ABSTRACT

Title of Dissertation: SITE-SPECIFIC INVESTIGATIONS OF UBIQUITIN ACTIVATION, CARBAMYLATION, AND INTERACTIONS WITH UBIQUITIN-BINDING DOMAINS

Westley Pawloski
Doctor of Philosophy, 2023

Dissertation directed by: Professor David Fushman
Department of Chemistry and Biochemistry

The post-translational modification of proteins with ubiquitin (Ub) can induce a multitude of cellular signaling processes. Ubiquitination involves the attachment of the C-terminus of Ub to lysines on the substrate protein through an isopeptide linkage. This process is facilitated by a multitude of enzymes which work in concert to write and erase these linkages. The power of Ub signaling is that Ub itself can be modified by additional Ub units to generate polyubiquitin chains through any of the seven lysines or N-terminal amine, and each of these attachment points produces polyubiquitin (polyUb) chains with unique orientations of the internal Ub. This allows for K48 polyUb chains to mark a substrate for proteasomal degradation or K63 polyUb chains to trigger DNA repair and maintenance processes.

The Ub signaling system is an amalgamation of post-translational modifications, enzymatic activity, and carefully curated protein-protein binding interactions for this small 76 amino acid protein. My work presented in this disseration involves harnessing the power of nuclear magnetic resonance (NMR) experimentation to observe interactions of multiple components of the Ub system with site-specific resolution and selective kinetics.

To this end, I have implemented some standard and atypical NMR experiments to observe the potential for carbon dioxide carbamates to modulate the Ub signaling system. I have determined the kinetics of the enzymatic Ub-activating process, and this was extrapolated to understand how ubiquitiun-like proteins, which share a similar fold to Ub, are discriminated from erroneously taking the place of Ub. I have solved the solution structure of an unusual ubiquitin-like domain and explored how it interacts with Ub. Lastly, I will report on the implementation of an unnatural amino acid that is a photosensitive cross-linker and demonstrate that this technology can be used to detect novel ubiquitin-binding proteins.

SITE-SPECIFIC INVESTIGATIONS OF UBIQUITIN ACTIVATION, CARBAMYLATION, AND INTERACTIONS WITH UBIQUITIN-BINDING DOMAINS

by

Westley Pawloski

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:
Professor David Fushman, Chair
Professor John Orban
Assistant Professor Myles Poulin
Associate Professor Paul Paukstelis
Professor Jeffery Klauda, Dean's Representative

Acknowledgements

My cat and dog, Cat and Dog, for emotional support when they aren't sleeping all day.

David Poole, because you put me in yours, I guess I have to reciprocate. I agree, it's time to meet up again and catch up over a few beers in Germany now that we're going to be extraordinarily wealthy post-docs. Honestly though, it's always fun to think back to our nights at SF State doing nonsensical experiments on that poor 300 MHz magnet and drinking whiskey and D_2O. I don't know how divergent my life would be if you didn't start up a conversation about iron NMR all those years ago. I think I owe you a lot for sparking my interest in biochemistry. Hope to see you again soon.

My parents, for not getting too worried when I kept pushing reality further and further away by pursuing more schooling. No matter how hard things get I know I can count on both of you for love and support.

The University of Maryland Graduate School and the Department of Chemistry and Biochemistry, who supported me through fellowships and awards – including the Ann G. Wylie Dissertation Fellowship, G. Forrest Woods Graduate Fellowship, and multiple Summer research awards. My research was also supported by Dr. Fushman's NIH grant GM065334 and NSF grant MCB1818280.

Past and current members of the Fushman lab for not being too crazy. A special thanks to Apurva and Dulith for teaching me most of what I know about protein purification, and to Andrew for fielding numerous, and repetitious, questions about using the chromatography systems. Steven, thanks for being a cool dude through the

whole graduate school experience. It's been fun to grow into a proper scientist alongside you.

Alysia, for being a pretty good life partner. I guess I'm starting to get old because time is seemingly slipping away. I still vividly remember moving out here alone, and we weren't sure what was going to happen in our lives and if we would remain together. You made the bold decision to follow me across the country, away from all of our friends and family that provided your foundation, so that we could start a life together. Throughout the entirety of my time in the program you have been a steadfast pillar in my life, regardless of how many late nights I had to pull or when the stress of everything got to be too large for me to hide inside. Thank you for everything and I love you.

Dr. Fushman, thank you for your mentorship and for fostering my growth as a scientist. Most of all, I would like to thank you for your patience.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

19S RP	19S regulatory particle
APC/C	anaphase promoting complex C
Alloc	N-Allyloxycarbonyl
AMP	adenosine monophosphate
APP	Amyloid Precursor Protein-Binding Protein 1
ARIA	ambiguous restraints for iterative assignment
ATG	autophagy-related
ATP	adenosine triphosphate
AZF	*p*-azidophenylalanine
BBO	broad band observe
Boc	tert-butyloxycarbonyl
Bpa	*p*-benzoylphenylalanine
BRCA1	breast cancer gene 1
CCPNMR	collaborative computational project for NMR
CHES	N-cyclohexyl-2-aminoethanesulfonic acid
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid
CO_2	carbon dioxide
CoA	coenzyme A
CRL	cullin ring ubiquitin ligases
CSP	chemical shift perturbation
DBD	DNA binding domain
DDI1	DNA-damage inducible protein 1

DDI2	DNA-damage inducible protein 2
DPC	DNA-protein crosslinks
DSK2	dominant suppressor of kar1
DUB	deubiquitinase
E1	ubiquitin-like protein activating enzyme
E2	ubiquitin-like protein conjugating enzyme
E3	ubiquitin-like protein ligating enzyme
E4	ubqiuitin-like protein elongase
FCCH	first catalytic cysteine domain
GTP	guanosine triphosphate
GST	Glutathione-S-transferase
H2CN	$H\varepsilon 2C\varepsilon N_{\zeta}$
HADDOCK	high ambiguity driven protein-protein docking
$HCO3^-$	bicarbonate
HDD	helical domain of DDI1
hDDi1	human DNA-damage inducible protein 1
hDDi2	human DNA-damage inducible protein 2
hDDIs	both human DNA-damage inducible protein 1 and 2
HEK	human embryonic kidney
HSQC	heteronuclear single quantum coherence
hULD	human DDI1 ULD
ISG15	interferon-stimulated gene 15
K0	lysine-free Ub

K_D	dissociation constant
MjTyRS	*Methanococcus jannaschii* tyrosyl-tRNA synthetase
monoUb	monoubiquitination
MS^2	mass spectrometry based proteomics
MOM	mitochondrial outer membrane
NAE	NEDD8 activating enzyme
NaP	sodium phosphate
NaNCO	sodium cyanate
NEDD8	neuronal precursor cell expressed developmentally down-regulated protein 8
NFκB	nuclear factor kappa-light-chain-enhancer of activated B cells
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect sprectroscopy
NTA	N-terminal acetate
OTUB1	otubain-1
pUb	phosphorylated ubiquitin
PINK1	PTEN Induced Kinase 1
pK_a	acid dissociation constant
polyUb	polymeric ubiquitin.
PTM	post-translational modification
$tRNA^{Tyr}$	orthogonal suppressor tyrosyl tRNA
RAD23	radiation sensitive proteins-23

RAP80	BRCA1-A complex subunit RAP80
RPN	regulatory particle non-ATPase
RUB1	related to ubiquitin 1
RuBisCo	ribulose-1,5-bisphosphate carboxylase-oxygenase
RVP	retroviral-like protease domain
SAE1	SUMO activating enzyme 1
SAE2	SUMO activating enzyme 2
SCCH	second catalytic cysteine domain
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
smPTM	small molecule post-translational modification
SOFAST	band-Selective Optimized-Flip-Angle Short-Transient
STUbL	SUMO-targeted ubiquitin ligases
STUbP	SUMO-targeting Ub proteases
SUMO	small ubiquitin-like modifiers
TEO	triethyloxonium
TOCSY	totatal correlation spectroscopy
TUBES	tandem ubiquitin binding entities
UAE	ubiquitin activating enzyme
Ub	ubiquitin
Ub2	ubiquitin dimer
UBA	ubiquitin associated domain
UBA3	ubiquitin activating enzyme 3

UBAy	UBA domain of human ubiquilin-1 protein
UBC9	ubiquitin-conjugating enzyme 9
UBD	ubiquitin binding domain
UBE2C	ubiquitin conjugating enzyme E2 C
UBE2K	ubiquitin conjugating enzyme E2 K
UBE2S	ubiquitin conjugating enzyme E2 S
UBL	ubiquitin-like protein
UFD	ubiquitin fold domain
UIM	ubiquitin interacting motif
ULD	ubiquitin-like domain
UPS	ubiquitin proteasome system.
yDDI1	yeast DDI1
yULD	ULD domain of yeast DDI1

Chapter 1. Literature review

1.1 Ubiquitin is a proteinaceous post-translational modification

Ubiquitin (Ub) is a small 76 amino acid protein utilized extensively in a multitude of cellular signaling pathways.[1] The Ub signal is written by the enzyme-catalyzed attachment of the Ub C-terminus to lysines on the target proteins through an isopeptide linkage. Ub, or protein-attached Ub, can additionally be ubiquitinated into polymeric Ub (polyUb) chains through one of the seven lysines (K6, K11, K27, K29, K33, K48, K63) or N-terminal amine (M1). The ubiquitin signal can therefore be represented as monoubiquitination (monoUb) or multiple-monoUb at different lysines, linear polyUb conjugated through a singular repeating lysine, linear polyUb conjugated through multiple lysines, or branching polyUb with multiple lysines on a Ub forming a polymerization nexus.[2–5] This high degree of modularity in the construction of Ub chains underlines the diversity of Ub signaling and showcases that by using a conserved signaling protein and enzymatic machinery many different signals can be written while limiting the metabolic cost.

Ubiquitin signaling was initially described through the discovery that ubiquitination was responsible for directing substrate proteins to the proteasome for proteolysis.[6–14] Subsequent work identified linear K48-linked polyUb chains as the de facto signal for facilitating this process[15], and it was further elaborated that a polyUb chain length of four Ub is optimal for binding to the regulatory subunit of the proteasome.[16] Linear polyUb chains linked through K63 were soon after identified as

being integrated with processes involved with DNA repair and maintenance and protein trafficking, while not inducing proteasomal degradation of substrates.[17]

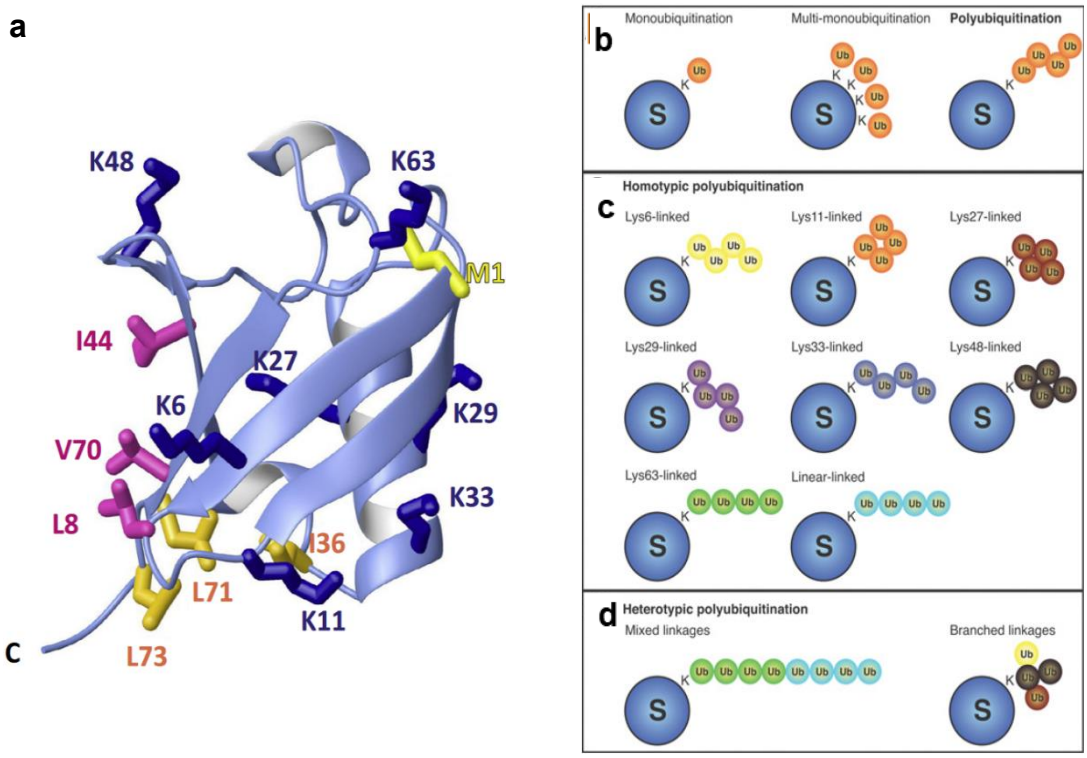


Figure 1. Structure of monoUb and orientation of each domain in polyUb chains. (**A**) The position of the N-terminal amine (yellow) is juxtaposed with the lysine sidechains (blue). The predominant hydrophobic binding patch is highlighted for I44, V70, and L8 in purple and an alternative patch consisting of L73, L71, and I36 in dark yellow. (**B – D**) The multiple forms of mono- or polyUb are represented in two dimensions. In **C** the experimental or predicted geometries are depicted for each amine based on the location in Ub.
(A) is adapted from: Alfano, C., Faggiano, S. & Pastore, A. The Ball and Chain of Polyubiquitin Structures. *Trends in Biochemical Sciences* **41**, 371–385 (2016).
(**B-D**) is adapated from: Komander, D. The emerging complexity of protein ubiquitination. *Biochemical Society Transactions* **37**, 937–953 (2009).

