through binding to polyubiquitin chains. *Molecular cell* **2004,** *15* (4), 535-48.
51. Sato, Y.; Yoshikawa, A.; Yamashita, M.; Yamagata, A.; Fukai, S., Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by NZF domains of TAB2 and TAB3. *The EMBO journal* **2009,** *28* (24), 3903-9.
52. Emmerich, C. H.; Ordureau, A.; Strickson, S.; Arthur, J. S.; Pedrioli, P. G.; Komander, D.; Cohen, P., Activation of the canonical IKK complex by K63/M1-linked hybrid ubiquitin chains. *Proc Natl Acad Sci U S A* **2013,** *110* (38), 15247-52.
53. Chen, Z. J.; Sun, L. J., Nonproteolytic functions of ubiquitin in cell signaling. *Molecular cell* **2009,** *33* (3), 275-86.
54. Bennett, E. J.; Harper, J. W., DNA damage: ubiquitin marks the spot. *Nature structural & molecular biology* **2008,** *15* (1), 20-2.
55. Morris, J. R.; Solomon, E., BRCA1 : BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. *Hum Mol Genet* **2004,** *13* (8), 807-17.

56. Nishikawa, H.; Ooka, S.; Sato, K.; Arima, K.; Okamoto, J.; Klevit, R. E.; Fukuda, M.; Ohta, T., Mass spectrometric and mutational analyses reveal Lys-6-linked polyubiquitin chains catalyzed by BRCA1-BARD1 ubiquitin ligase. *The Journal of biological chemistry* **2004,** *279* (6), 3916-24.
57. Ordureau, A.; Sarraf, S. A.; Duda, D. M.; Heo, J. M.; Jedrychowski, M. P.; Sviderskiy, V. O.; Olszewski, J. L.; Koerber, J. T.; Xie, T.; Beausoleil, S. A.; Wells, J. A.; Gygi, S. P.; Schulman, B. A.; Harper, J. W., Quantitative proteomics reveal a feedforward mechanism for mitochondrial PARKIN translocation and ubiquitin chain synthesis. *Molecular cell* **2014,** *56* (3), 360-75.
58. Gatti, M.; Pinato, S.; Maiolica, A.; Rocchio, F.; Prato, M. G.; Aebersold, R.; Penengo, L., RNF168 promotes noncanonical K27 ubiquitination to signal DNA damage. *Cell reports* **2015,** *10* (2), 226-38.
59. Fei, C.; Li, Z.; Li, C.; Chen, Y.; Chen, Z.; He, X.; Mao, L.; Wang, X.; Zeng, R.; Li, L., Smurf1-mediated Lys29-linked nonproteolytic polyubiquitination of axin negatively regulates Wnt/beta-catenin signaling. *Mol Cell Biol* **2013,** *33* (20), 4095-105.
60. Huang, H.; Jeon, M. S.; Liao, L.; Yang, C.; Elly, C.; Yates, J. R., 3rd; Liu, Y. C., K33-linked polyubiquitination of T cell receptor-zeta regulates proteolysis-independent T cell signaling. *Immunity* **2010,** *33* (1), 60-70.
61. Al-Hakim, A. K.; Zagorska, A.; Chapman, L.; Deak, M.; Peggie, M.; Alessi, D. R., Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains. *Biochem J* **2008,** *411* (2), 249-60.
62. Yuan, W. C.; Lee, Y. R.; Lin, S. Y.; Chang, L. Y.; Tan, Y. P.; Hung, C. C.; Kuo, J. C.; Liu, C. H.; Lin, M. Y.; Xu, M.; Chen, Z. J.; Chen, R. H., K33-Linked Polyubiquitination of Coronin 7 by Cul3-KLHL20 Ubiquitin E3 Ligase Regulates Protein Trafficking. *Molecular cell* **2014,** *54* (4), 586-600.
63. Wenzel, D. M.; Stoll, K. E.; Klevit, R. E., E2s: structurally economical and functionally replete. *Biochem J* **2011,** *433* (1), 31-42.
64. Lee, I.; Schindelin, H., Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. *Cell* **2008,** *134* (2), 268-78.
65. Rape, M.; Reddy, S. K.; Kirschner, M. W., The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell* **2006,** *124* (1), 89-103.
66. Christensen, D. E.; Brzovic, P. S.; Klevit, R. E., E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages. *Nature structural & molecular biology* **2007,** *14* (10), 941-8.
67. Petroski, M. D.; Zhou, X.; Dong, G.; Daniel-Issakani, S.; Payan, D. G.; Huang, J., Substrate modification with lysine 63-linked ubiquitin chains through the UBC13-UEV1A ubiquitin-conjugating enzyme. *The Journal of biological chemistry* **2007,** *282* (41), 29936-45.
68. Verma, R.; Feldman, R. M.; Deshaies, R. J., SIC1 is ubiquitinated in vitro by a pathway that requires CDC4, CDC34, and cyclin/CDK activities. *Mol Biol Cell* **1997,** *8* (8), 1427-37.
69. Petroski, M. D.; Deshaies, R. J., Mechanism of lysine 48-linked ubiquitin-chain synthesis by the cullin-RING ubiquitin-ligase complex SCF-Cdc34. *Cell* **2005,** *123* (6), 1107-20.
70. Gazdoiu, S.; Yamoah, K.; Wu, K.; Pan, Z. Q., Human Cdc34 employs distinct sites to coordinate attachment of ubiquitin to a substrate and assembly of polyubiquitin chains. *Mol Cell Biol* **2007,** *27* (20), 7041-52.
71. Eddins, M. J.; Carlile, C. M.; Gomez, K. M.; Pickart, C. M.; Wolberger, C., Mms2-Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. *Nature structural & molecular biology* **2006,** *13* (10), 915-20.
72. Spratt, D. E.; Walden, H.; Shaw, G. S., RBR E3 ubiquitin ligases: new structures, new insights, new questions. *Biochem J* **2014,** *458* (3), 421-37.
73. Reyes-Turcu, F. E.; Ventii, K. H.; Wilkinson, K. D., Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annu. Rev. Biochem* **2009,** *78*, 363-97.

74. Komander, D.; Clague, M. J.; Urbe, S., Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol* **2009,** *10* (8), 550-63.
75. Mevissen, T. E. T.; Komander, D., Mechanisms of Deubiquitinase Specificity and Regulation. *Annual review of biochemistry* **2017,** *86*, 159-192.
76. Nijman, S. M.; Luna-Vargas, M. P.; Velds, A.; Brummelkamp, T. R.; Dirac, A. M.; Sixma, T. K.; Bernards, R., A genomic and functional inventory of deubiquitinating enzymes. *Cell* **2005,** *123* (5), 773-86.
77. Komander, D.; Clague, M. J.; Urbe, S., Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol* **2009,** *10* (8), 550-63.
78. 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., The differential modulation of USP activity by internal regulatory domains, interactors and eight ubiquitin chain types. *Chemistry & biology* **2011,** *18* (12), 1550-61.
79. Ritorto, M. S.; Ewan, R.; Perez-Oliva, A. B.; Knebel, A.; Buhrlage, S. J.; Wightman, M.; Kelly, S. M.; Wood, N. T.; Virdee, S.; Gray, N. S.; Morrice, N. A.; Alessi, D. R.; Trost, M., Screening of DUB activity and specificity by MALDI-TOF mass spectrometry. *Nature communications* **2014,** *5*, 4763.
80. Komander, D.; Lord, C. J.; Scheel, H.; Swift, S.; Hofmann, K.; Ashworth, A.; Barford, D., The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module. *Molecular cell* **2008,** *29* (4), 451-64.
81. Bingol, B.; Tea, J. S.; Phu, L.; Reichelt, M.; Bakalarski, C. E.; Song, Q.; Foreman, O.; Kirkpatrick, D. S.; Sheng, M., The mitochondrial deubiquitinase USP30 opposes parkin-mediated mitophagy. *Nature* **2014,** *510* (7505), 370-5.
82. Sun, S. C., CYLD: a tumor suppressor deubiquitinase regulating NF-kappaB activation and diverse biological processes. *Cell Death Differ* **2010,** *17* (1), 25-34.
83. Sato, Y.; Goto, E.; Shibata, Y.; Kubota, Y.; Yamagata, A.; Goto-Ito, S.; Kubota, K.; Inoue, J.; Takekawa, M.; Tokunaga, F.; Fukai, S., Structures of CYLD USP with Met1- or Lys63-linked diubiquitin reveal mechanisms for dual specificity. *Nature structural & molecular biology* **2015,** *22* (3), 222-9.
84. Mevissen, T. E.; Hospenthal, M. K.; Geurink, P. P.; Elliott, P. R.; Akutsu, M.; Arnaudo, N.; Ekkebus, R.; Kulathu, Y.; Wauer, T.; El Oualid, F.; Freund, S. M.; Ovaa, H.; Komander, D., OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell* **2013,** *154* (1), 169-84.
85. Keusekotten, K.; Elliott, P. R.; Glockner, L.; Fiil, B. K.; Damgaard, R. B.; Kulathu, Y.; Wauer, T.; Hospenthal, M. K.; Gyrd-Hansen, M.; Krappmann, D.; Hofmann, K.; Komander, D., OTULIN antagonizes LUBAC signaling by specifically hydrolyzing Met1-linked polyubiquitin. *Cell* **2013,** *153* (6), 1312-26.
86. Juang, Y. C.; Landry, M. C.; Sanches, M.; Vittal, V.; Leung, C. C.; Ceccarelli, D. F.; Mateo, A. R.; Pruneda, J. N.; Mao, D. Y.; Szilard, R. K.; Orlicky, S.; Munro, M.; Brzovic, P. S.; Klevit, R. E.; Sicheri, F.; Durocher, D., OTUB1 co-opts Lys48-linked ubiquitin recognition to suppress E2 enzyme function. *Molecular cell* **2012,** *45* (3), 384-97.
87. Mevissen, T. E. T.; Kulathu, Y.; Mulder, M. P. C.; Geurink, P. P.; Maslen, S. L.; Gersch, M.; Elliott, P. R.; Burke, J. E.; van Tol, B. D. M.; Akutsu, M.; Oualid, F. E.; Kawasaki, M.; Freund, S. M. V.; Ovaa, H.; Komander, D., Molecular basis of Lys11-polyubiquitin specificity in the deubiquitinase Cezanne. *Nature* **2016,** *538* (7625), 402-405.
88. Sokratous, K.; Hadjisavvas, A.; Diamandis, E. P.; Kyriacou, K., The role of ubiquitin-binding domains in human pathophysiology. *Crit Rev Clin Lab Sci* **2014,** *51* (5), 280-90.
89. Husnjak, K.; Dikic, I., Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu Rev Biochem* **2012,** *81*, 291-322.
90. Popovic, D.; Vucic, D.; Dikic, I., Ubiquitination in disease pathogenesis and treatment. *Nat Med* **2014,** *20* (11), 1242-53.