These divergent outcomes for the Ub-tagged proteins are attributed to K48- and K63-linked polyUb chains displaying unique recognition sites for their respective binding partners. Ub adopts a β-grasp fold consisting of a 5-stranded β-sheet interacting with a 3.5-turn α-helix which are connected through a 3_{10} helix.[18–20] The center of this beta-sheet primarily consists of cationic and uncharged polar residues but is interspersed with three hydrophobic residues which create the so-called hydrophobic

surface patch (L8, I44, V70) that is the primary interaction site for most Ub binding partners.[18,21–23] Each polyUb chain type, length, and the extent of branching places the Ub in different orientations relative to each other and the substrate protein. This allows for many different combinations of lysine linkages and chain lengths to generate unique binding surfaces. [24–30]

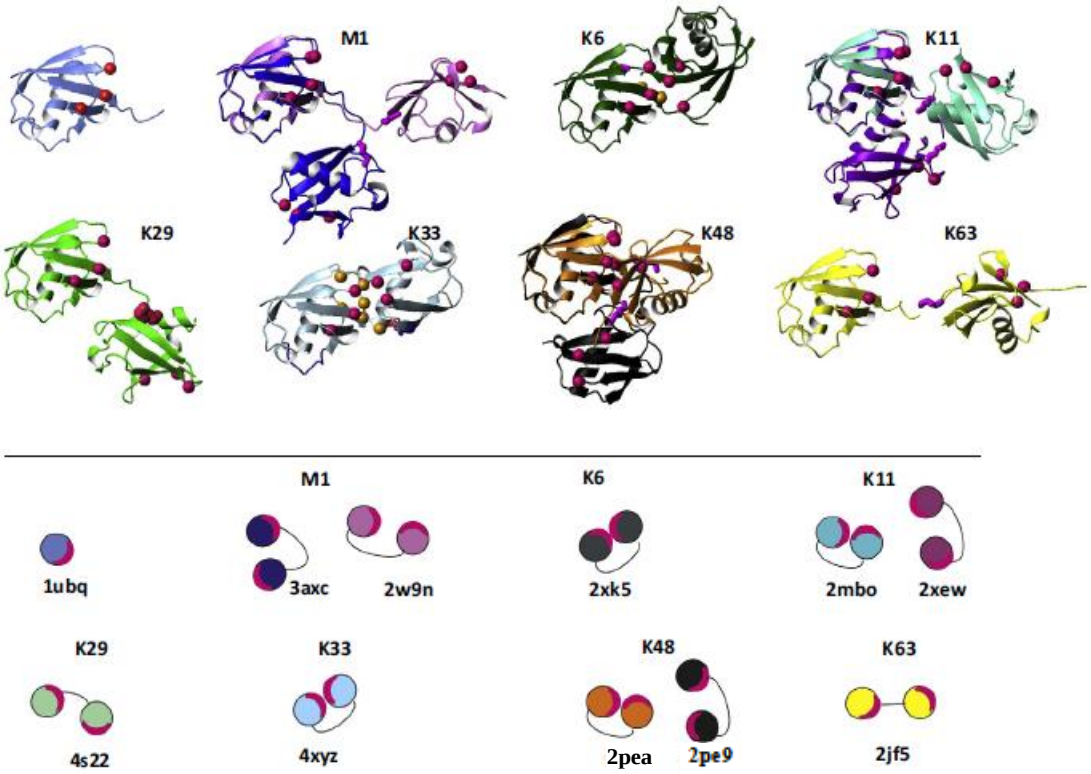


Figure 2. Ub dimers adopt unique geometries. The crystal structure of monoUb is shown adjacent to the crystal structures of Ub dimers. The lysine linkage is indicated for each structure and the hydrophobic patch residues are represented as purple spheres. In the bottom panel the polyUb type is indicated above the structure and the PDB accession codes are cited below. The hydrophobic patch of each Ub is highlighted in purple.
This figure is from: Alfano, C., Faggiano, S. & Pastore, A. The Ball and Chain of Polyubiquitin Structures. *Trends in Biochemical Sciences* **41**, 371–385 (2016)

The success of ubiquitin signaling requires that the desired Ub chains are built only when necessary and with a high fidelity to limit off-target signaling. Three classes of enzymes regulate all aspects of writing the ubiquitination signal with different levels

of specificity. Ubiquitin activating enzymes (E1) bind and charge the Ub C-terminus with a reactive thioester in an ATP-dependent reaction.[31–33] Ubiquitin conjugating enzymes (E2) extract the activated thioester[34–38] and with ubiquitin ligating enzymatic activity (E3) connect the activated Ub to the substrate protein or growing Ub chain.[39–43] E3 enzymes are the primary driver of substrate specificity and hundreds have been identified,[44,45] whereas only dozens of E2s are necessary owing to their ability to function with multiple E3s.[46–49] Ub activation, in contrast, is nearly universally achieved with a single E1 enzyme ubiquitin-activating enzyme 1 (UAE).[50,51] The Ub signal can be erased or pruned through the activity of deubiquitinases (DUBs) which cleave after the Ub C-terminus and regenerate free Ub. Some of these DUBs have general activity for any ligated Ub but others are only active for specific linkage types.[52–55]

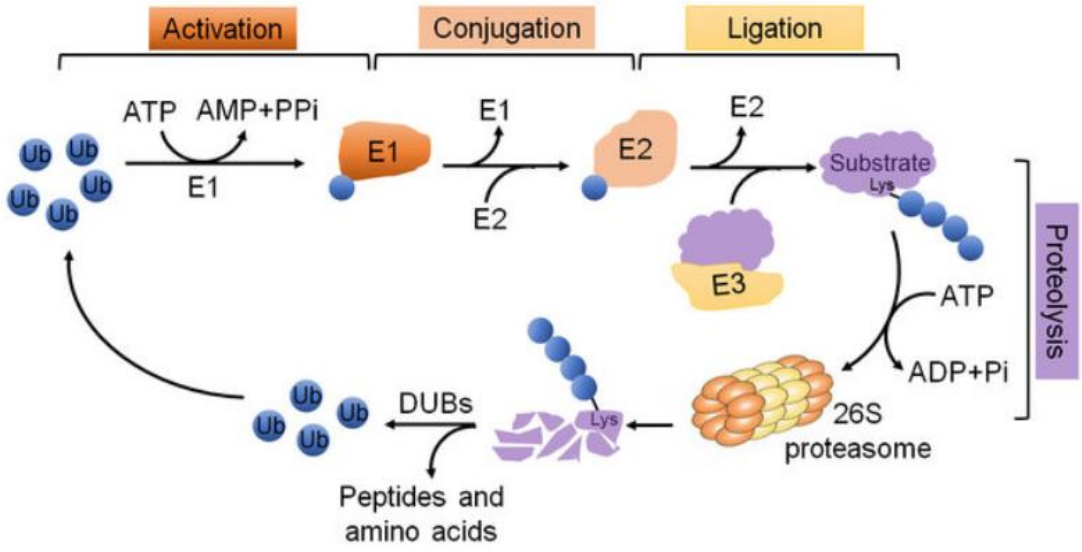


Figure 3. Enzymatic machinery of the ubiquitin proteasome system. After the activation of ubiquitin by an ATP-dependent process an enzymatic relay from E1 to E2 to E3 enzymes leads to ubiquitination of substrates. Deubiquitinates conversely remove conjugated ubiquitin.
This figure is from T. Zhou. The involvement of ubiquitin machinery in cell cycle regulation and cancer progression. Int. j. mol. Sci. (2021)

1.1.1 polyUb chains influence primary and secondary signaling pathways

Although the breadth of Ub signaling can be readily observed in aggregate, it can be difficult to nail down precisely how different polyUb linkage types or the amount and types of branching affect individual signaling pathways. The discovery that the Ub C-terminus can be cleaved by trypsin (Ub_{74} | GG) and leave behind a GG isopeptide linkage has allowed for a survey of ubiquitination substrates, the relative amounts of linkage types, and estimates of the amount of chain branching when studied by mass spectrometry based proteomics (MS^2).[56–69] In addition to Ub, Ub-chain, and GlyGly antibodies which can be used to enrich samples prior to or after digestion,[70–72] tandem ubiquitin binding entities (TUBES), consisting of multiple ubiquitin binding domains (UBDs), have been developed to enhance Ub chain enrichment and identify chain-specific substrates.[73–77] These technologies have allowed for scanning the Ub landscape in total, focusing on the ubiquitination of specific substrates, or analyzing how ubiquitination profiles change in response to different cellular conditions or environmental stressors.

MonoUb, or multiple monoUb on a single substrate, can regulate the function of substrate proteins by multiple ways. MonoUb can expand binding domains to include the Ub surface, block a binding interaction by steric occlusion, or block functionality by binding to a UBD on the substrate protein. MonoUb can therefore play a diverse role in modifying many proteins similar to small-molecule modifications like acetylation and phosphorylation. MonoUb is currently recognized to primarily influence protein trafficking by subcellular localization or endocytosis and modify protein-protein interactions leading to activation or inhibition of enzymatic activity.[78–

[80] Ubiquitination has been shown to disrupt the dynamics of substrate proteins near the sites of attachment leading to localized unfolding[81,82], and the conjugation of a single Ub to specific lysines is sufficient to induce proteasomal degradation of some substrate proteins.[83–86]

M1 linkages, or head-to-tail polyUb chains, are polyUb chains formed through the N-terminal amine. In contrast to polyUb chains conjugated through lysine sidechains, M1 linkages are generated by a single E3 enzyme which is highly specific for interfacing with the N-terminal amine. These chains adopt an "extended" conformation, similar to K63 polyUb,[87] where the individual Ub units do not form a contiguous binding interface. This shared morphology between M1 and K63 polyUb allows both chain types to have a complementary action in regulating the activation of NFκB, although by way of unique E3 and DUB enzymes.[88–93] M1 polyUb chains are commonly encountered in effectors of apoptotic pathways and seemingly act as stabilizing agents to prevent apoptosis signals.[94,95] K6 chains currently do not have a defined function and are poorly studied, but similarly to K63 they are found near DNA damage as a product of BRCA1 E3 activity. These chains are thought to either stabilize BRCA1 or bind to an unidentified component involved in DNA damage repair.[96,97] K6 polyUb chains are also evidenced to play a role in efficient mitophagy and mitochondrial homeostasis, and the improper regulation of mitophagy has been linked to the genesis of neurodegenerative diseases.[98–102]

K11 linkages are primarily linked to degradative functions in concert with K48 polyUb chains. K11 polyUb chains are produced by the E3 Anaphase Promoting Complex C (APC/C) to control cell cycle progression by causing the degradation of

cyclins, and thereby inhibiting cyclin dependent kinases and other proteins with functions relating to progression of the cell cycle. The APC/C can generate both K48 or K11 linkages through the two E2s UBE2C and UBE2S, but the latter is specific for generating K11 chains.[103–108] K11 and K48 polyUb chains can both bind to regulatory domains on the proteasome but with unique affinities, and are cleaved by individual DUBs at different rates. K11 chains are preferentially cleaved by the DUB Cezanne, which is similarly regulated by cell cycle progression, whereas K48 chains are resistant to this cleavage.[109,110] Because the E2s associated with the APC/C can build K11 and K48 polyUb linkages, branched K11/K48 chains can be produced on substrates. These mixed K11/K48 chains are considered to be a more efficient degradation signal than linear K11 or K48 polyUb due to having multiple binding interfaces, K11-like, K48-like, or mixed K11-K48, as well as being resistant to cleavage by linkage-specific DUBs.[111–115]

K27 is the only Ub lysine which is not solvent exposed and forms a unique polyUb structure with the isopeptide linkage buried in the hydrophobic core. This unique feature imparts K27 polyUb with a greater resistance to DUB activity.[116,117] K27 polyUb, similarly to K6 and K63 conjugates, is involved in the response to DNA damage and has been shown to recruit p53 binding protein 1.[118–120] K27 linkages are also connected to the innate immune system against microbial and viral infections including signal transduction and the intensity of the immune response.[121,122] K29 polyUb are similarly involved with many divergent pathways. Tangentially K29 linkages have been associated with Wnt/β-catenin signaling, mRNA stability, the inhibition of autophagy, and in the progression of neurodegenerative diseases such as

Huntington's, Alzheimer's, and Parkinson's Disease.[121,123,124] K33 polyUb have the least amount of experimental data and little information is available on interactions with the proteome. Currently K33 linkages have a sparse association with intracellular substrate trafficking.[125,126]

K48 is the only Ub lysine which is essential for yeast growth owing to the necessity of the Ub proteasome system (UPS) in cellular homeostasis.[127] Multiple Ub recognition sites in the 19S regulatory particle (19S RP) of the proteasome, located on the regulatory particle non-ATPase (RPN) subunits RPN1, RPN10, and RPN13, can all bind and orient monoUb or K48 polyUb chains so that proteasome-associated DUBs cleave the polyUb while the partially unfolded substrate is translocated into the proteolytic 20S core particle.[128–134] The extended conformation of K63 allows for two Ub units in K63 polyUb chains to bind the tandem ubiquitin interacting motif (tUIM) of RAP80, recruit a DNA repair complex including BRCA1, and localize this complex to the site of DNA damage. Additionally, E3s which are known to generate K63 polyUb chains have confirmed interactions with substrates involved with NFκB signaling and the endocytosis of membrane proteins.[66,135–141]

1.1.2 Ubiquitin-proteasome shuttle proteins

Ub shuttle proteins are generally defined as multiple domain proteins which contain a ubiquitin-like domain (ULD), a linker region, sometimes with defined functions, and one or more ubiquitin-associated domains (UBA).[142–145] These were initially proposed to function primarily as additional factors in promoting K48 polyUb degradation signals. By this mechanism the N-terminal ULD, which has a similar structure and binding interface to Ub, can associate with RPN1, RPN10, or RPN13

while the C-terminal UBA domain binds to Ub or polyUb chains with a preference for K48 linkages. The ULDs of proteasomal shuttles have unique binding affinities for the RPNs and this grants an additional mechanism for modifying Ub-dependent proteolysis. [143,146–149] The three primary Ub shuttles were initially identified in yeast as radiation sensitive proteins-23 (RAD23), dominant suppressor of kar1 (DSK2), and DNA-damage inducible protein 1 (DDI1).[131,150–154]