91. Atkin, G.; Paulson, H., Ubiquitin pathways in neurodegenerative disease. *Front Mol Neurosci* **2014,** *7*, 63.
92. Nath, S. R.; Lieberman, A. P., The Ubiquitination, Disaggregation and Proteasomal Degradation Machineries in Polyglutamine Disease. *Front Mol Neurosci* **2017,** *10*, 78.
93. Hofmann, R. M.; Pickart, C. M., In vitro assembly and recognition of Lys-63 polyubiquitin chains. *J Biol Chem* **2001,** *276* (30), 27936-43.
94. Piotrowski, J.; Beal, R.; Hoffman, L.; Wilkinson, K. D.; Cohen, R. E.; Pickart, C. M., Inhibition of the 26 S proteasome by polyubiquitin chains synthesized to have defined lengths. *J Biol Chem* **1997,** *272* (38), 23712-21.
95. Castaneda, C. A.; Kashyap, T. R.; Nakasone, M. A.; Krueger, S.; Fushman, D., Unique structural, dynamical, and functional properties of k11-linked polyubiquitin chains. *Structure* **2013,** *21* (7), 1168-81.
96. Kumar, K. S.; Spasser, L.; Erlich, L. A.; Bavikar, S. N.; Brik, A., Total chemical synthesis of di-ubiquitin chains. *Angew Chem Int Ed Engl* **2010,** *49* (48), 9126-31.
97. El Oualid, F.; Merkx, R.; Ekkebus, R.; Hameed, D. S.; Smit, J. J.; de Jong, A.; Hilkmann, H.; Sixma, T. K.; Ovaa, H., Chemical synthesis of ubiquitin, ubiquitin-based probes, and diubiquitin. *Angew Chem Int Ed Engl* **2010,** *49* (52), 10149-53.
98. Kumar, K. S.; Bavikar, S. N.; Spasser, L.; Moyal, T.; Ohayon, S.; Brik, A., Total chemical synthesis of a 304 amino acid K48-linked tetraubiquitin protein. *Angew Chem Int Ed Engl* **2011,** *50* (27), 6137-41.
99. Spasser, L.; Brik, A., Chemistry and biology of the ubiquitin signal. *Angew Chem Int Ed Engl* **2012,** *51* (28), 6840-62.
100. Geurink, P. P.; van Tol, B. D.; van Dalen, D.; Brundel, P. J.; Mevissen, T. E.; Pruneda, J. N.; Elliott, P. R.; van Tilburg, G. B.; Komander, D.; Ovaa, H., Development of Diubiquitin-Based FRET Probes To Quantify Ubiquitin Linkage Specificity of Deubiquitinating Enzymes. *ChemBioChem* **2016,** *17* (9), 816-20.
101. Haj-Yahya, N.; Haj-Yahya, M.; Castaneda, C. A.; Spasser, L.; Hemantha, H. P.; Jbara, M.; Penner, M.; Ciechanover, A.; Fushman, D.; Brik, A., Modifying the vicinity of the isopeptide bond to reveal differential behavior of ubiquitin chains with interacting proteins: organic chemistry applied to synthetic proteins. *Angew Chem Int Ed Engl* **2013,** *52* (42), 11149-53.
102. Virdee, S.; Ye, Y.; Nguyen, D. P.; Komander, D.; Chin, J. W., Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. *Nat Chem Biol* **2010,** *6* (10), 750-7.
103. Castaneda, C.; Liu, J.; Chaturvedi, A.; Nowicka, U.; Cropp, T. A.; Fushman, D., Nonenzymatic assembly of natural polyubiquitin chains of any linkage composition and isotopic labeling scheme. *J Am Chem Soc* **2011,** *133* (44), 17855-68.
104. Virdee, S.; Kapadnis, P. B.; Elliott, T.; Lang, K.; Madrzak, J.; Nguyen, D. P.; Riechmann, L.; Chin, J. W., Traceless and site-specific ubiquitination of recombinant proteins. *Journal of the American Chemical Society* **2011,** *133* (28), 10708-11.
105. Pickart, C. M.; Raasi, S., Controlled synthesis of polyubiquitin chains. *Methods Enzymol.* **2005,** *399*, 21-36.
106. Raasi, S.; Pickart, C. M., Ubiquitin chain synthesis. *Methods Mol Biol* **2005,** *301*, 47-55.
107. Varadan, R.; Walker, O.; Pickart, C.; Fushman, D., Structural properties of polyubiquitin chains in solution. *J Mol Biol* **2002,** *324* (4), 637-47.
108. Trang, V. H.; Valkevich, E. M.; Minami, S.; Chen, Y. C.; Ge, Y.; Strieter, E. R., Nonenzymatic polymerization of ubiquitin: single-step synthesis and isolation of discrete ubiquitin oligomers. *Angew Chem Int Ed Engl* **2012,** *51* (52), 13085-8.
109. Valkevich, E. M.; Guenette, R. G.; Sanchez, N. A.; Chen, Y. C.; Ge, Y.; Strieter, E. R., Forging isopeptide bonds using thiol-ene chemistry: site-specific coupling of ubiquitin molecules for studying the activity of isopeptidases. *J Am Chem Soc* **2012,** *134* (16), 6916-9.

110. Zhang, X.; Smits, A. H.; van Tilburg, G. B.; Jansen, P. W.; Makowski, M. M.; Ovaa, H.; Vermeulen, M., An Interaction Landscape of Ubiquitin Signaling. *Mol Cell* **2017,** *65* (5), 941-955 e8.
111. Flierman, D.; van der Heden van Noort, G. J.; Ekkebus, R.; Geurink, P. P.; Mevissen, T. E.; Hospenthal, M. K.; Komander, D.; Ovaa, H., Non-hydrolyzable Diubiquitin Probes Reveal Linkage-Specific Reactivity of Deubiquitylating Enzymes Mediated by S2 Pockets. *Cell chemical biology* **2016,** *23* (4), 472-82.
112. Hemelaar, J.; Galardy, P. J.; Borodovsky, A.; Kessler, B. M.; Ploegh, H. L.; Ovaa, H., Chemistry-based functional proteomics: mechanism-based activity-profiling tools for ubiquitin and ubiquitin-like specific proteases. *J Proteome Res* **2004,** *3* (2), 268-76.
113. Love, K. R.; Pandya, R. K.; Spooner, E.; Ploegh, H. L., Ubiquitin C-terminal electrophiles are activity-based probes for identification and mechanistic study of ubiquitin conjugating machinery. *ACS chemical biology* **2009,** *4* (4), 275-87.
114. Flavell, R. R.; Muir, T. W., Expressed protein ligation (EPL) in the study of signal transduction, ion conduction, and chromatin biology. *Acc Chem Res* **2009,** *42* (1), 107-16.
115. Borodovsky, A.; Kessler, B. M.; Casagrande, R.; Overkleeft, H. S.; Wilkinson, K. D.; Ploegh, H. L., A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. *The EMBO journal* **2001,** *20* (18), 5187-96.
116. Ovaa, H.; Kessler, B. M.; Rolen, U.; Galardy, P. J.; Ploegh, H. L.; Masucci, M. G., Activity-based ubiquitin-specific protease (USP) profiling of virus-infected and malignant human cells. *Proc Natl Acad Sci U S A* **2004,** *101* (8), 2253-8.
117. Mulder, M. P.; Witting, K.; Berlin, I.; Pruneda, J. N.; Wu, K. P.; Chang, J. G.; Merkx, R.; Bialas, J.; Groettrup, M.; Vertegaal, A. C.; Schulman, B. A.; Komander, D.; Neefjes, J.; El Oualid, F.; Ovaa, H., A cascading activity-based probe sequentially targets E1-E2-E3 ubiquitin enzymes. *Nature chemical biology* **2016,** *12* (7), 523-30.
118. Hewings, D. S.; Flygare, J. A.; Bogyo, M.; Wertz, I. E., Activity-based probes for the ubiquitin conjugation-deconjugation machinery: new chemistries, new tools, and new insights. *FEBS J* **2017,** *284* (10), 1555-1576.
119. Ekkebus, R.; Flierman, D.; Geurink, P. P.; Ovaa, H., Catching a DUB in the act: novel ubiquitin-based active site directed probes. *Curr Opin Chem Biol* **2014,** *23*, 63-70.
120. Haj-Yahya, N.; Hemantha, H. P.; Meledin, R.; Bondalapati, S.; Seenaiah, M.; Brik, A., Dehydroalanine-based diubiquitin activity probes. *Organic letters* **2014,** *16* (2), 540-3.
121. Weber, A.; Elliott, P. R.; Pinto-Fernandez, A.; Bonham, S.; Kessler, B. M.; Komander, D.; El Oualid, F.; Krappmann, D., A Linear Diubiquitin-Based Probe for Efficient and Selective Detection of the Deubiquitinating Enzyme OTULIN. *Cell Chem Biol* **2017,** *24* (10), 1299-1313 e7.
122. McGouran, J. F.; Gaertner, S. R.; Altun, M.; Kramer, H. B.; Kessler, B. M., Deubiquitinating enzyme specificity for ubiquitin chain topology profiled by di-ubiquitin activity probes. *Chemistry & biology* **2013,** *20* (12), 1447-55.
123. Paudel, P.; Zhang, Q.; Leung, C.; Greenberg, H. C.; Guo, Y.; Chern, Y. H.; Dong, A.; Li, Y.; Vedadi, M.; Zhuang, Z.; Tong, Y., Crystal structure and activity-based labeling reveal the mechanisms for linkage-specific substrate recognition by deubiquitinase USP9X. *Proc Natl Acad Sci U S A* **2019,** *116* (15), 7288-7297.
124. Li, G.; Liang, Q.; Gong, P.; Tencer, A. H.; Zhuang, Z., Activity-based diubiquitin probes for elucidating the linkage specificity of deubiquitinating enzymes. *Chemical communications (Cambridge, England)* **2014,** *50* (2), 216-8.
125. Nguyen, D. P.; Elliott, T.; Holt, M.; Muir, T. W.; Chin, J. W., Genetically encoded 1,2-aminothiols facilitate rapid and site-specific protein labeling via a bio-orthogonal cyanobenzothiazole condensation. *J Am Chem Soc* **2011,** *133* (30), 11418-21.

126. Lang, K.; Chin, J. W., Cellular incorporation of unnatural amino acids and bioorthogonal labeling of proteins. *Chem Rev* **2014,** *114* (9), 4764-806.
127. Liu, C. C.; Schultz, P. G., Adding new chemistries to the genetic code. *Annu Rev Biochem* **2010,** *79*, 413-44.
128. Wang, L.; Schultz, P. G., Expanding the genetic code. *Angewandte Chemie (International ed. in English)* **2004,** *44* (1), 34-66.
129. Nguyen, T. A.; Cigler, M.; Lang, K., Expanding the Genetic Code to Study Protein-Protein Interactions. *Angewandte Chemie (International ed. in English)* **2018,** *57* (44), 14350-14361.
130. Chen, H.; Venkat, S.; McGuire, P.; Gan, Q.; Fan, C., Recent Development of Genetic Code Expansion for Posttranslational Modification Studies. *Molecules (Basel, Switzerland)* **2018,** *23* (7).
131. Pham, N. D.; Parker, R. B.; Kohler, J. J., Photocrosslinking approaches to interactome mapping. *Current opinion in chemical biology* **2013,** *17* (1), 90-101.
132. Lee, K. J.; Kang, D.; Park, H. S., Site-Specific Labeling of Proteins Using Unnatural Amino Acids. *Molecules and cells* **2019,** *42* (5), 386-396.
133. Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G., A genetically encoded photocaged amino acid. *Journal of the American Chemical Society* **2004,** *126* (44), 14306-7.
134. Huguenin-Dezot, N.; Alonzo, D. A.; Heberlig, G. W.; Mahesh, M.; Nguyen, D. P.; Dornan, M. H.; Boddy, C. N.; Schmeing, T. M.; Chin, J. W., Trapping biosynthetic acyl-enzyme intermediates with encoded 2,3-diaminopropionic acid. *Nature* **2019,** *565* (7737), 112-117.
135. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G., Expanding the genetic code of Escherichia coli. *Science* **2001,** *292* (5516), 498-500.
136. Wang, L., Amersham Prize winner. Expanding the genetic code. *Science* **2003,** *302* (5645), 584-5.
137. Doctor, B. P.; Mudd, J. A., SPECIES SPECIFICITY OF AMINO ACID ACCEPTOR RIBONUCLEIC ACID AND AMINOACYL SOLUBLE RIBONUCLEIC ACID SYNTHETASES. *The Journal of biological chemistry* **1963,** *238*, 3677-81.
138. Wang, L., Engineering the Genetic Code in Cells and Animals: Biological Considerations and Impacts. *Acc Chem Res* **2017,** *50* (11), 2767-2775.
139. Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G., Designing logical codon reassignment - Expanding the chemistry in biology. *Chem Sci* **2015,** *6* (1), 50-69.
140. Hao, B.; Gong, W.; Ferguson, T. K.; James, C. M.; Krzycki, J. A.; Chan, M. K., A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science* **2002,** *296* (5572), 1462-6.
141. Srinivasan, G.; James, C. M.; Krzycki, J. A., Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **2002,** *296* (5572), 1459-62.
142. Kavran, J. M.; Gundllapalli, S.; O'Donoghue, P.; Englert, M.; Soll, D.; Steitz, T. A., Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proc Natl Acad Sci U S A* **2007,** *104* (27), 11268-73.
143. Nguyen, D. P.; Lusic, H.; Neumann, H.; Kapadnis, P. B.; Deiters, A.; Chin, J. W., Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA Synthetase/tRNA(CUA) pair and click chemistry. *Journal of the American Chemical Society* **2009,** *131* (25), 8720-1.
144. Takimoto, J. K.; Xiang, Z.; Kang, J. Y.; Wang, L., Esterification of an unnatural amino acid structurally deviating from canonical amino acids promotes its uptake and incorporation into proteins in mammalian cells. *ChemBioChem* **2010,** *11* (16), 2268-72.
145. Bi, X.; Pasunooti, K. K.; Lescar, J.; Liu, C. F., Thiazolidine-Masked alpha-Oxo Aldehyde Functionality for Peptide and Protein Modification. *Bioconjug Chem* **2017,** *28* (2), 325-329.
146. Bi, X.; Pasunooti, K. K.; Tareq, A. H.; Takyi-Williams, J.; Liu, C. F., Genetic incorporation of 1,2-aminothiol functionality for site-specific protein modification via thiazolidine formation. *Organic & biomolecular chemistry* **2016,** *14* (23), 5282-5.

147. Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S., Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem Biophys Res Commun* **2008,** *371* (4), 818-22.
148. Villain, M.; Vizzavona, J.; Rose, K., Covalent capture: a new tool for the purification of synthetic and recombinant polypeptides. *Chemistry & biology* **2001,** *8* (7), 673-9.
149. Batjargal, S.; Wang, Y. J.; Goldberg, J. M.; Wissner, R. F.; Petersson, E. J., Native chemical ligation of thioamide-containing peptides: development and application to the synthesis of labeled alpha-synuclein for misfolding studies. *Journal of the American Chemical Society* **2012,** *134* (22), 9172-82.
150. Lewis, Y. E.; Galesic, A.; Levine, P. M.; De Leon, C. A.; Lamiri, N.; Brennan, C. K.; Pratt, M. R., O-GlcNAcylation of alpha-Synuclein at Serine 87 Reduces Aggregation without Affecting Membrane Binding. *ACS chemical biology* **2017,** *12* (4), 1020-1027.
151. Wang, P.; Dong, S.; Shieh, J. H.; Peguero, E.; Hendrickson, R.; Moore, M. A. S.; Danishefsky, S. J., Erythropoietin derived by chemical synthesis. *Science* **2013,** *342* (6164), 1357-1360.
152. Siman, P.; Karthikeyan, S. V.; Nikolov, M.; Fischle, W.; Brik, A., Convergent chemical synthesis of histone H2B protein for the site-specific ubiquitination at Lys34. *Angewandte Chemie (International ed. in English)* **2013,** *52* (31), 8059-63.
153. Seenaiah, M.; Jbara, M.; Mali, S. M.; Brik, A., Convergent versus sequential protein synthesis: the case of ubiquitinated and glycosylated H2B. *Angewandte Chemie (International ed. in English)* **2015,** *54* (42), 12374-8.
154. Perler, F. B., InBase: the Intein Database. *Nucleic Acids Res* **2002,** *30* (1), 383-4.
155. Kane, P. M.; Yamashiro, C. T.; Wolczyk, D. F.; Neff, N.; Goebl, M.; Stevens, T. H., Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase. *Science* **1990,** *250* (4981), 651-7.
156. Vila-Perello, M.; Muir, T. W., Biological applications of protein splicing. *Cell* **2010,** *143* (2), 191-200.
157. Shah, N. H.; Muir, T. W., Inteins: Nature's Gift to Protein Chemists. *Chem Sci* **2014,** *5* (1), 446-461.
158. Southworth, M. W.; Amaya, K.; Evans, T. C.; Xu, M. Q.; Perler, F. B., Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein. *Biotechniques* **1999,** *27* (1), 110-4, 116, 118-20.
159. Vila-Perello, M.; Liu, Z.; Shah, N. H.; Willis, J. A.; Idoyaga, J.; Muir, T. W., Streamlined expressed protein ligation using split inteins. *Journal of the American Chemical Society* **2013,** *135* (1), 286-92.
160. Dawson, P. E.; Kent, S. B., Synthesis of native proteins by chemical ligation. *Annual review of biochemistry* **2000,** *69*, 923-60.
161. Conibear, A. C.; Watson, E. E.; Payne, R. J.; Becker, C. F. W., Native chemical ligation in protein synthesis and semi-synthesis. *Chem Soc Rev* **2018,** *47* (24), 9046-9068.
162. Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoet, M.; Monbaliu, J. M.; Melnyk, O., Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chemical reviews* **2019,** *119* (12), 7328-7443.
163. Jin, K.; Li, X., Advances in Native Chemical Ligation-Desulfurization: A Powerful Strategy for Peptide and Protein Synthesis. *Chemistry* **2018,** *24* (66), 17397-17404.
164. Wan, W.; Huang, Y.; Wang, Z.; Russell, W. K.; Pai, P. J.; Russell, D. H.; Liu, W. R., A facile system for genetic incorporation of two different noncanonical amino acids into one protein in Escherichia coli. *Angewandte Chemie (International ed. in English)* **2010,** *49* (18), 3211-4.
165. Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Davis, B. G., Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chem Sci* **2011,** *2* (9), 1617-1868.