RAD23 contains an additional UBA domain which has a preference for binding K48 polyUb chains but additionally shows a weaker affinity for binding K63 polyUb chains.[135] This multi-functionality imparted by the two UBAs of RAD23 is consistent with its dual roles of localizing to DNA to facilitate DNA repair processes through the K63 polyUb UBA while assisting in the removal of extraneous proteins at the sites of

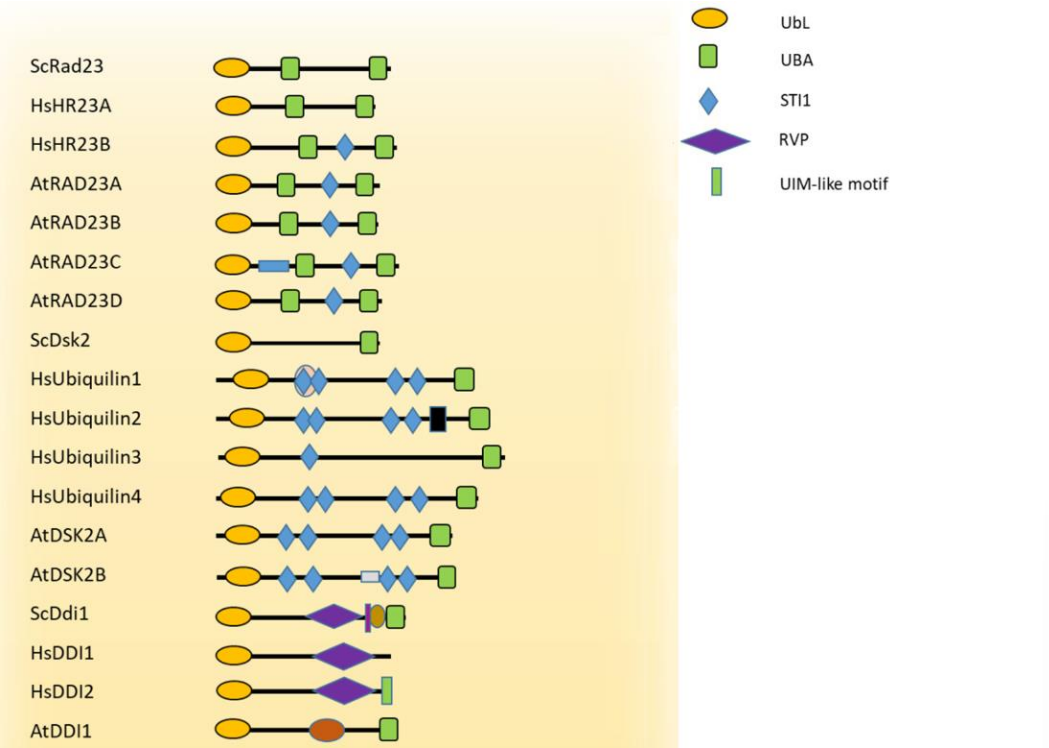


Figure 4. The domain types in canonical ubiquitin-proteasome shuttle proteins. The domain types relative to this dissertation are indicated in the top right.
This figure is adapted from: Zientara-Rytter, K. & Subramani, S. The Roles of Ubiquitin-Binding Protein Shuttles in the Degradative Fate of Ubiquitinated Proteins in the Ubiquitin-Proteasome System and Autophagy. *Cells* **8**, 40 (2019)

DNA repair through the K48 polyUb UBA.[155,156] The ULD of DSK2, in contrast, has a weaker affinity for the proteasome than RAD23[153] and the singular UBA preferentially binds to K48 polyUb.[157,158] Both RAD23 and DSK2, paradoxically, have been shown to block interactions between proteasomal DUBs and K48 polyUb chains, and thus act to prolong the life of K48 polyUb marked substrates.[146,159–161] Mammalian homologues of RAD23 (HR23A/B) and DSK2 (Ubiquilin1-4) have additional domains which resemble heat-shock chaperone domains (STI1) which create additional binding sites and may indicate a role in protein quality control outside of proteolysis. [162–167]

1.1.3 DDI1 is an enigmatic proteasomal shuttle

Distinct from the ULDs of RAD23 and DSK2, which share a ~30% sequence identity with Ub, the sequence of the ULD from DDI1 is completely different from Ub. Most prominently this manifests as the surface of the β-sheet, which contains the putative hydrophobic binding patch in Ub, displaying a majority of acidic rather than basic functional groups.[168] The DDI1 ULD retains the capacity to bind to receptors on the proteasome, albeit at unique locations, and the UBA domain retains the negatively charged binding surface similar to the other shuttle proteins.[169,170] In contrast to the other shuttle proteins which can adopt an autoinhibitory structure where the ULD and UBA domains can bind to each other, both intra- or intermolecularly, the negative binding surface of the DDI1 ULD prohibits similar interactions.[171] Surprisingly it was reported that the negatively charged binding surface of the DDI1 ULD has a physiologically relevant binding affinity for monoUb, and even binds stronger than the UBA. This observation put into question the ability for DDI1 to effectively bind to Ub and the proteasome simultaneously.[168,172]

DDI1 in humans is present as two homologues DDI1/2 and both have lost the N-terminal UBA and rely on other components for specificity of function. Both yeast and human DDI1 have a highly conserved retroviral-like protease domain (RVP) which is similar to HIV-1 protease and is a domain-swapped homodimer in the active form. In addition both DDI1/2 have a DNA binding domain (DBD) and a so-called helical domain of DDI1 (HDD) which are thought to be primarily for interfacing with non-Ub substrates.[172–174] The ability to dimerize in a head to tail configuration suggested that DDI1/2 could still function as a proteasomal shuttle by using both ULD domains of the dimer to bind Ub and the proteasome simultaneously.[168] This appears to not be the case for DDI2, as the ULD was shown to have lost its significantly negative surface charge and has physiologically irrelevant binding affinities for Ub or any Ub_2. In studies with the full length yeast DDI1 the ULD and HDD, but not the UBA, work in concert to enhance binding to polyUb chains, but it has yet to be confirmed if human DDI1 remains a Ub binding protein.[172,174]

The RVP domain of the HIV virus is crucial for infectivity, as it processes nascent viral polypeptides at multiple sites to generate all components of the mature virion. Following autoproteolysis from the transcribed polypeptide, the protease homodimerizes into its active form. In this configuration the two catalytic aspartates are positioned adjacently within the active site pocket, with one functioning as a general base while the other remains protonated due to the microenvironment.[175] Small molecule drugs have been developed to target this unique environment and thereby reduce the infectivity and severity of HIV-induced acquired immunodeficiency syndrome. Serendipitously, patients receiving treatment regimens containing these

small molecules exhibited resistance to infection by opportunistic pathogens, including parasitic worms (helminths) and protozoans. The only known RVP-like domain-containing protein in these organisms that could be affected by these drugs is DDI1.[176,177]

DDI1 is a multifunctional protein known to regulate cellular homeostasis through several signaling pathways. It has been proposed to act as a ULD/UBA proteasomal shuttle and enhance the capacity for protein turnover. DDI1 is also recognized to be responsible for the cleavage and activation of NRF1, a transcription factor. In its nascent form, NRF1 is rapidly ubiquitinated and degraded by the proteasome. However, when facilitated by DDI1, NRF1 evades degradation, translocates to the nucleus, and induces the expression of proteasomal components.[178,179] The induction of DDI1 by DNA damage suggests a primary role in DNA repair and maintenance, and DDI1 has been implicated in alleviating DNA-protein crosslinks (DPC). SPRTN, or WSS1 in yeast, is a DNA-activated metalloprotease known to specifically degrade DPCs and enables homologous recombination and nucleotide excision repair on the liberated DNA.[180] DDI1, SPRTN, and the proteasome are all negative genetic interactors for clearing hyperpolyubiquitinated substrates, indicating that DDI1 can stimulate protein degradation independent of the proteasome. In addition to these central roles in cellular homeostasis, DDI1 also regulates protein secretion by inhibiting the assembly of SNARE complexes and vesicle-membrane fusion, as well as cell cycle regulation through the degradation of HO endonuclease and F-box protein UFO1.[173,180]

The observed reduction in polyubiquitinated substrates and the canonical ULD/UBA domain orientation associated DDI1 with the proteasomal degradation of proteins. In yeast, the ULD domain can bind to RPN1 / RPN10, while the UBA domain binds to Ub and K48 polyUb chains.[181] However, the enigma of DDI1 is that either of these domains is absent in higher organisms[182], rendering these constructs incapable of acting as a canonical shuttle. The discovery that the ULD domain of yeast DDI1 (yDDI1) could bind to Ub led to a proposed mechanism in which dimerization through the RVP domain positions the ULDs, promiscuous for both Ub and proteasome receptors, on opposite sides of the complex.[168]

Human cells express two distinct DDI proteins, hDDI1 and hDDI2 (collectively referred to as hDDIs). hDDI1 lacks a C-terminal UBD whereas the C-terminus of hDDI2 adopts a helical structure resembling a UIM. Deletion of either hDDI results in the accumulation of high molecular weight polyUb chains, suggesting both hDDIs are complementary and share a conserved role in homeostasis with scDDI1. For example, the hyperpolyubiquitination phenotype observed with hDDI2 knockouts *in cellulo* was shown to be resolved upon the addition of *p. falciparum* DDI1.[183,184] Despite these observations, hDDI2 was found to bind to monoUb or Ub dimers (Ub2) weakly, and both the ULD and UIM exhibited minimal binding to Ub constructs *in vitro*. The ULD is specifically weakened for Ub binding, as the β-sheet surface has lost much of the negatively charged sidechains that facilitate Ub binding. [181]

Despite extensive efforts to identify proteolytic targets of DDI1, the RVP domain appeared inactive towards peptide libraries at physiological pH, cleaving only select peptides under acidic conditions. For both hDDIs the conserved HDD domain

was expected to impart substrate specificity while simultaneously activating the proteolytic domain.[184] Surprisingly, proteolysis by scDDI1 was found to be induced by binding to polyUb chains, requiring long polyUb chains of at least 8 units for efficient cleavage. Optimal binding necessitated both the HDD and ULD domains, with a HDD deletion decreasing the binding affinity by 10-fold and ULD deletion resulting in a 10^3-fold decrease in the binding affinity *in vitro*. Deletion of the UBA domain had no effect on binding. In contrast with DUBs, which cleave isopeptide linkages, scDDI1 cuts the substrate protein in the proximity of the isopeptide linkage with an unknown sequence specificity.[172,185] Therefore, aside from the proteasome, DDI1 is the only endopeptidase that acts on ubiquitinated proteins. Currently, there is no established sequence specificity for proteolysis, but the proteolytic domain of DDI1 is larger than that in HIV, potentially allowing for multiple sequence motifs to be accommodated.[186]

1.2 Ubiquitin-like signaling

Ubiquitin-like proteins (UBL) and domains (ULD) are broadly defined by having a similar structure to Ub. Many UBLs are utilized as post-translational modifications in the same manner as Ub (UBL signaling), with similar E1, E2, and E3 enzymatic machinery that generate the isopeptide linkage to target proteins. Notably, many of these E1 enzymes are heterodimers of the adenylation and thioester domains, rather than a single polypeptide like the Ub E1, but retain a conserved activation mechanism.[187–190]

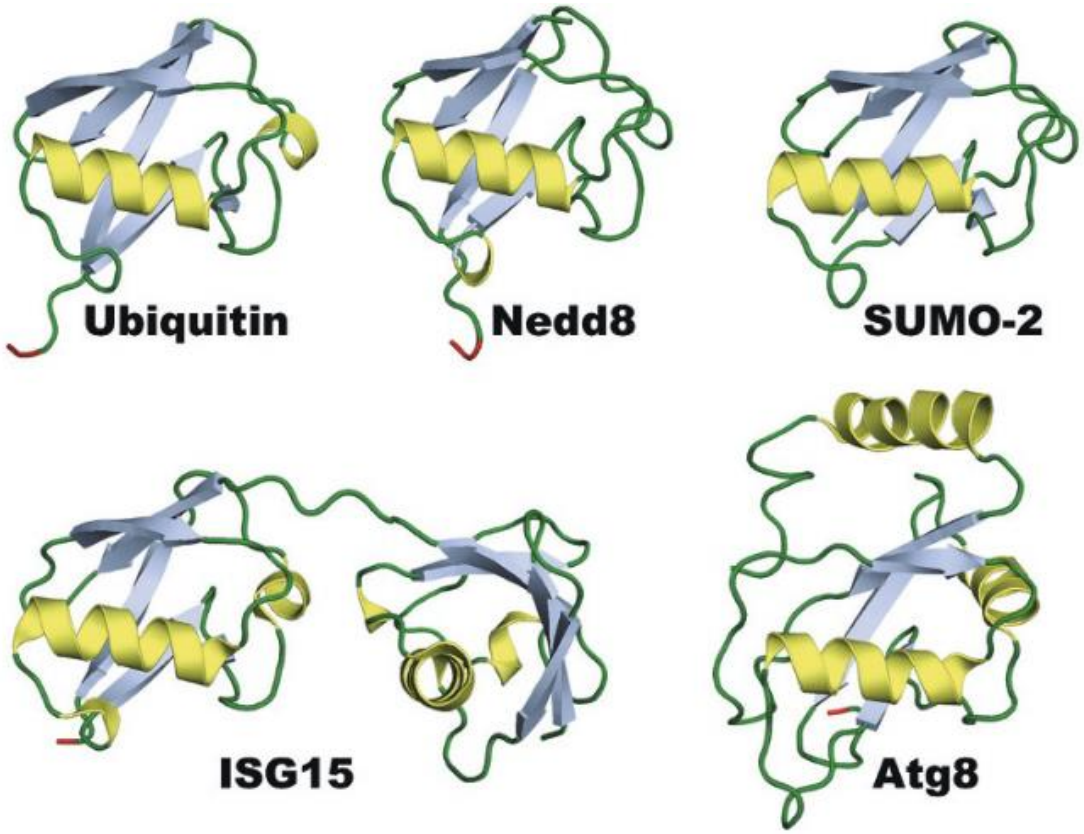


Figure 5. A side-by-side arrangement of UBL structures. The beta-grasp fold is conserved in Ub and UBLs. ISG15 contains two ubiquitin-like domains connected through a linker.
This figure is from: Ronau, J. A., Beckmann, J. F. & Hochstrasser, M. Substrate specificity of the ubiquitin and Ubl proteases. *Cell Res.* **26**, 441–456 (2016).

Although many UBL signaling pathways are much less studied than Ub, current evidence suggests that most lack the ability to form polymeric chains and therefore exist as monomeric or multi-monomeric post-translational modifications. Examples include members of the Autophagy-related (ATG) 8 family of proteins in vertebrates, encompassing approximately 7 UBLs, which have been found both conjugated to proteins and to the amine of phosphatidylethanolamine in autophagic membranes. These UBLs stimulate many processes involved with protein trafficking and selective autophagy including mitophagy and ribophagy.[191–194] Interferon-stimulated gene 15 (ISG15) contains two UBL domains each with ~30% sequence identity to Ub that are

fused through a short linker. The structure of ISG15 is similar to the head-to-tail M1-linked Ub dimer although ISG15 does not contain the Ub-terminus GG motif between the two UBL domains.[195–197] ISG15 functions principally to combat viral infections, and after its induction by type 1 interferons is found conjugated to a multitude of proteins in the host and to foreign proteins. Many of these targets have an unknown function or consequence, but some of the viral protein targets are known to inhibit viral replication when tagged with ISG15 whereas the modified host proteins are involved with membrane fusion and impact retroviral release processes.[198–203] ISG15 conjugates have also been observed on nascent polypeptide chains which disrupt proper folding pathways. These modifications appear to lack substrate or sequence specificity and inhibit viral replication.[204] The indiscriminate attachment of ISG15, after stimulation by interferons, may act as a panacea to combat all manner of infections.

1.2.1 SUMO signaling pathways similarly include polymerized UBL

The small ubiquitin-like modifiers (SUMOs) are UBLs which, in aggregate, are the most like Ub signaling in the breadth of substrates they modify and their polymerization into polySUMO chains. In line with other UBL signaling SUMO chains are constructed with an E1, E2, and E3 enzymatic pathway, however SUMO polymerization relies heavily on SUMO elongases (E4) which work via a different mechanism than canonical SUMO E3s.[205–209] Four SUMO paralogues are expressed in vertebrates denoted as SUMO1, SUMO2/3, which differ in sequence by only 4% and have the same functionality, and SUMO4 which has limited expression in basal conditions. SUMO2/3 have a sequence similarity to SUMO1 of ~45% and can share the same SUMO activating enzyme heterodimer (polymerization relies heavily on

SUMO elongases (E4) which work via a different mechanism than SAE1/SAE2) and the sole SUMO E2 enzyme Ubiquitin-conjugating enzyme 9 (Ubc9). The final substrate targeting by SUMO E3s have known preferences between conjugating SUMO1 or SUMO2/3, and SUMO deconjugating enzymes have similar biases for deconjugating either of the two species.[210–212]

Only SUMO2/3 is known to interact with E3s and E4s to produce polySUMO chains. As there are significantly less SUMO E3s and deconjugases than is currently known for Ub signaling pathways, current evidence suggests that synergistic interactions with other PTMs including ubiquitination, phosphorylation, and acetylation ultimately sculpt SUMO signaling.[213–215] In one of the few instances of directed UBL crosstalk, proteins known as SUMO-targeted ubiquitin ligases (STUbL) have SUMO-interacting motif domains, which bind to SUMO, as well as Ub E3 ligase activity targeting either SUMO or the SUMOylated protein. While some STUbLs are predicted to generate K63 polyUb chains, thus modifying the SUMO signal by presenting non-proteolytic polyUb chains, the primary effect of STUbLs appears to be generating K48 polyUb chains and marking the SUMOylated protein for proteasomal degradation.[216–221] SUMO-targeting Ub proteases (STUbP) work in contrast to cleave Ub conjugated to SUMO and create purely SUMO-tagged substrates.[215,222]

1.2.2 NEDD8, the cousin of Ub, can be processed by Ub conjugation machinery

Similarly to the relationship between SUMO1 and SUMO2/3, Ub shares an ~60% sequence identity and ~80% sequence similarity with the UBL related to ubiquitin 1 (RUB1) in *saccharomyces cerevisiae* and neuronal precursor cell expressed developmentally down-regulated protein 8 (NEDD8) in vertebrates. Although Ub is

highly conserved from yeast to human evolution, and only differs by three residues, the similarity of RUB1 to NEDD8 is about the same as either protein to Ub individually.

Regardless, due to this high sequence similarity the overall folds, lysine positions, and key hydrophobic patch recognition sequence are nearly identical between the three proteins.[223] One glaring example of nonuniformity is that Ub, RUB1, and NEDD8 all share the same ten C-terminal residues except for the residue at position 72. Both RUB1 and NEDD8 have small and uncharged sidechains at position 72, a threonine or alanine respectively, but Ub contains a cationic arginine at this residue. As this location is close to the C-terminus it would be expected to be involved in processes relating to the C-terminal activation by E1s and transfer of the activated UBL to E2 enzymes.[224]

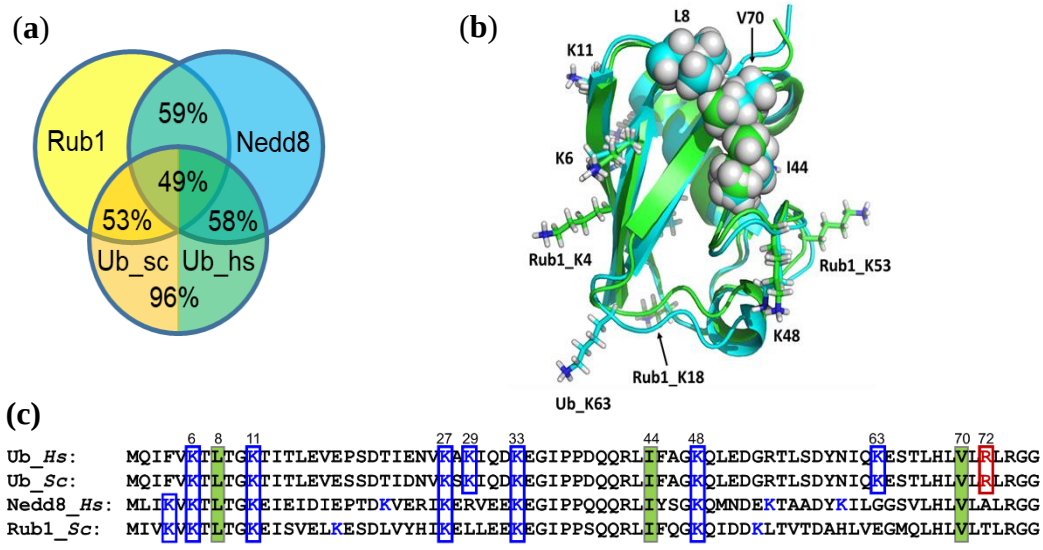


Figure 6. RUB1/NEDD8 shares a high sequence similarity to Ub. (a) The sequence identity between RUB1, NEDD8, and Ub. **(b)** Overlay of the structures of RUB1 and Ub highlighting the conserved hydrophobic patch and lysine residues. (**c**) Sequence alignment of the three proteins highlighting the conserved hydrophobic patch residues (L8, I44, V70) in green, conservation or variability of lysine residues in blue, and the critical Arg 72 residue in Ub which prevents binding to the NEDD8 E1.
This figure is from: Gurevich, S. Z. *et al.* Rub1/NEDD8, a ubiquitin-like modifier, is also a ubiquitin modifier. 2020.06.18.159145 Preprint at https://doi.org/10.1101/2020.06.18.159145 (2020).

Similarly to the SUMO pathway, NEDD8 has a single E1 (APPBP1 / UBA3), two known E2s, and ~10 E3s, but unlike SUMO signaling NEDD8 has limited conjugation targets. The primary role of RUB1 / NEDD8 (Simply referred to as NEDD8 for the remainder of this chapter unless specified) conjugation is to activate Cullins, a component of a specific class of Ub E3 ligases known as cullin ring ubiquitin ligases (CRL).[225–229] The neddylation of Cullins activates the CRL to ligate substrates with primarily K48 polyUb chains leading to proteasomal degradation, and it is estimated that degradation of upwards of 20% of the human proteome is regulated by CRLs. The activation of CRLs is known to depend only on mono-neddylation, but there is evidence for additional NEDD8 targets and polymeric NEDD8 chains to potentially contribute to other signaling pathways.[225,230–232]

Due to the high degree of similarity between NEDD8 and Ub, the UAE has the capability to bind, activate, and transfer NEDD8 to Ub E2s. Studies *in vitro* have demonstrated that this charged NEDD8 can be added to polyUb chains and generate chimeric Ub-NEDD8 polymers. The Ub E2s have been demonstrated to manipulate the NEDD8 C-terminal thioester, and conjugate it onto Ub lysines, but are unable to recognize NEDD8 as a substrate and cannot conjugate Ub or NEDD8 to NEDD8.[233,234] This phenomenon has led to theories that Ub E2s could interface with unknown E3s or substrates to generate atypical neddylation, either by novel signaling pathways or to modify canonical Ub / NEDD8 pathways.[224,235,236] The relative concentrations of NEDD8 and Ub fluctuate wildly during the cell cycle progression or in response to stressors, but non-ligated Ub and NEDD8 have been shown to be present in roughly the same concentrations in some stages of the cell cycle. In stress conditions where the

flux of proteins is accelerated, the intracellular pool of monomeric Ub is depleted as Ub is polymerized into K48 polyUb chains.[237,238] The K48 polyUb degradation signal has been shown to depend on a minimum number of four K48-linked Ub, and can be modestly enhanced with up to eight Ub,[16] but beyond this point the Ub are not utilized effectively. The now relatively large concentration of NEDD8 can be processed by the UAE, transferred to Ub E2s, and used to cap growing K48 chains to help to preserve the pool of monomeric Ub for further polymerization.[224,234]

If the opposite reaction were to occur, and Cullins could be activated by Ub infiltrating the NEDD8 signaling pathway, many scenarios exist for runaway ubiquitination due to spurious E3 ligase activity. It was recognized soon after the characterization of NEDD8 and the NEDD8 activating enzyme (NAE) that the activation mechanism was congruent with the Ub pathway.[239,240] Later structural studies confirmed the similarity of the adenylation sites, aside from the major addition of a positively charged arginine where the C-terminus of NEDD8 binds. This permutation appears specifically designed to exclude Ub, with a cationic arginine at a key position in the C-terminus, from fully docking into the activation site. This appears to be the primary discrimination against Ub from entering the NEDD8 pathway and prevents Ub from being transferred to NEDD8 E2 and E3 enzymes and ultimately activating CRLs.[241] Mutating this key residue to a hydrophobic leucine was sufficient to bypass this gating mechanism and allow for the NAE to activate Ub.[240]

1.3 Ubiquitin-like signaling is gated by E1 enzymes

The E1 of Ub, UAE, is a 118 kDa polypeptide which contains five domains: two adenylation domains, only one of which is functional, two halves of a catalytic cysteine domain (first (FCCH) and second (SCCH) catalytic cysteine domains, and a ubiquitin fold domain (UFD)).[242–244]

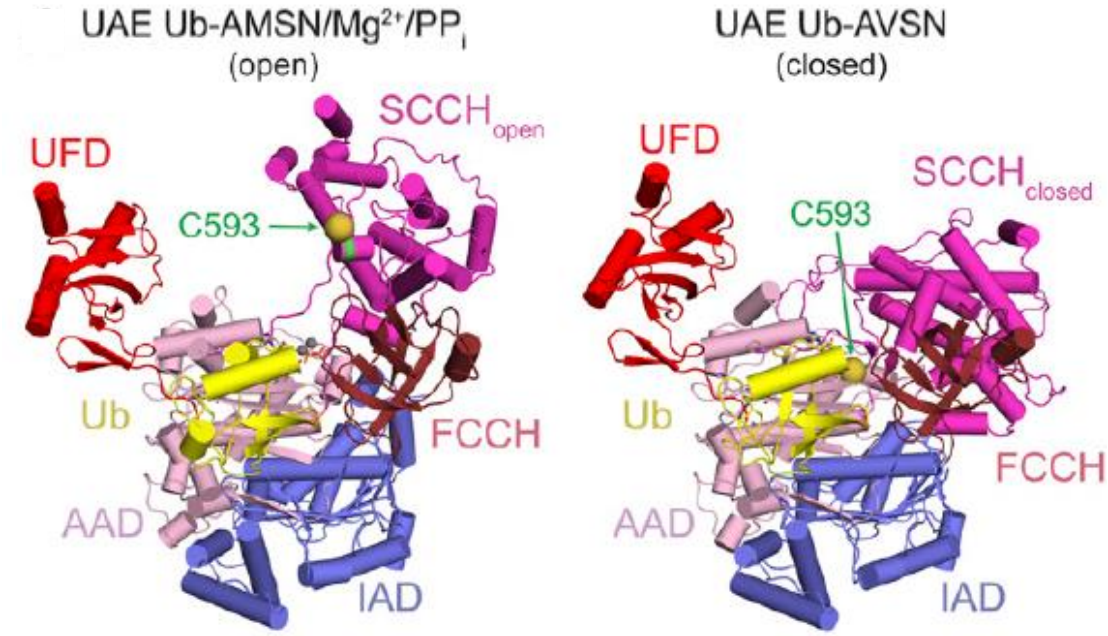


Figure 7. Structure of the ubiquitin E1 (UAE) in the 'open' or 'closed' conformations. Two ubiquitin constructs with AMP mimetic functional groups were used to trap the Ub E1 in two conformers to elucidate the intermediate structures of the UAE bound to Ub. The five E1 domains and Ub are colored on the crystal structures and the catalytic cysteine is labeled.
The figure was obtained from: Hann, Z. S. *et al.* Structural basis for adenylation and thioester bond formation in the ubiquitin E1. *Proc. Natl. Acad. Sci.* **116**, 15475–15484 (2019).