166. Ryu, Y.; Schultz, P. G., Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat Methods* **2006,** *3* (4), 263-5.
167. Cravatt, B. F.; Wright, A. T.; Kozarich, J. W., Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annual review of biochemistry* **2008,** *77*, 383-414.
168. van Kasteren, S. I.; Florea, B. I.; Overkleeft, H. S., Activity-Based Protein Profiling: From Chemical Novelty to Biomedical Stalwart. *Methods Mol Biol* **2017,** *1491*, 1-8.
169. Liu, Y.; Patricelli, M. P.; Cravatt, B. F., Activity-based protein profiling: the serine hydrolases. *Proc Natl Acad Sci U S A* **1999,** *96* (26), 14694-9.
170. Kidd, D.; Liu, Y.; Cravatt, B. F., Profiling serine hydrolase activities in complex proteomes. *Biochemistry* **2001,** *40* (13), 4005-15.
171. Thornberry, N. A.; Peterson, E. P.; Zhao, J. J.; Howard, A. D.; Griffin, P. R.; Chapman, K. T., Inactivation of interleukin-1 beta converting enzyme by peptide (acyloxy)methyl ketones. *Biochemistry* **1994,** *33* (13), 3934-40.
172. Faleiro, L.; Kobayashi, R.; Fearnhead, H.; Lazebnik, Y., Multiple species of CPP32 and Mch2 are the major active caspases present in apoptotic cells. *The EMBO journal* **1997,** *16* (9), 2271-81.
173. Ranjitkar, P.; Perera, B. G.; Swaney, D. L.; Hari, S. B.; Larson, E. T.; Krishnamurty, R.; Merritt, E. A.; Villen, J.; Maly, D. J., Affinity-based probes based on type II kinase inhibitors. *Journal of the American Chemical Society* **2012,** *134* (46), 19017-25.
174. Chakrabarty, S.; Kahler, J. P.; van de Plassche, M. A. T.; Vanhoutte, R.; Verhelst, S. H. L., Recent Advances in Activity-Based Protein Profiling of Proteases. *Curr Top Microbiol Immunol* **2019,** *420*, 253-281.
175. Fu, J.; Wu, M.; Liu, X., Proteomic approaches beyond expression profiling and PTM analysis. *Anal Bioanal Chem* **2018,** *410* (17), 4051-4060.
176. Wang, S.; Tian, Y.; Wang, M.; Wang, M.; Sun, G. B.; Sun, X. B., Advanced Activity-Based Protein Profiling Application Strategies for Drug Development. *Front Pharmacol* **2018,** *9*, 353.
177. Sadler, N. C.; Wright, A. T., Activity-based protein profiling of microbes. *Current opinion in chemical biology* **2015,** *24*, 139-44.
178. Dadova, J.; Galan, S. R.; Davis, B. G., Synthesis of modified proteins via functionalization of dehydroalanine. *Current opinion in chemical biology* **2018,** *46*, 71-81.
179. Yang, B.; Wang, N.; Schnier, P. D.; Zheng, F.; Zhu, H.; Polizzi, N. F.; Ittuveetil, A.; Saikam, V.; DeGrado, W. F.; Wang, Q.; Wang, P. G.; Wang, L., Genetically Introducing Biochemically Reactive Amino Acids Dehydroalanine and Dehydrobutyrine in Proteins. *Journal of the American Chemical Society* **2019,** *141* (19), 7698-7703.
180. Meledin, R.; Mali, S. M.; Singh, S. K.; Brik, A., Protein ubiquitination via dehydroalanine: development and insights into the diastereoselective 1,4-addition step. *Organic & biomolecular chemistry* **2016,** *14* (21), 4817-23.
181. Xu, L.; Fan, J.; Wang, Y.; Zhang, Z.; Fu, Y.; Li, Y. M.; Shi, J., An activity-based probe developed by a sequential dehydroalanine formation strategy targets HECT E3 ubiquitin ligases. *Chemical communications (Cambridge, England)* **2019,** *55* (49), 7109-7112.
182. Rowan, F. C.; Richards, M.; Bibby, R. A.; Thompson, A.; Bayliss, R.; Blagg, J., Insights into Aurora-A kinase activation using unnatural amino acids incorporated by chemical modification. *ACS chemical biology* **2013,** *8* (10), 2184-91.
183. Chooi, K. P.; Galan, S. R.; Raj, R.; McCullagh, J.; Mohammed, S.; Jones, L. H.; Davis, B. G., Synthetic phosphorylation of p38alpha recapitulates protein kinase activity. *Journal of the American Chemical Society* **2014,** *136* (5), 1698-701.
184. Haldeman, M. T.; Xia, G.; Kasperek, E. M.; Pickart, C. M., Structure and function of ubiquitin conjugating enzyme E2-25K: the tail is a core-dependent activity element. *Biochemistry* **1997,** *36* (34), 10526-37.

185. Altun, M.; Walter, T. S.; Kramer, H. B.; Herr, P.; Iphofer, A.; Bostrom, J.; David, Y.; Komsany, A.; Ternette, N.; Navon, A.; Stuart, D. I.; Ren, J.; Kessler, B. M., The human otubain2-ubiquitin structure provides insights into the cleavage specificity of poly-ubiquitin-linkages. *PloS one* **2015,** *10* (1), e0115344.
186. Nanao, M. H.; Tcherniuk, S. O.; Chroboczek, J.; Dideberg, O.; Dessen, A.; Balakirev, M. Y., Crystal structure of human otubain 2. *EMBO reports* **2004,** *5* (8), 783-8.
187. Saldana, M.; VanderVorst, K.; Berg, A. L.; Lee, H.; Carraway, K. L., Otubain 1: a non-canonical deubiquitinase with an emerging role in cancer. *Endocr Relat Cancer* **2019,** *26* (1), R1-r14.
188. Zhang, Z.; Du, J.; Wang, S.; Shao, L.; Jin, K.; Li, F.; Wei, B.; Ding, W.; Fu, P.; van Dam, H.; Wang, A.; Jin, J.; Ding, C.; Yang, B.; Zheng, M.; Feng, X. H.; Guan, K. L.; Zhang, L., OTUB2 Promotes Cancer Metastasis via Hippo-Independent Activation of YAP and TAZ. *Molecular cell* **2019,** *73* (1), 7-21.e7.
189. Wilkinson, K. D.; Tashayev, V. L.; O'Connor, L. B.; Larsen, C. N.; Kasperek, E.; Pickart, C. M., Metabolism of the polyubiquitin degradation signal: structure, mechanism, and role of isopeptidase T. *Biochemistry* **1995,** *34* (44), 14535-46.
190. Reyes-Turcu, F. E.; Shanks, J. R.; Komander, D.; Wilkinson, K. D., Recognition of polyubiquitin isoforms by the multiple ubiquitin binding modules of isopeptidase T. *The Journal of biological chemistry* **2008,** *283* (28), 19581-92.
191. Bekes, M.; van der Heden van Noort, G. J.; Ekkebus, R.; Ovaa, H.; Huang, T. T.; Lima, C. D., Recognition of Lys48-Linked Di-ubiquitin and Deubiquitinating Activities of the SARS Coronavirus Papain-like Protease. *Molecular cell* **2016,** *62* (4), 572-85.
192. Lee, H.; Cao, S.; Hevener, K. E.; Truong, L.; Gatuz, J. L.; Patel, K.; Ghosh, A. K.; Johnson, M. E., Synergistic inhibitor binding to the papain-like protease of human SARS coronavirus: mechanistic and inhibitor design implications. *ChemMedChem* **2013,** *8* (8), 1361-72.
193. Liu, Y.; Soetandyo, N.; Lee, J. G.; Liu, L.; Xu, Y.; Clemons, W. M., Jr.; Ye, Y., USP13 antagonizes gp78 to maintain functionality of a chaperone in ER-associated degradation. *Elife* **2014,** *3*, e01369.
194. Russell, N. S.; Wilkinson, K. D., Deubiquitinating enzyme purification, assay inhibitors, and characterization. *Methods Mol Biol* **2005,** *301*, 207-19.
195. Jin, L.; Wang, W.; Fang, G., Targeting protein-protein interaction by small molecules. *Annu Rev Pharmacol Toxicol* **2014,** *54*, 435-56.
196. Tang, X.; Bruce, J. E., Chemical cross-linking for protein-protein interaction studies. *Methods Mol Biol* **2009,** *492*, 283-93.
197. Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R., Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Molecular & cellular proteomics : MCP* **2010,** *9* (8), 1634-49.
198. Cigler, M.; Muller, T. G.; Horn-Ghetko, D.; von Wrisberg, M. K.; Fottner, M.; Goody, R. S.; Itzen, A.; Muller, M. P.; Lang, K., Proximity-Triggered Covalent Stabilization of Low-Affinity Protein Complexes In Vitro and In Vivo. *Angewandte Chemie (International ed. in English)* **2017,** *56* (49), 15737-15741.
199. Williamson, A.; Wickliffe, K. E.; Mellone, B. G.; Song, L.; Karpen, G. H.; Rape, M., Identification of a physiological E2 module for the human anaphase-promoting complex. *Proc Natl Acad Sci U S A* **2009,** *106* (43), 18213-8.
200. Wu, T.; Merbl, Y.; Huo, Y.; Gallop, J. L.; Tzur, A.; Kirschner, M. W., UBE2S drives elongation of K11-linked ubiquitin chains by the anaphase-promoting complex. *Proc Natl Acad Sci U S A* **2010,** *107* (4), 1355-60.
201. Song, L.; Rape, M., Regulated degradation of spindle assembly factors by the anaphase-promoting complex. *Molecular cell* **2010,** *38* (3), 369-82.