The overall structure of UBA1, without the UFD, is similar to the functional homodimers observed for the bacterial proteins MoeB and ThiF, which bind and activate specific UBLs (MoeD, ThiS) with charged thioesters to facilitate bioorganic chemistry. MoeB and ThiF each contain similar adenylation and catalytic cysteine

domains, but are active only after homodimerization into a reciprocal structure where each monomer activates the ATP binding domain of the other.[245,246]

These activated adenylation domains can bind to both ATP and a Mg^{2+} counterion and the unstructured C-terminus of the UBLs. The ATP binding pocket anchors the adenine ring through hydrophobic contacts and the α-phosphate through electrostatic interactions to the Mg^{2+} counterion, and this configuration induces steric strain between the α- and β-phosphates.[247] When MoeD or ThiS are bound to this site the highly flexible C-terminal GG motif is positioned near the posterior of the bound α-phosphate and react to form a mixed anhydride intermediate before shedding pyrophosphate to generate the stable adenyl ester.[188] After the release of pyrophosphate a large conformational shift of >20 Å occurs which brings the catalytic cysteine in close proximity to the adenyl ester. Nucleophilic attack by the thiol displaces the AMP leaving group to generate the enzyme-bound thioester.[242,244] In contrast to UBL signaling, where the protein is activated for downstream processes, MoeD and ThiS act solely as sulfur atom donors for the synthesis of molybdopterin or thiamine, respectively. To fulfill this role they are displaced from the thioester as thiocarboxylates and subsequently transport the labile sulfur atom to additional enzymes.[248,249]

Owing to the similar strategy of C-terminal adenyl esterification to prime UBLs for thioesterification, the adenylation domains of E1s share a ~20-30% sequence identity with ThiF.[250] The highest degree of sequence conservation is in the ATP binding domain, including the putative G-X-G-X-X-G nucleotide binding motif and most of the catalytic adenylation residues, which highlights the conservation of this adenylation mechanism between the biological kingdoms.[247] In UBL signaling

pathways the E1s are primarily responsible for segregating the pathway to only the target UBL, as the transthiolation reactions with E2s are seemingly less discriminative of the substrate. In the case of NEDD8 and Ub two arginine sidechains near the Ub C-terminus, R42 and R74, were shown to engage in a hydrogen bonding network and electrostatic interactions within the adenylation site, as mutations to either of these impaired the activation rates.[251,252] Although NEDD8 and Ub share a 60% sequence identity the NEDD8 E1 adenylation domain has evolved a cationic arginine in the active site to gate the binding pocket against the R72 of Ub.[253] In studies where the Ub C-terminal residue 72 was mutated from a cationic arginine to a leucine, mimicking the hydrophobic alanine of NEDD8, this mutant Ub was readily activated by the NEDD8 E1.[240] In other cases where the UBLs are less similar the C-termini would be excluded from binding and being adenylated.

The increased complexity of charging a protein with a high energy thioester which must be transferred to multiple enzymes is evidenced in the evolution from MoeB and ThiF homodimers into the single polypeptide or heterodimers found in the E1s of UBL signaling. In E1s a single adenylation site binds ATP, Mg^{2+}, and the UBL, and forms a similar UBL-AMP adenyl ester. Accompanying the release of pyrophosphate a dramatic rearrangement of the enzyme occurs where the SCCH domain, containing the catalytic cysteine, rotates ~115° – 130° to position the cysteine for nucleophilic attack. This conformational change includes a remodeling of the adenylation active site to eliminate many of the contacts necessary to bind ATP and facilitate the adenylation reaction. After AMP is released following the formation of the thioester in this “closed” conformation the enzyme reopens to display the activated

UBL for transfer to E2 enzymes.[247,254,255] The affinity of the E2s to bind to the E1 is predicated on the presence of the UFD which features a similar β-grasp fold and a complementary binding domain to the UBL.[256,257] In this "open" state the enzyme can become doubly loaded with an additional UBL bound with an equivalent of ATP and Mg^{2+} at the regenerated adenylation site. The accompanying conformational change of double loading the E1 increases the efficiency of the thioester transfer from E1s to E2s.[189,258]

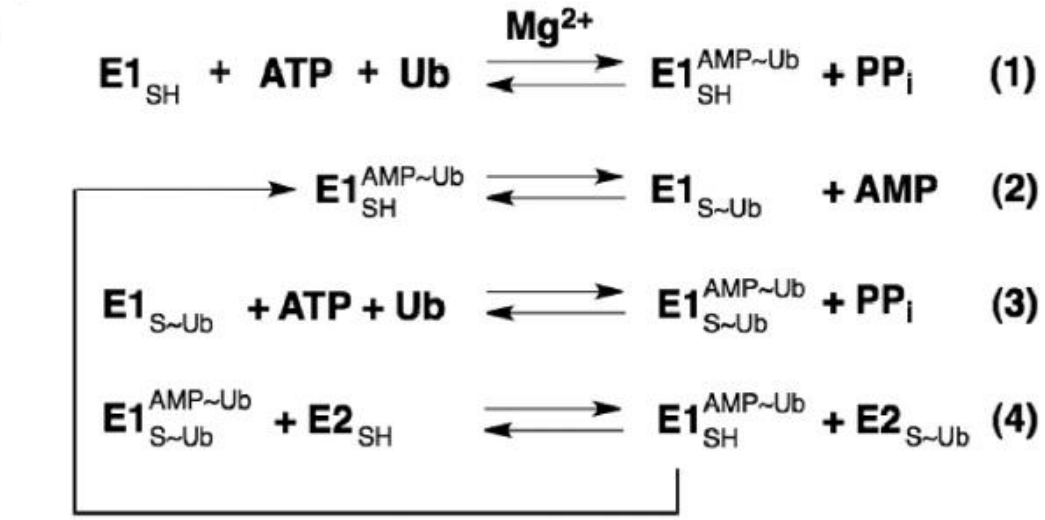


Figure 8. UBL activation mechanism. The ATP dependent loading and adenylation / catalytic cysteine transfer steps are shown.
This figure is reproduced from: An, H. & Statsyuk, A. V. Development of Activity-Based Probes for Ubiquitin and Ubiquitin-like Protein Signaling Pathways. *J. Am. Chem. Soc.* **135**, 16948–16962 (2013).

1.4 Ub signaling can be affected by post-translational modifications

Ub and UBL signaling are the best known proteinaceous post-translational modifications (PTMs) but in total the vast majority of PTMs involve small molecules (smPTM).[259–264] The phosphorylation of, primarily, alcohols by kinases and the acetylation of amines by acetylases make up the majority of these modifications.[265–270] Phosphorylation is the most abundant PTM in eukaryotes and is present on one-third of eukaryotic proteins under basal conditions.[271] For phosphorylation to occur the

kinase is charged with ATP or GTP and binds to threonine or serine sidechains on the substrate in an orientation to facilitate the cleavage and transfer of the γ-phosphate to the deprotonated alcohol.[272–277]

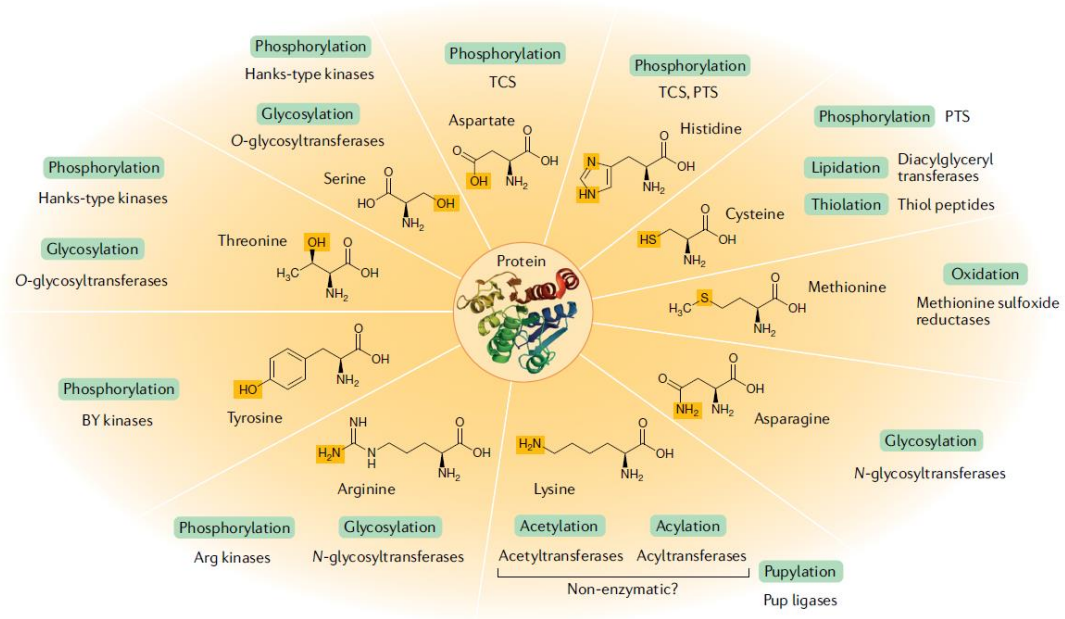


Figure 9. Protein post-translational modifications. The common modifications for sidechains are shown.
This figure is from: Macek, B. *et al.* Protein post-translational modifications in bacteria. *Nat. Rev. Microbiol.* **17**, 651–664 (2019)

This phosphoester introduces a significant negative charge from the two deprotonated oxygens, at physiological pH, which can induce local as well as tertiary structural perturbations. The primary effect of phosphorylation is to disrupt and or create a local hydrogen bonding network while forming new electrostatic interactions or repelling hydrophobic sidechains. Phosphorylation can therefore act as an allosteric regulator due to widespread structural changes that can modify the protein stability, substrate kinetics, or dynamics.[278–282] Phosphatases work in contrast to cleave the phosphate esters and regenerate the native sidechain. [283–286]

Estimates of acetylation propensity have indicated that upwards of 80% of the soluble proteome can be expressed as N-terminal acetates (NTA) on the nascent polypeptides in mammalian cells. Six N-terminal AcetylTransferases have been identified in humans and have a tight association with ribosomes to perform these modifications.[287–293] In many cases NTAs are removed after translation, and the N-terminus is cleaved to the proper starting residue, however NTAs can also impact subcellular localization and mark some proteins for secretion. NTAs can directly impart functionality as well, and have been shown to facilitate proper protein folding and increase the overall protein stability through rigidifying the N-terminus.[294–300] Some enzymatic processes rely on NTAs, and all NEDD8 E2s and the Ub E2 UBC12 require NTA to bind to E3s.[301,302] Lysine sidechain amines can similarly be modified by acetylation post-translationally and these modifications can be removed by the activity of deacetylases.[271,303–311] As with phosphorylation, which can modify the tertiary structure of a protein, acetylation adds steric bulk, can modify hydrogen bonding networks, and modifies the electrostatic nature of the sidechain by changing the cationic amine sidechain to an uncharged acetamide.[312–314] While the majority of acetylation is induced by acetyltransferases, using acetyl CoA as the acetate source, amines can be nucleophilic within the range of pH utilized in biological systems. In some cases, especially in environments with an elevated pH such as within the mitochondrial matrix, non-enzymatic acetylation can occur with acetyl CoA or other biomolecules with electrophilic acetate groups.[315–319]

The average pKa for N-terminal amines (Nα) lies within the range of ~7-9 and lysine sidechains have a pKa greater than ~10.5, with hydrogen bonded or salt-bridged amines having larger pKas and those in hydrophobic environments having depressed pKas.[320–322] Aside from the mitochondrial matrix which has an elevated pH due to proton flux to the inner membrane, most organelles and the cytosol have environments which are only slightly alkaline or moderately acidic.[323–332]

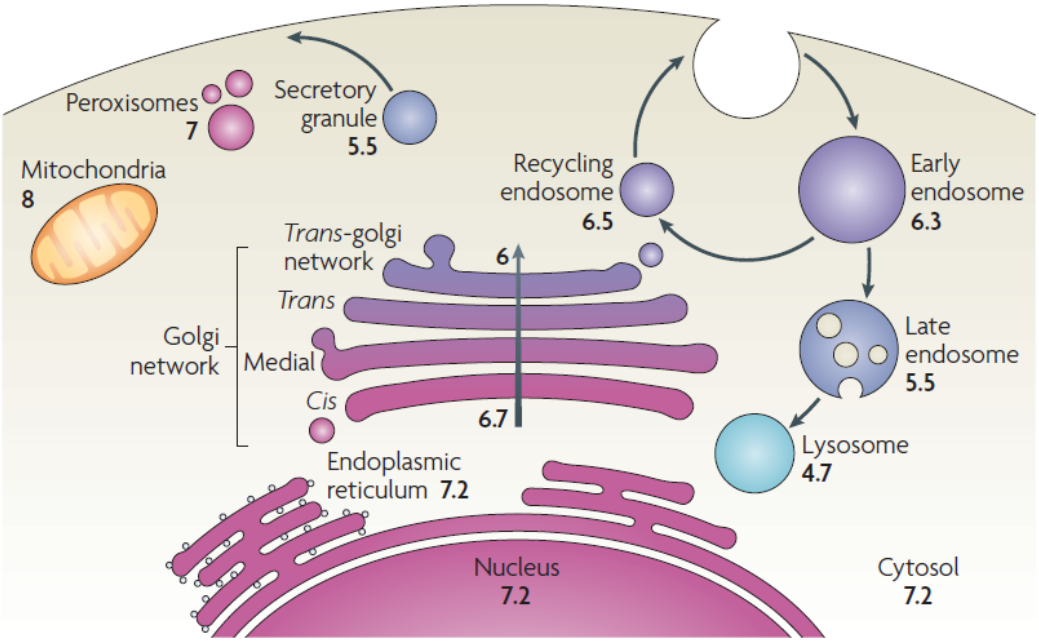


Figure 10. The pH of intracellular organelles. The primary organelles in eukaryotic cells are labeled with the common pH in their environments.
Figure from: Casey, J. R., Grinstein, S. & Orlowski, J. Sensors and regulators of intracellular pH. *Nat. Rev. Mol. Cell Biol.* **11**, 50–61 (2010).

N_α are therefore expected to be moderately reactive to electrophilic biomolecules whereas the majority of lysine sidechains (N_ζ) would be protonated and nonreactive. Despite this, proteomic analyses have detected a multitude of N_ζ modifications which are not enzyme catalyzed, most of which have implications for age-related degeneration and the progression of diseases. Four of the most common

nonenzymatic PTMs of amines are glycation with sugars, carbamylation with isocyanic acid, deamidation, and reactions with oxidation products generated from reactive oxygen species.[333,334]

1.4.1 Small-molecule post-translational modifications affect ubiquitin signaling pathways

Ubiquitin signaling pathways are modified by PTMs at the substrate, enzyme, and individual Ub level. Aside from polymeric ubiquitination, where Ub is a PTM of Ub, phosphorylation of Ub and the E3 Parkin is the best studied interconnection between smPTMs and Ub signaling. In response to mitochondrial outer membrane (MOM) depolarization the kinase PINK1 is recruited to the mitochondrion and phosphorylates Ub at S65 (pUb). S65 is adjacent to the hydrophobic recognition patch and phosphorylation modifies this binding site by introducing a negative charge.[335] Parkin contains a slightly modified UBD which specifically binds to pUb and causes Parkin to localize to the MOM. Parkin contains a ULD which is recognized and phosphorylated at S65 by PINK1, and this modifications induces a structural change in Parkin which activates the ligase functionality. The activated Parkin further ubiquitinates components of the MOM with polyUb chains which can interface with the autophagosomal membrane and allow for efficient mitophagy. This process is beneficial for cellular homeostasis, and mutations to Parkin or PINK1 have direct links to the severity of several neurodegenerative disorders including Parkinson's disease.

smPTMs have been detected on many other Ub E3 and E2 enzymes but most appear to be sporadic and are not associated with a gain of function. In many cases these PTMs perturb the tertiary structure to reduce the ability to bind Ub, transfer

activated Ub from one enzyme to the next, or generate the proper E3, E2, and substrate interface for Ub ligation. The polymorphism of E3 and E2 enzymes limits any general mechanisms for smPTMs to modify the Ub conjugation machinery, aside from a few PTMs of E2s occurring near the conserved catalytic cysteine site that interfaces with the E1.[336] Due to the large size of the Ub E1 many PTMs have been recorded although most occur on the surface sporadically. In cancerous cells the frequency of PTMs on the E1 are increased and include the enrichment of internal modifications affecting functionality depending on the cancer type.[337] Ubiquitination, acetylation, and phosphorylation are predicted to affect the activity in various ways including obstruction of the hydrophobic pocket for the adenine base of ATP, sterically block ATP binding, inhibit E2 binding to the UFD to correctly orient its catalytic cysteine with the Ub thioester, and prevent the conformational changes accompanying ATP binding or Ub transfer to the SCCH catalytic cysteine.[338]

Ub itself is readily modified by smPTMs (Figure 11) and phosphorylation is known to occur on all serines, 7/8 threonines, and the single tyrosine.[339,340] These modifications become more prevalent in cancer-derived cell lines but the phosphorylation of S65[341] and T12[342] are linked to the cellular processes of oxidative stress or the DNA damage response, respectively. Ub S65 phosphorylation has been studied in depth using both pUb and the Asp or Glu phosphomimetics to ascertain how the modification affects the Ub structure and recognition by binding partners.[335,341,343] S65E was preferentially introduced into K6 and K11 polyUb chains and excluded from K27 polyUb chains, and polyUb chains comprised of S65E were protected from DUB activity. E3 and E2 activities with the phosphomimetics were impaired *in cellulo* but

E1 loading was only modestly decreased.[343] Acetylation of all lysines except for K29 has been detected *in vivo*[344] but acetylated K6, K48, and K63 were the major population found in conjugated Ub.[345] Acetylation was shown to disrupt polyUb chain formation either directly through converting the free amine to an acetamide, precluding its conjugation, or by masking the positive charge binding interactions with the enzymatic machinery are disrupted. Histone deacetylase 6 is the only known Ub-active deacetylase and no acetylases have currently been identified.[338,345,346] Ub amines have been shown to react non-enzymatically with acetyl-CoA *in vitro*[347] which provides a possible mechanism for Ub signaling to be regulated dynamically depending on pH, intracellular localization, and in response to stressors.

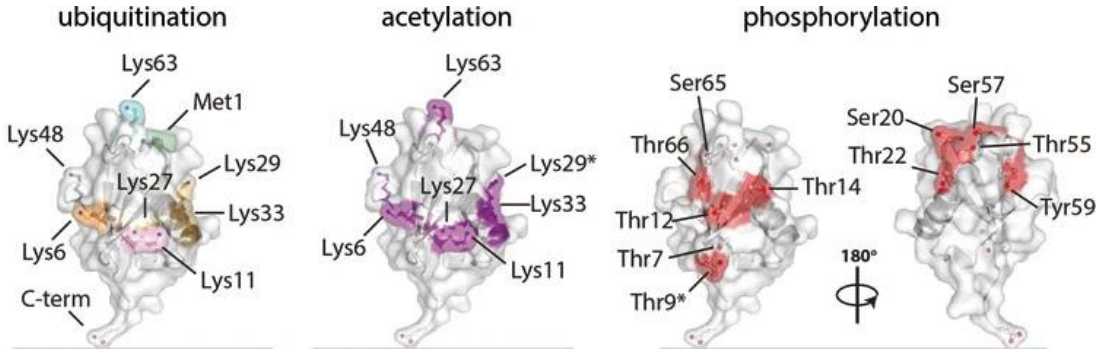


Figure 11. Many ubiquitin post-translational modifications have been detected. Ubiquitin is primarily modified through ubiquitination, phosphorylation, and acetylation at the indicated positions. This figure is from: Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).

1.5 Carbon dioxide is a poorly understood post-translational modifier with widespread physiological effects

High concentrations of carbon dioxide, or hypercapnia, can have profound effects on animal physiology including hyperventilation and an increased heart rate in mammals[348,349] and the breakdown of neuronal signaling capacity in fish.[350–352] On the microbiological level CO_2 is produced as a natural byproduct of the citric acid cycle

required for cellular respiration[353,354] and is utilized to maintain intracellular pH through the buffering capacity of the $HCO3^-$ / H^+ pair formed when dissolved CO2 reacts with water and generates carbonic acid.[355,356] Biochemically CO_2 is known to modify the activities of hemoglobin[357–360] and ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCo)[361–364] by forming stabilized carbamates on specific amines which affect O_2 binding affinities or activate the process of carbon fixation, respectively. However, in model organisms the widespread physiological effects of induced hypercapnia cannot be explained by any of the few known CO_2 binding proteins nor by perturbations to intracellular pH due to the increased concentration of aqueous CO_2.[365,366]

In *D. melanogaster* exposed to elevated CO_2 ~500 genes were up- or down-regulated compared to a control group. This manifested in a marked susceptibility to bacterial infections, rates of mortality due to bacterial infections, and issues with embryonic development including morphological defects, failures to hatch, and a reduction in the number of laid eggs.[367,368] Similar physiological consequences were observed in *C. elegans*, and after exposure to an environment with a large concentration of CO_2 the expression of ~500 genes were up- or downregulated by at least a factor of 2. Long term exposure to hypercapnic conditions similarly reduced development, egg production, and motility, but additionally increased the lifespan by 20-30% through unknown processes.[369,370] In the lungs of mice exposed to hypercapnia the expression profiles of ~700 genes were affected by elevated CO_2, but the physiological consequences were less discernible. Mice gestated with hypercapnia had an increased kidney size, predicted to be a consequence of bicarbonate reabsorption,

and thinner alveoli in the lungs.[371,372] Although less pronounced, similar changes in gene expression were observed in cultured mammalian cells.[373,374]

The similarity in the observable phenotypes of *D. melanogaster* and *C. elegans* alludes to a common biochemical response to CO_2. As only a few proteins are known to bind to bicarbonate directly, and many of the responses to hypercapnia can be replicated in pH controlled environments, the most likely effector would be transient carbamates formed with proteinaceous amines. Similar to nonenzymatic acetylation, all nucleophilic amines would be expected to react with CO_2 to form carbamates in some capacity. The formation of CO_2 carbamates, however, is readily reversible unless cationic functional groups or bound metal ions are positioned to stabilize the negative charge of the carbamate.[375] This property of carbamates would render the modification inconsequential to most proteins as any structural changes actuated by the cationic to acidic conversion would have a short lifetime and low population at physiological pH. However, for proteins which require certain structural features or unmodified amines to initiate signaling cascades even modest amounts of carbamylation could have widespread effects.

1.6 Research motivation and specific aims

The Ub signaling system is a complex network that encompasses the interplay of PTMs, UBL crosstalk, enzymatic processing, and the selective recognition of diverse polyUb chain topologies by UBDs. Perturbations in any of these components can alter the populations or relative proportions of polyUb chain types, potentially impairing cellular signaling pathways or disrupting the normal function of organelles. While

conventional biochemical techniques enable the study of direct interactions between these components, they often fail to provide detailed structural information or site-specific resolution. In this dissertation, I will demonstrate how I employed nuclear magnetic resonance (NMR) spectroscopy experiments and site-directed mutagenesis with a photocrosslinker to unravel some of the intricacies of interactions within Ub signaling pathways. To that end I accomplished four specific aims:

Specific aim 1: Analyze how carbon dioxide acts as a Ub PTM.

In collaboration with the group of Dr. Martin Cann at Druham University I ascertained how carbon dioxide binds to Ub. They were able to demonstrate that, surprisingly, Ub K33 and K48 were detected to be carbamylated by CO_2 in a novel proteomics assay. While the functions of K33 signaling are less well known, K48 polyUb is responsible for facilitating the proteasomal degradation of proteins which is key for maintaining cellular homeostasis. The observation that CO_2 interacts with Ub as a carbamate modification, and blocks polymerization through K48, points to a potential mechanism for the intracellular localization of Ub or extreme stress conditions to passively regulate Ub signaling.

Specific aim 2: Characterize the binding efficiency of UAE and NAE for Ub and RUB1.

The activation of UBLs by E1 enzymes is the first step to catalyze UBL-signaling pathways. E1s are implicated as being the primary gating mechanism for UBLs entering the various signaling pathways. Ub and NEDD8 have a high degree of sequence similarity, and NEDD8 has been shown to be modified by the UAE. Despite this, the processing efficiencies of E1s have only been measured indirectly, either

through the flux of radiolabeled ATP or by competition assays with fluorophores. I proposed that NMR spectroscopy can be used to directly observe the activation rates for UBLs with kinetic resolution. I explored this methodology with Ub and NEDD8 E1 enzymes to tease out the intricacies of residue 72 on the activation rates of Ub, NEDD8, and RUB1.

Specific aim 3: Characterize the structure and Ub-binding ability of hDDI1 ULD.

The ULD / UBA containing protein DDI1 was first identified as a proteasomal shuttle in yeast. Unexpectantly, the ULD was shown to bind to Ub. In higher organisms DDI has shed the UBA domain, and it was proposed that the ULD could maintain a Ub binding affinity to allow for DDI1 to associate with polyUb chains. In humans DDI1 is present as two paralogues, DDI1 and DDI2. DDI2 was shown to be incapable of binding to polyUb with meaningful affinities, thus putting into question whether the DDIs retain the Ub-binding functionality of the yeast ortholog. In this project I determined the solution structure of the DDI1 ULD and established the Ub-binding properties to determine how human DDI1 is associated with the UPS.

Specific aim 4: Develop a methodology to trap weak Ub-binding proteins through photo-crosslinking.

Many UBDs bind only weakly to Ub or polyUb chains to allow for Ub-signaling to be robust and highly dynamic. Capturing these binding events by immunoprecipitation or other affinity methods is impractical due to the low binding affinities. In addition, many of the current methodologies for UBD identification lack chain-specificity, which is necessary to place the UBDs into their proper contexts. We

have proposed that photo-crosslinking functional groups could be used to permanently trap these otherwise invisible associations. Using non-canonical amino acid incorporation I will demonstrate that Ub multimers can be generated with photo-crosslinking functional groups at multiple locations, and within different Ub chain types, and used to identify UBDs.

Chapter 2. Carbon dioxide is a reversible post-translational modification of Ub

Chapter 2 was adapted from two manuscripts:

Linthwaite, V. L. *et al.* Ubiquitin is a carbon dioxide–binding protein. *Sci. Adv.* **7**, eabi5507 (2021).

Pawloski, W., Komiyama, T., Kougentakis, C., Majumdar, A. & Fushman, D. Site-Specific Detection and Characterization of Ubiquitin Carbamylation. *Biochemistry* **61**, 712–721 (2022).