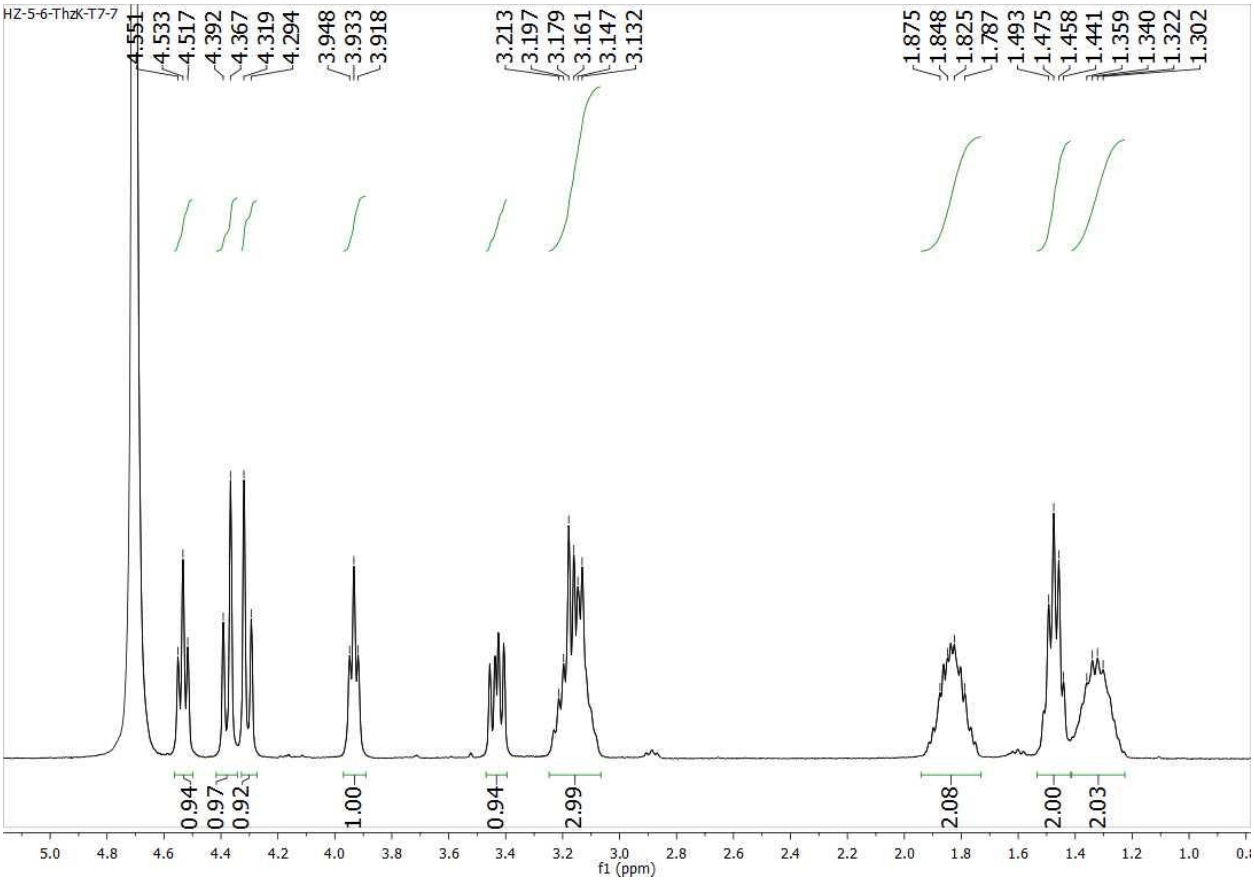
202. Wickliffe, K. E.; Lorenz, S.; Wemmer, D. E.; Kuriyan, J.; Rape, M., The mechanism of linkage-specific ubiquitin chain elongation by a single-subunit E2. *Cell* **2011,** *144* (5), 769-81.
203. Lorenz, S.; Bhattacharyya, M.; Feiler, C.; Rape, M.; Kuriyan, J., Crystal Structure of a Ube2S-Ubiquitin Conjugate. *PloS one* **2016,** *11* (2), e0147550.
204. Liess, A. K. L.; Kucerova, A.; Schweimer, K.; Yu, L.; Roumeliotis, T. I.; Diebold, M.; Dybkov, O.; Sotriffer, C.; Urlaub, H.; Choudhary, J. S.; Mansfeld, J.; Lorenz, S., Autoinhibition Mechanism of the Ubiquitin-Conjugating Enzyme UBE2S by Autoubiquitination. *Structure (London, England : 1993)* **2019,** *27* (8), 1195-1210.e7.
205. Kadokura, H.; Katzen, F.; Beckwith, J., Protein disulfide bond formation in prokaryotes. *Annual review of biochemistry* **2003,** *72*, 111-35.
206. Bulleid, N. J., Disulfide bond formation in the mammalian endoplasmic reticulum. *Cold Spring Harb Perspect Biol* **2012,** *4* (11).
207. Ellman, G. L., Tissue sulfhydryl groups. *Arch Biochem Biophys* **1959,** *82* (1), 70-7.
208. Owen, J. B.; Butterfield, D. A., Measurement of oxidized/reduced glutathione ratio. *Methods Mol Biol* **2010,** *648*, 269-77.
209. Riener, C. K.; Kada, G.; Gruber, H. J., Quick measurement of protein sulfhydryls with Ellman's reagent and with 4,4'-dithiodipyridine. *Anal Bioanal Chem* **2002,** *373* (4-5), 266-76.
210. Elmore, Z. C.; Donaher, M.; Matson, B. C.; Murphy, H.; Westerbeck, J. W.; Kerscher, O., Sumo-dependent substrate targeting of the SUMO protease Ulp1. *BMC Biol* **2011,** *9*, 74.
211. Mossessova, E.; Lima, C. D., Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Molecular cell* **2000,** *5* (5), 865-76.
212. Huang, L.; Kinnucan, E.; Wang, G.; Beaudenon, S.; Howley, P. M.; Huibregtse, J. M.; Pavletich, N. P., Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade. *Science* **1999,** *286* (5443), 1321-6.
213. Verdecia, M. A.; Joazeiro, C. A.; Wells, N. J.; Ferrer, J. L.; Bowman, M. E.; Hunter, T.; Noel, J. P., Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase. *Molecular cell* **2003,** *11* (1), 249-59.
214. Kamadurai, H. B.; Souphron, J.; Scott, D. C.; Duda, D. M.; Miller, D. J.; Stringer, D.; Piper, R. C.; Schulman, B. A., Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex. *Molecular cell* **2009,** *36* (6), 1095-102.
215. Wang, M.; Pickart, C. M., Different HECT domain ubiquitin ligases employ distinct mechanisms of polyubiquitin chain synthesis. *The EMBO journal* **2005,** *24* (24), 4324-33.
216. Kim, H. C.; Huibregtse, J. M., Polyubiquitination by HECT E3s and the determinants of chain type specificity. *Mol Cell Biol* **2009,** *29* (12), 3307-18.
217. 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.; Schulman, B. A., Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. *Elife* **2013,** *2*, e00828.
218. Maspero, E.; Valentini, E.; Mari, S.; Cecatiello, V.; Soffientini, P.; Pasqualato, S.; Polo, S., Structure of a ubiquitin-loaded HECT ligase reveals the molecular basis for catalytic priming. *Nature structural & molecular biology* **2013,** *20* (6), 696-701.
219. Scheffner, M.; Kumar, S., Mammalian HECT ubiquitin-protein ligases: biological and pathophysiological aspects. *Biochim Biophys Acta* **2014,** *1843* (1), 61-74.
220. Li, L.; Martinez, S. S.; Hu, W.; Liu, Z.; Tjian, R., A specific E3 ligase/deubiquitinase pair modulates TBP protein levels during muscle differentiation. *Elife* **2015,** *4*, e08536.

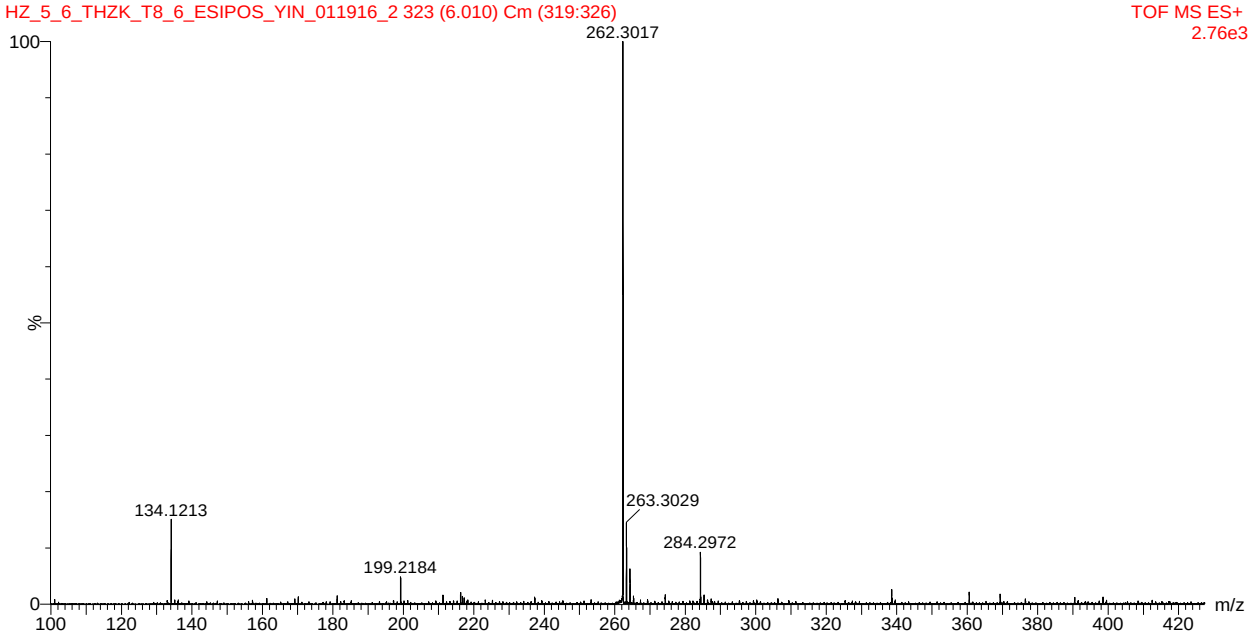
221. Chen, L. J.; Xu, W. M.; Yang, M.; Wang, K.; Chen, Y.; Huang, X. J.; Ma, Q. H., HUWE1 plays important role in mouse preimplantation embryo development and the dysregulation is associated with poor embryo development in humans. *Sci Rep* **2016,** *6*, 37928.
222. Myant, K. B.; Cammareri, P.; Hodder, M. C.; Wills, J.; Von Kriegsheim, A.; Gyorffy, B.; Rashid, M.; Polo, S.; Maspero, E.; Vaughan, L.; Gurung, B.; Barry, E.; Malliri, A.; Camargo, F.; Adams, D. J.; Iavarone, A.; Lasorella, A.; Sansom, O. J., HUWE1 is a critical colonic tumour suppressor gene that prevents MYC signalling, DNA damage accumulation and tumour initiation. *EMBO Mol Med* **2017,** *9* (2), 181-197.
223. Inoue, S.; Hao, Z.; Elia, A. J.; Cescon, D.; Zhou, L.; Silvester, J.; Snow, B.; Harris, I. S.; Sasaki, M.; Li, W. Y.; Itsumi, M.; Yamamoto, K.; Ueda, T.; Dominguez-Brauer, C.; Gorrini, C.; Chio, II; Haight, J.; You-Ten, A.; McCracken, S.; Wakeham, A.; Ghazarian, D.; Penn, L. J.; Melino, G.; Mak, T. W., Mule/Huwe1/Arf-BP1 suppresses Ras-driven tumorigenesis by preventing c-Myc/Miz1-mediated down-regulation of p21 and p15. *Genes Dev* **2013,** *27* (10), 1101-14.
224. Chen, D.; Kon, N.; Li, M.; Zhang, W.; Qin, J.; Gu, W., ARF-BP1/Mule is a critical mediator of the ARF tumor suppressor. *Cell* **2005,** *121* (7), 1071-83.
225. Jackl, M.; Stollmaier, C.; Strohaker, T.; Hyz, K.; Maspero, E.; Polo, S.; Wiesner, S., beta-Sheet Augmentation Is a Conserved Mechanism of Priming HECT E3 Ligases for Ubiquitin Ligation. *J Mol Biol* **2018,** *430* (18 Pt B), 3218-3233.
226. Ohtake, F.; Saeki, Y.; Ishido, S.; Kanno, J.; Tanaka, K., The K48-K63 Branched Ubiquitin Chain Regulates NF-kappaB Signaling. *Molecular cell* **2016,** *64* (2), 251-266.
227. Pandya, R. K.; Partridge, J. R.; Love, K. R.; Schwartz, T. U.; Ploegh, H. L., A structural element within the HUWE1 HECT domain modulates self-ubiquitination and substrate ubiquitination activities. *The Journal of biological chemistry* **2010,** *285* (8), 5664-73.
228. Sander, B.; Xu, W.; Eilers, M.; Popov, N.; Lorenz, S., A conformational switch regulates the ubiquitin ligase HUWE1. *Elife* **2017,** *6*.
229. Zou, X.; Levy-Cohen, G.; Blank, M., Molecular functions of NEDD4 E3 ubiquitin ligases in cancer. *Biochim Biophys Acta* **2015,** *1856* (1), 91-106.
230. Trotman, L. C.; Wang, X.; Alimonti, A.; Chen, Z.; Teruya-Feldstein, J.; Yang, H.; Pavletich, N. P.; Carver, B. S.; Cordon-Cardo, C.; Erdjument-Bromage, H.; Tempst, P.; Chi, S. G.; Kim, H. J.; Misteli, T.; Jiang, X.; Pandolfi, P. P., Ubiquitination regulates PTEN nuclear import and tumor suppression. *Cell* **2007,** *128* (1), 141-56.
231. Dou, H.; Buetow, L.; Sibbet, G. J.; Cameron, K.; Huang, D. T., BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nature structural & molecular biology* **2012,** *19* (9), 876-83.
232. Plechanovova, A.; Jaffray, E. G.; Tatham, M. H.; Naismith, J. H.; Hay, R. T., Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **2012,** *489* (7414), 115-20.
233. Brown, N. G.; Watson, E. R.; Weissmann, F.; Jarvis, M. A.; VanderLinden, R.; Grace, C. R. R.; Frye, J. J.; Qiao, R.; Dube, P.; Petzold, G.; Cho, S. E.; Alsharif, O.; Bao, J.; Davidson, I. F.; Zheng, J. J.; Nourse, A.; Kurinov, I.; Peters, J. M.; Stark, H.; Schulman, B. A., Mechanism of polyubiquitination by human anaphase-promoting complex: RING repurposing for ubiquitin chain assembly. *Molecular cell* **2014,** *56* (2), 246-260.
234. Brown, N. G.; VanderLinden, R.; Watson, E. R.; Qiao, R.; Grace, C. R.; Yamaguchi, M.; Weissmann, F.; Frye, J. J.; Dube, P.; Ei Cho, S.; Actis, M. L.; Rodrigues, P.; Fujii, N.; Peters, J. M.; Stark, H.; Schulman, B. A., RING E3 mechanism for ubiquitin ligation to a disordered substrate visualized for human anaphase-promoting complex. *Proc Natl Acad Sci U S A* **2015,** *112* (17), 5272-9.
235. Brown, N. G.; VanderLinden, R.; Watson, E. R.; Weissmann, F.; Ordureau, A.; Wu, K. P.; Zhang, W.; Yu, S.; Mercredi, P. Y.; Harrison, J. S.; Davidson, I. F.; Qiao, R.; Lu, Y.; Dube, P.; Brunner, M. R.; Grace, C. R. R.; Miller, D. J.; Haselbach, D.; Jarvis, M. A.; Yamaguchi, M.; Yanishevski, D.;