2.1 Carbon dioxide binds to certain ubiquitin amines

Aside from proteins that have biochemical mechanisms requiring the presence of carbamylated amines, little is known about how carbamylation affects protein structures and functions. The crux of the issue is that carbamylation is inherently a reversible reaction that occurs and dissipates independent of enzymatic activity. Any effect of carbamylation would therefore have to be observed in retrospect, and this makes the modification invisible to standard methods of detection by proteomics. [363,364] Carbamate esters are significantly more stable against hydrolysis than carbamate acids, and the group of Dr. Martin Cann at Durham University developed a method to trap CO_2 carbamates as ethyl esters that works in aqueous solutions and under physiological conditions.[378] In these reactions, cells were exposed to controlled levels of CO_2 in pH-buffered solutions followed by the addition of the alkylating agent triethyloxonium (TEO) as shown in Figure 12**Error! Reference source not found.**. TEO, applied as Meerwein's salt, was classically used to alkylate carboxyl sidechains of proteins, and was shown to effectively produce the ethyl carbamate ester in vitro. This methodology

was successfully used to pick out the ethylated carbamate of hemoglobin, in addition to the carboxylic acid of E7 being esterified in the tryptic peptide, after tryptic digestion and MS^2 analysis. This methodology was subsequently used *in vivo* to survey carbamates in *Arabidopsis thaliana* leaves.[378]

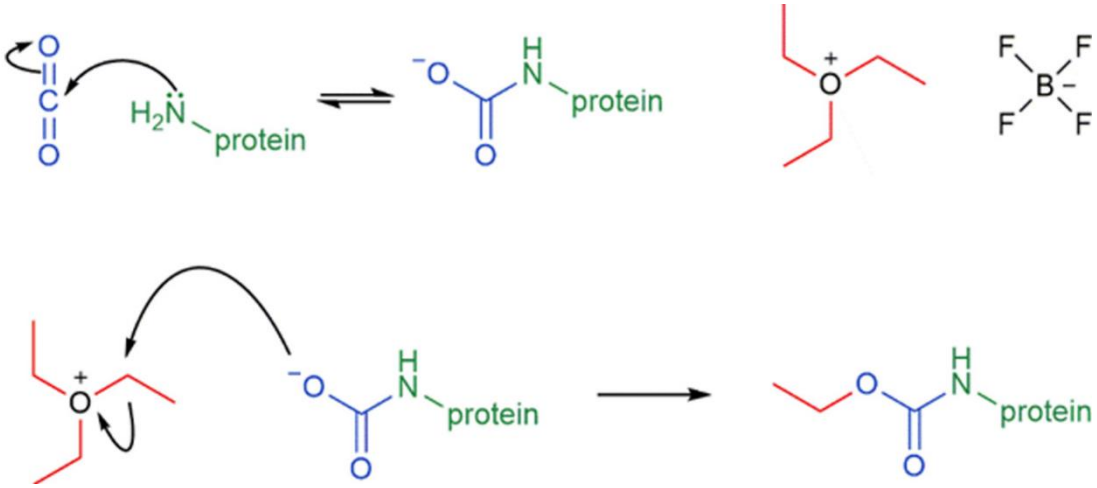


Figure 12. Carbamates can be stabilized as carbamate esters. The mechanism of carbamate formation is shown on the top, followed by the ethylation reaction with triethyloxonium.

The carbamate of RuBisCo, which makes up 40% of the weight of *Arabidopsis* leaves, was primarily detected, but six proteins previously having no functional association with carbamylation were discovered. Upon further investigations, Ub and a few additional proteins were suspected of containing novel carbamates of CO_2. *In vitro* analyses further corroborated that the ε-amines of K48 and K33 of Ub are the favored carbamylation sites. As there are no obvious structural features that would stabilize a carbamate on these amines, we postulated that these lysines have depressed pKas correlating to an enhanced reactivity with CO_2 in general.

My initial work on this project was to determine the pKa of the sidechain amines of Ub. NMR experiments report the chemical shift of nuclei, and these are influenced by their electronic environment. For acidic or basic sidechains the solution

pH affects the relative populations of the charged or uncharged states, corresponding to the Henderson-Hasselbalch equation shown in Equation 1, and the observed population-weighted chemical shift (δ_{obs}). The calculated fit parameters were the protonated (δ_i) and unprotonated (δ_f) chemical shift, pK_a, and the Hill coefficient (n_H).

$$\delta_{obs} = \delta_i - (\delta_i - \delta_f) \times \frac{10^{(pH-pKa)n_H}}{1 + 10^{(pH-pKa)n_H}} \quad (1)$$

Direct measurements of the chemical shift of the ionizable moieties provide the least ambiguous results, and both ^{15}N or ^{13}C chemical shifts of amines or carboxylic acids in sidechains have been monitored by one-dimensional (1D) NMR experiments at variable pH.[379,380] These experiments are often insufficient to provide site-specific resolution due to spectral crowding within the chemical shift ranges of the functional group. Additionally, the decreased gyromagnetic ratio of ^{13}C and ^{15}N nuclei relative to protons necessitates lengthy experiment times to account for the inherent lower sensitivity. To circumvent these issues the pKa of protein functional groups have classically been ascertained by following the ^{1}H, or ^{13}C chemical shifts of neighboring methylene groups by conventional NMR experiments. These types pKa derivations are prone to errors, especially those relying on proton chemical shifts, due to systemic effects on the chemical shift from structural rearrangements or the ionization states of distal functional groups.[381]

These issues are often overcome in protein NMR by using multidimensional experiments which start and end with magnetization on protons but are labeled with the ^{15}N or ^{13}C frequencies via scalar couplings. This is most effective for one-bond

magnetization transfers, but sidechain carboxylic acids lack an attached proton. Rapid solvent exchange in the *epsilon* amines of lysines or the *eta* N-H pair of the guanidino groups of arginines precludes sampling the ^{15}N frequency through one-bond couplings. To overcome these limitations multidimensional NMR experiments have been developed to directly sample the ionizing heteroatoms by passing magnetization from the protons of the neighboring methylene groups. [382–385] These experiments are generalized as Hε2CεNζ (H2CN) for Lys / Arg or H2CCO for Asp / Glu.

2.2 Determination of ubiquitin amine pKas by H2CN NMR experiments

For my work I utilized a pulse sequence written by Dr. Ananya Majumdar from Johns Hopkins University based on the premise of the original H2CN experiment.[384] At pH 6.8 I was able to observe crosspeaks corresponding to all seven lysines of Ub and the crosspeak corresponding to the N-terminal amine (Figure 13a). The crosspeaks pertaining to lysine residues were assigned by a H2CN experiment modified with a TOCSY mixing block which linked the H_ε-N_ζ crosspeak to the β, γ, and δ ^1H chemical shifts of the lysine sidechain, which I was able to obtain by a conventional sidechain assignment. The signal pertaining to the N-terminal amine was referenced from previous work. From these assignments I initially determined the Ub lysine pK_as by titrating Ub through the pH range of 8.6 to 12.3 over 15 points and following the $^1H_\varepsilon$-$^{15}N_\zeta$ chemical shifts recorded by H2CN experiments (Figure 13b). Six of the seven lysines were observable throughout the entirety of this pH titration but the signal for K27 decreased at elevated pH due to exchange broadening. Additionally, in an effort

to maximize the experiment for lysine detection, the intensity for the N-terminal amine was too low to be accurately followed.

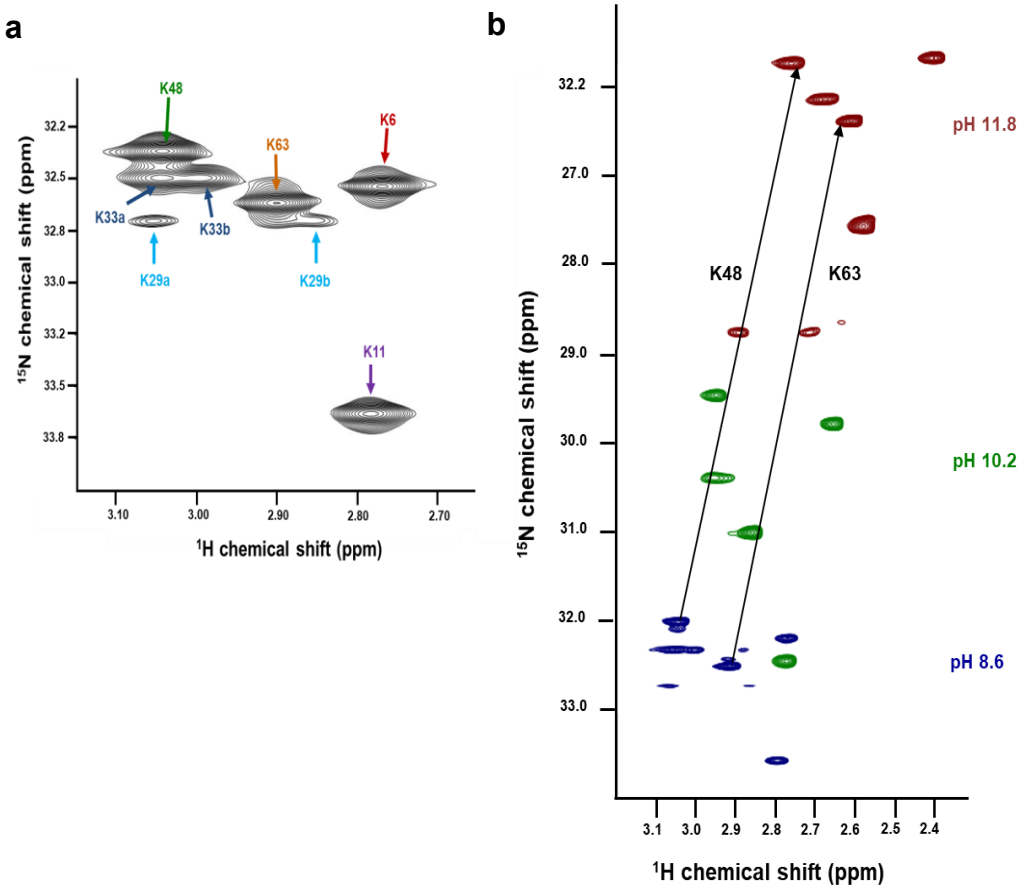


Figure 13. H2CN NMR experiments were used to determine the ubiquitin lysine pK$_a$s. (a) The H2CN spectrum recorded for ubiquitin at pH 8.6. K29 and K33 are observed with magnetically inequivalent protons at low pH. **(b)** H2CN spectra of ubiquitin measured at three pH values. The chemical shift trajectories for K48 and K63 are highlighted.

To complete the pKa measurements I performed an additional ^{1}H-^{13}C HSQC-based titration over the same pH range and sufficiently resolved crosspeaks were monitored to independently assess pH-dependent effects on Ub sidechain resonances. $^{1}H_{\varepsilon}$-$^{13}C_{\varepsilon}$ crosspeaks could be resolved for K6, K11, K33, and K27, and the N-terminal $^{1}H_{\alpha}$-$^{13}C_{\alpha}$ crosspeak could be followed for M1. The pH-dependent chemical shifts from

the ^{1}H-^{15}N (H2CN) or ^{1}H-^{13}C (HSQC) crosspeaks were fit to a Henderson-Hasselbalch equation modified with a Hill coefficient and these fits are shown in Figure 14.

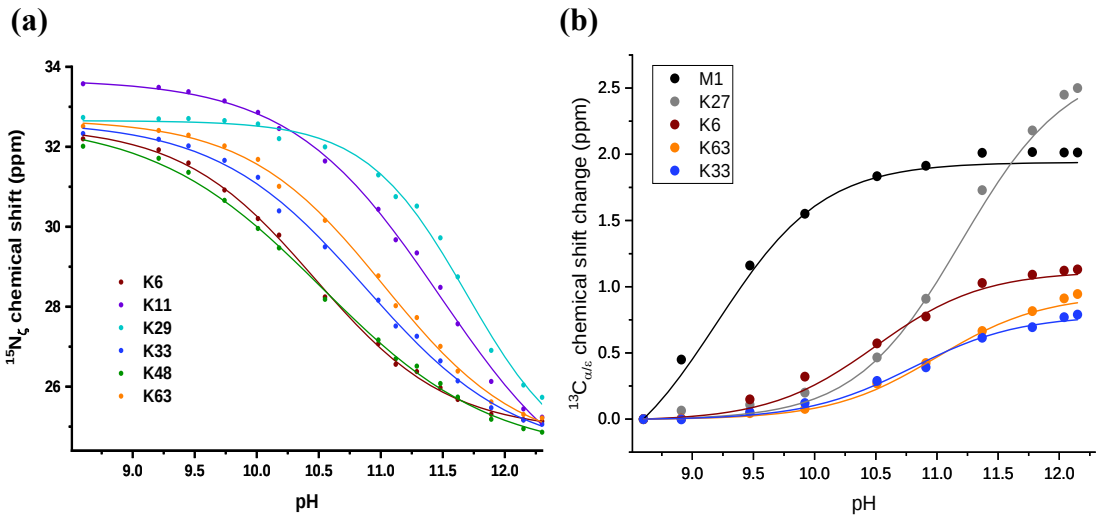


Figure 14. Ubiquitin lysine sidechain ^{15}N / ^{13}C chemical shift changes through the pH titration. (a) The $^{15}N_\zeta$ chemical shift recorded at each pH (dots) was fit to a Henderson-Hasselbalch equation modified with the Hill coefficient (solid lines) to determine the sidechain pK_a. **(b)** The ^{13}C chemical shifts for the well-resolved residues were similarly fit for pK_a.