Petzold, G.; Sidhu, S. S.; Kuhlman, B.; Kirschner, M. W.; Harper, J. W.; Peters, J. M.; Stark, H.; Schulman, B. A., Dual RING E3 Architectures Regulate Multiubiquitination and Ubiquitin Chain Elongation by APC/C. *Cell* **2016,** *165* (6), 1440-1453.
236. Petroski, M. D.; Deshaies, R. J., Function and regulation of cullin-RING ubiquitin ligases. *Nat Rev Mol Cell Biol* **2005,** *6* (1), 9-20.
237. Lydeard, J. R.; Schulman, B. A.; Harper, J. W., Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO reports* **2013,** *14* (12), 1050-61.
238. Soucy, T. A.; Smith, P. G.; Milhollen, M. A.; Berger, A. J.; Gavin, J. M.; Adhikari, S.; Brownell, J. E.; Burke, K. E.; Cardin, D. P.; Critchley, S.; Cullis, C. A.; Doucette, A.; Garnsey, J. J.; Gaulin, J. L.; Gershman, R. E.; Lublinsky, A. R.; McDonald, A.; Mizutani, H.; Narayanan, U.; Olhava, E. J.; Peluso, S.; Rezaei, M.; Sintchak, M. D.; Talreja, T.; Thomas, M. P.; Traore, T.; Vyskocil, S.; Weatherhead, G. S.; Yu, J.; Zhang, J.; Dick, L. R.; Claiborne, C. F.; Rolfe, M.; Bolen, J. B.; Langston, S. P., An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **2009,** *458* (7239), 732-6.
239. Williams, K. M.; Qie, S.; Atkison, J. H.; Salazar-Arango, S.; Alan Diehl, J.; Olsen, S. K., Structural insights into E1 recognition and the ubiquitin-conjugating activity of the E2 enzyme Cdc34. *Nature communications* **2019,** *10* (1), 3296.
240. Goldenberg, S. J.; Cascio, T. C.; Shumway, S. D.; Garbutt, K. C.; Liu, J.; Xiong, Y.; Zheng, N., Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. *Cell* **2004,** *119* (4), 517-28.
241. Chen, H.; Shen, Y.; Tang, X.; Yu, L.; Wang, J.; Guo, L.; Zhang, Y.; Zhang, H.; Feng, S.; Strickland, E.; Zheng, N.; Deng, X. W., Arabidopsis CULLIN4 Forms an E3 Ubiquitin Ligase with RBX1 and the CDD Complex in Mediating Light Control of Development. *Plant Cell* **2006,** *18* (8), 1991-2004.
242. Li, T.; Chen, X.; Garbutt, K. C.; Zhou, P.; Zheng, N., Structure of DDB1 in complex with a paramyxovirus V protein: viral hijack of a propeller cluster in ubiquitin ligase. *Cell* **2006,** *124* (1), 105-17.
243. Angers, S.; Li, T.; Yi, X.; MacCoss, M. J.; Moon, R. T.; Zheng, N., Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* **2006,** *443* (7111), 590-3.
244. Li, T.; Robert, E. I.; van Breugel, P. C.; Strubin, M.; Zheng, N., A promiscuous alpha-helical motif anchors viral hijackers and substrate receptors to the CUL4-DDB1 ubiquitin ligase machinery. *Nature structural & molecular biology* **2010,** *17* (1), 105-11.
245. Olsen, S. K.; Lima, C. D., Structure of a ubiquitin E1-E2 complex: insights to E1-E2 thioester transfer. *Molecular cell* **2013,** *49* (5), 884-96.

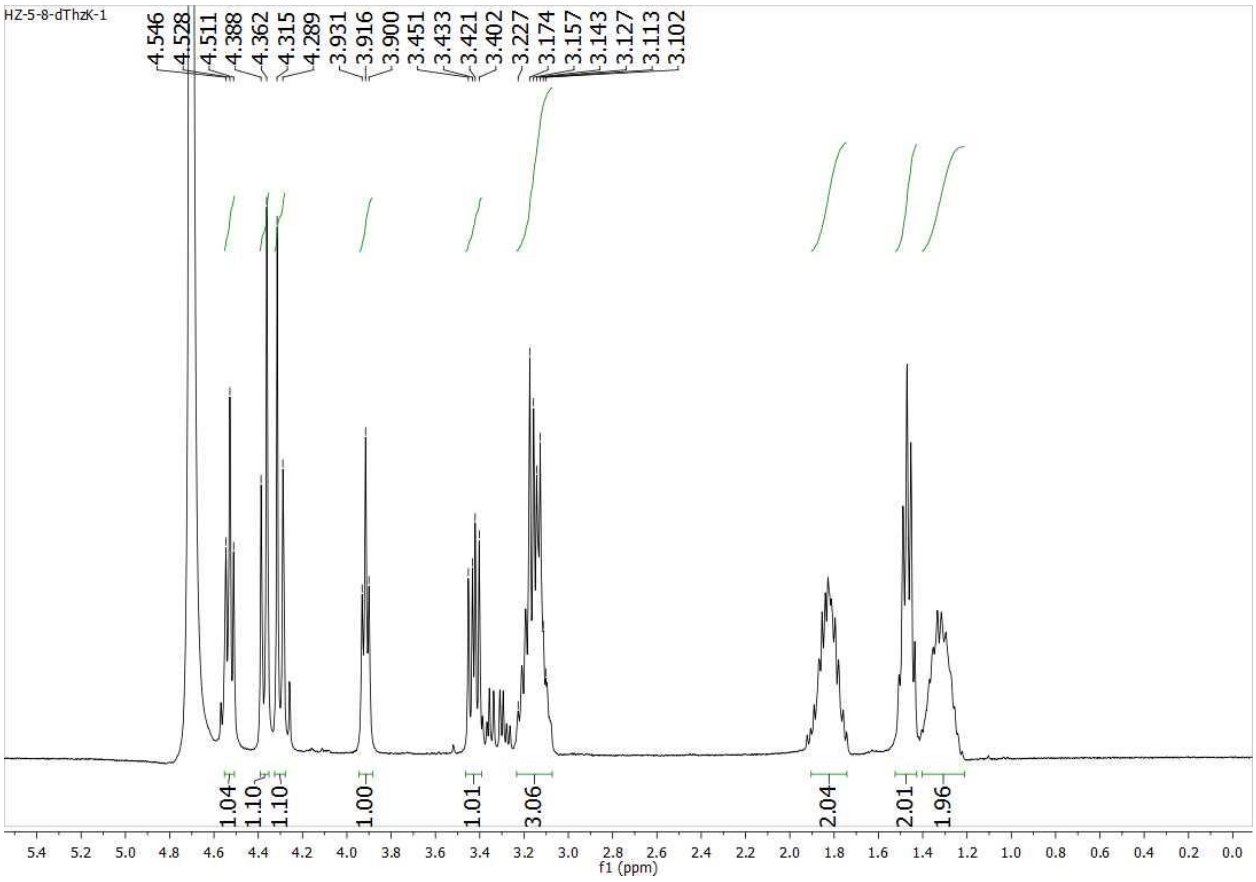
APPENDICES

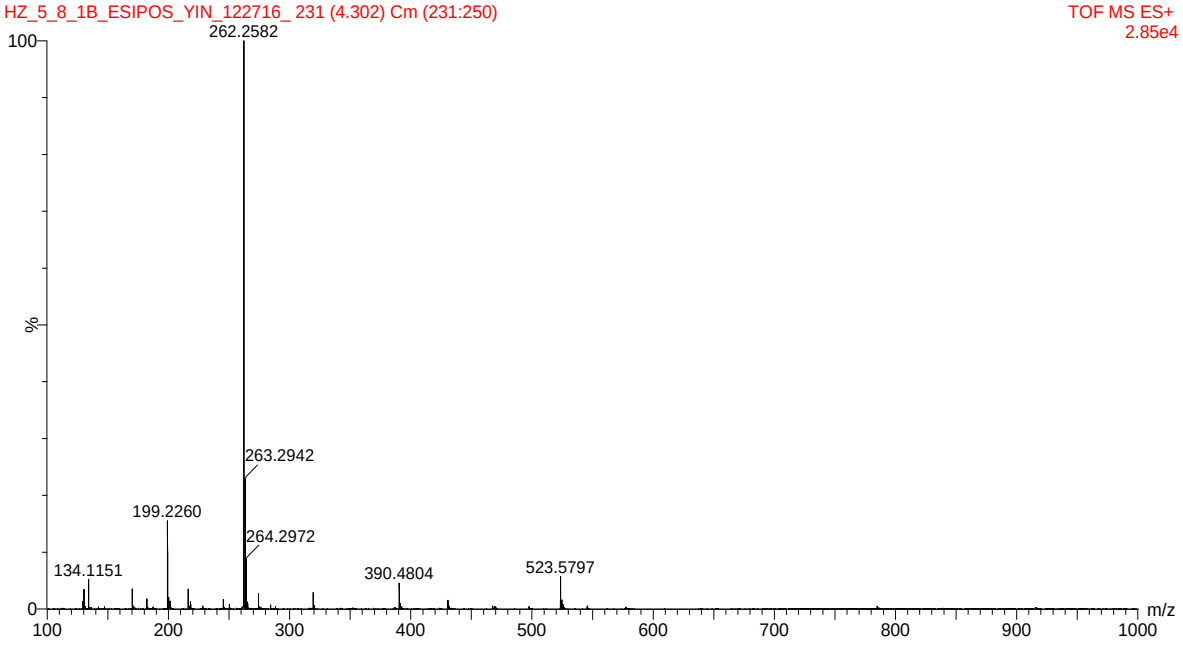
Appendix A NMR and ESI of unnatural amino acids.



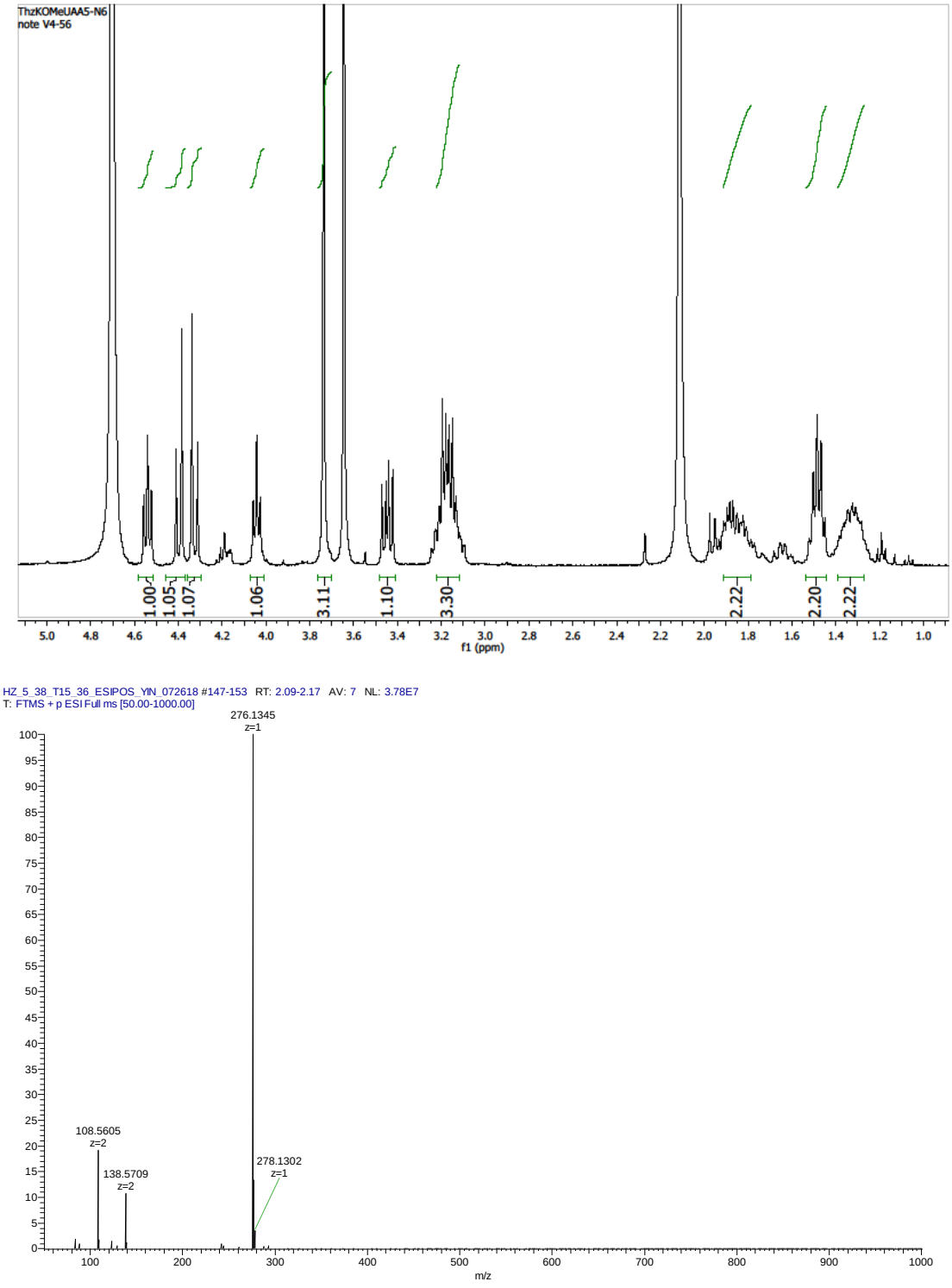


Figures S1. ^{1}H-NMR and MS(ESI+) of L-ThzK-OH (**7a**).

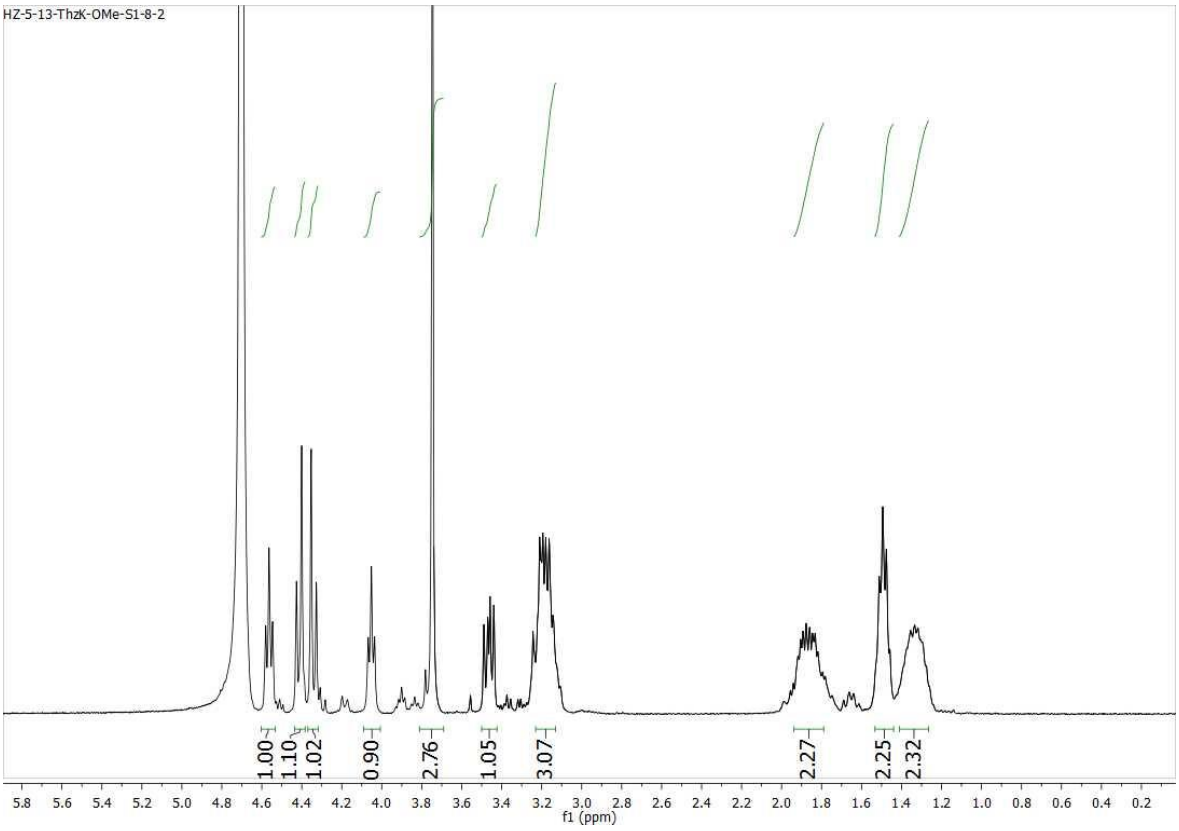


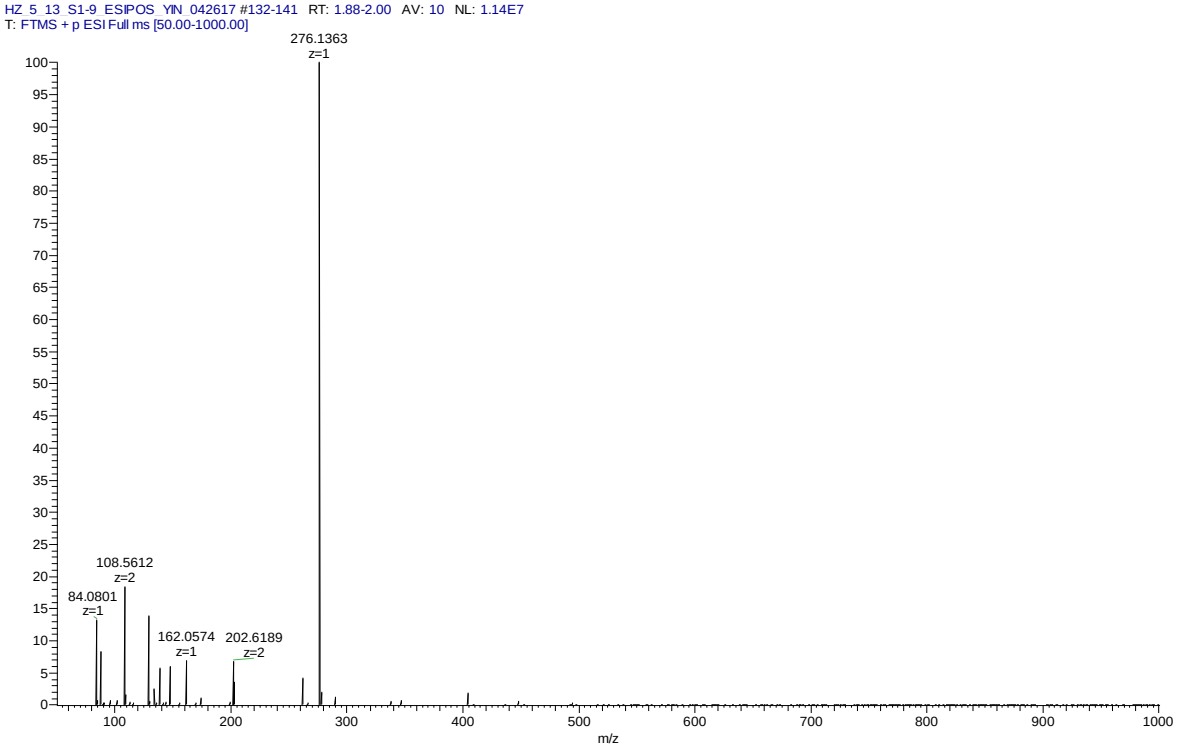


Figures S2. ^{1}H-NMR and MS(ESI+) of D-ThzK-OH (**8a**).



Figures S3. ^{1}H-NMR and MS(ESI+) of L-ThzK-OMe (**7b**).





Figures S4. ^{1}H-NMR and MS(ESI+) of D-ThzK-OMe (**8b**).

Appendix B NMR and ESI for α, α′′-Di-bromo-adipyl(bis)amide (DBAA)