In principle, both ^{1}H and ^{13}C / ^{15}N pH dependent chemical shift trajectories can be fit to determine sidechain pK_as. For both ^{1}H-^{15}N H2CN and the ^{1}H-^{13}C HSQC titrations the pK_as calculated from proton chemical shifts were in good agreement with the primary reporting nuclei, aside from discrepancies in the HSQC experiments due to peak overlap for K63 and pH-dependent splitting of the proton resonances for K27 (Table 1). Overall, we were able to obtain the pK_as for all the Ub amines using $^{15}N_\zeta$ and/or $^{13}C_{\varepsilon/\alpha}$ reporters and standard U-^{15}N / ^{13}C labeling procedures. Where applicable we consider the ^{15}N resonances to be the most accurate reporters due to direct sampling of the ionizable nuclei. However, for three amines with both $^{15}N_\zeta$ and $^{13}C_\varepsilon$ reporters,

K6, K11, and K33, pK_as derived from both ^{1}H-^{15}N H2CN or ^{1}H-^{13}C HSQC series were in good agreement (Table 2). The Hill coefficient is commonly applied in pKa measurements to account for non-Henderson-Hasselbalch behavior due to charge coupling interactions between ionizable groups. For most amines the calculated pKas were not significantly altered by the addition of the Hill coefficient, but for K29 the fit and calculated pKa were significantly improved due to the difficulty of fitting to the high pKa. These are discussed further in the appendix (Appendix 1.1).

Table 1. pKa values for ubiquitin amines derived from pH titration series recorded by ^{1}H-^{15}N H2CN or ^{1}H-^{13}C HSQC experiments, using the indicated nuclei as reporters.

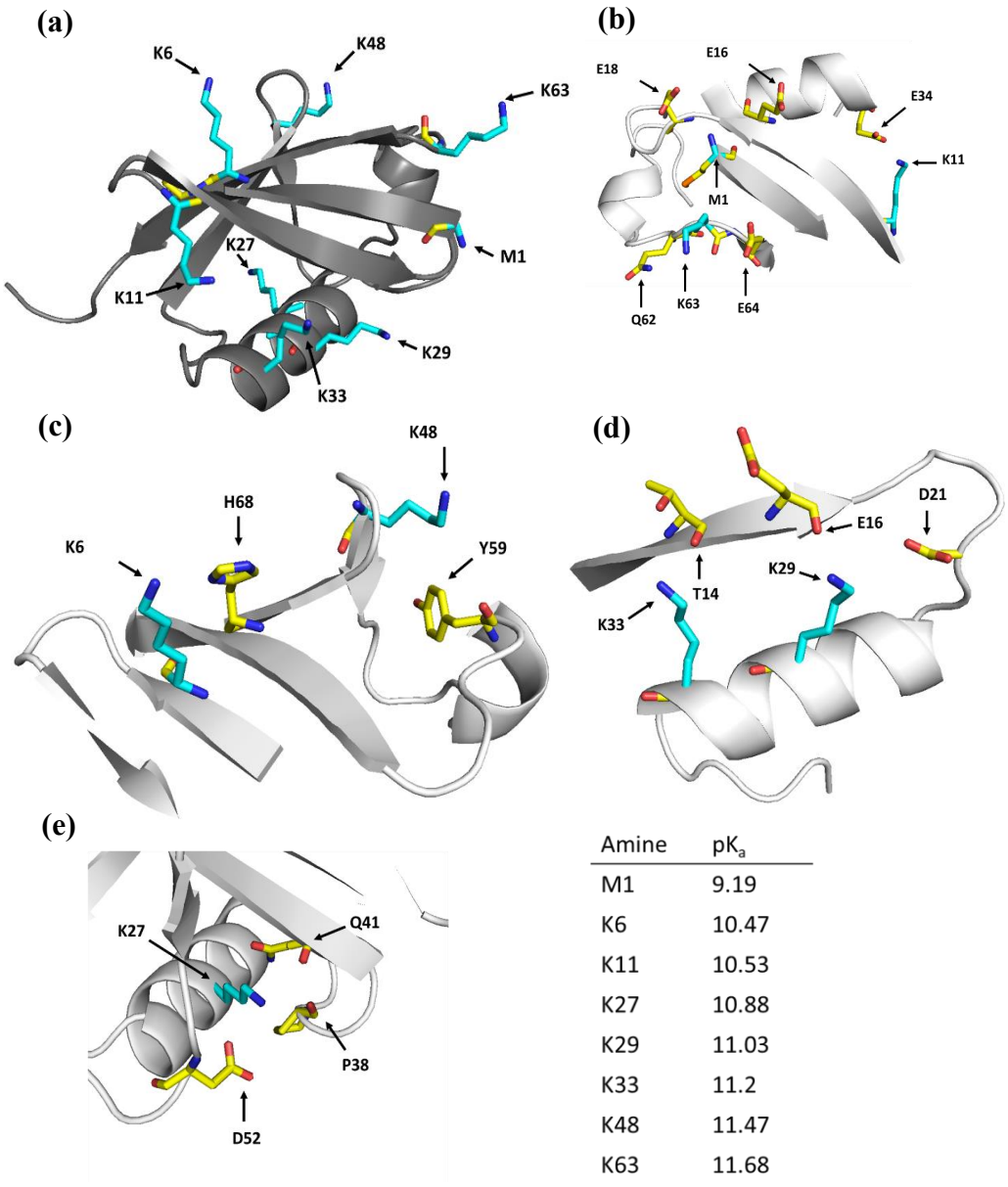
Amine	$^{15}N\zeta$	$^{1}H_{\varepsilon}$ (H2CN)	$^{13}C_{\varepsilon}$	$^{1}H_{\varepsilon}$ (HSQC)
M1	--	--	9.19 ± 0.05	9.21 ± 0.01
K6	10.47 ± 0.02	10.57 ± 0.03	10.5 ± 0.1	10.55 ± 0.04
K48	10.53 ± 0.02	10.72 ± 0.06	--	--
K33	10.88 ± 0.03	10.99 ± 0.03	10.8 ± 0.1	10.73 ± 0.06
K63	11.03 ± 0.02	11.06 ± 0.03	11.0 ± 0.1	11.5 ± 0.2[a]
K27	--	--	11.2 ± 0.1	11.5 ± 0.1[b]
K11	11.47 ± 0.03	11.46 ± 0.07	--	--
K29	11.68 ± 0.03	11.85 ± 0.03	--	--

[a] Peak overlap in the proton dimension of the HSQC experiments for K63 decreased the accuracy in calculating this pK_a. [b] The proton resonances for K27 are degenerate at lower pH values but begin to split into unique resonances as the pH is increased. This is combined with exchange broadening at elevated pH which limited an accurate proton chemical shift determination and decreased the accuracy in this calculated pK_a.

Table 2. pKa values for ubiquitin amines derived from pH titration series recorded by ^{1}H-^{15}N H2CN or ^{1}H-^{13}C HSQC experiments. The calculated Hill coefficients are included from both titrations.

Amine	$^{15}N_{\zeta}$ pK_a	Hill coefficient	$^{13}C_{\alpha/\varepsilon}$ pK_a	Hill coefficient
M1	--	--	9.19 ± 0.05	0.9 ± 0.1
K6	10.47 ± 0.02	0.80 ± 0.03	10.5 ± 0.1	0.8 ± 0.2
K48	10.53 ± 0.02	0.63 ± 0.03	--	--
K33	10.88 ± 0.03	0.72 ± 0.03	10.8 ± 0.1	0.7 ± 0.4
K63	11.03 ± 0.02	0.79 ± 0.03	11.0 ± 0.1	0.6 ± 0.3
K27	--	--	11.2 ± 0.1	0.9 ± 0.1
K11	11.47 ± 0.03	0.73 ± 0.03	--	--
K29	11.68 ± 0.03	1.09 ± 0.06	--	--

The electrostatic environment and hydrogen bonding network will affect the ionization state of solvent-exchangeable functional groups at a given pH. In Ub the sidechain amine pKas are unsurprising in this regard and reflect electrostatic and hydrogen bonds (Figure 15). Both K6 and K48 were demonstrated to have pKas near the theoretical value for a non-bonded *epsilon* amine of ~10.5. The pKa of K33 is the next highest at 10.88. Supposing that the efficiency of the carbamylation reaction is proportional to amine pKa, K6 should form a carbamate similarly to K33 and K48. M1, with a pKa of 9.19, would be expected to be the most carbamylated amine.



Amine	pK_a
M1	9.19
K6	10.47
K11	10.53
K27	10.88
K29	11.03
K33	11.2
K48	11.47
K63	11.68

Figure 15. Ubiquitin lysines and their immediate environment highlighted on the solution structure (1D3Z). (a) From the solution structure of Ubiquitin (Ub) it can be gleaned that the eight amines are positioned in unique environments. (b) Unusual for an N-terminal amine, in Ub M1 is gated by two glutamate sidechains which were reported to increase the pK_a to 9.14. K11 is the only surface-exposed lysine in an ionic bond to E34. K63 is gated by a glutamate and glutamine, however no significant bonding interactions were observed in MD simulations. (c) K6 and K48 neighbor an imidazole (H68) and phenol (Y59) functional group, respectively. Although limited information is available on the extent of the interactions between these groups, H68 is suspected to be an important component for the carbamylation of K6 by CO_2. (d) K33 and K29 are positioned at the interface between the helix and β-strand 2. K33 forms a hydrogen bond with the carbonyl of T14, whereas K29 is positioned in polar contacts to the backbone carbonyl of E16 and the sidechain of D21. (e) K27 features a sidechain amine buried into the hydrophobic core and is charge stabilized by D52, the carbonyl of P38, and a polar contact to the backbone carbonyl of Q41. The amine pK_as are shown in the bottom right table.

2.3 Ubiquitin carbamates were detected with 13C direct-detection NMR experiments

We used ^{13}C-NMR as an orthologous method to TEO trapping to confirm Ub CO_2-binding sites at K33 and K48, to investigate other sites for carbamate formation not identified by MS-MS, and as a direct demonstration of the carbamate PTM on the native protein. I initially mixed 1 mM $^{13}C/^{15}N$-labeled Ub with 100 mM $NaH^{13}CO_3$ and observed three peaks in 1D-^{13}C NMR spectra which were not present in spectra for either Ub or $NaH^{13}CO_3$ alone as shown in Figure 16a. The chemical shifts of these new signals: 163.25, 164.77, and 164.96 ppm, are consistent with the empirical range for carbamate PTMs.

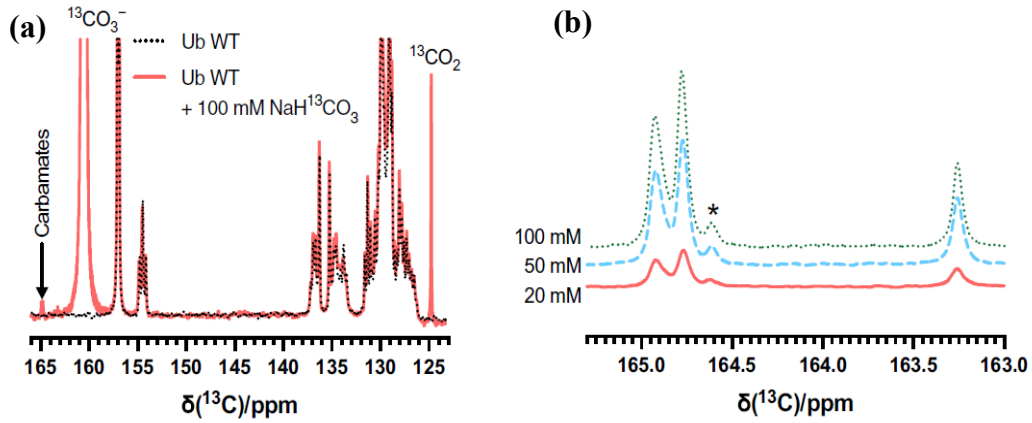


Figure 16. Carbon dioxide forms carbamates with ubiquitin. (a) After adding ^{13}C-enriched bicarbonate new signals corresponding to carbon dioxide, bicarbonate/carbonate, and ubiquitin carbamates are evident. **(b)** The intensity of these carbamate signals increases with bicarbonate concentration. The star indicates a low intensity carbamate signal only seen with 5 mM ubiquitin.

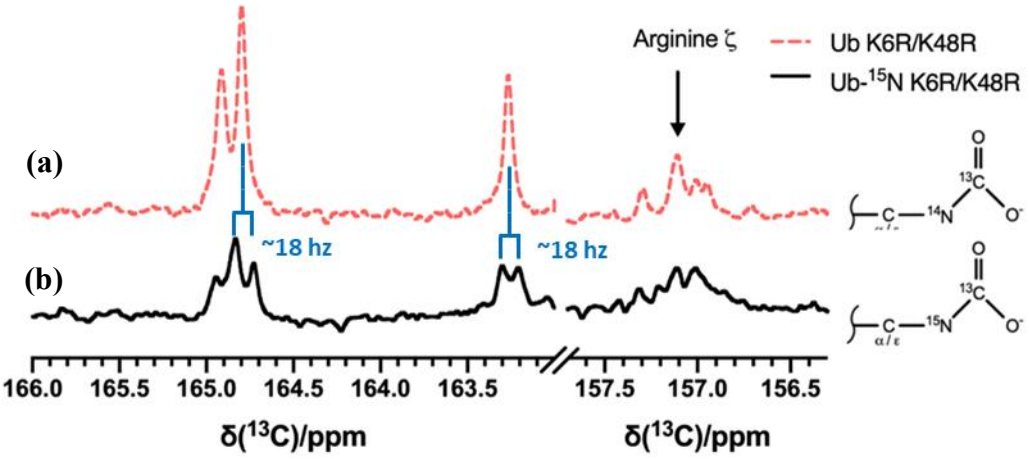
1 mM or 5 mM unlabeled Ub was subsequently exchanged into buffers containing 20, 50, or 100 mM $NaH^{13}CO_3$, and we observed that the intensities of these carbamate signals increased with increasing bicarbonate or Ub concentration

(Appendix 1.2), supporting the hypothesis that they are the product of reversible carbamate formation on Ub (Figure 16b).

The ratios of the peak intensities were unaltered at 20, 50, or 100 mM $NaH^{13}CO_3$. We therefore used 100 mM $NaH^{13}CO_3$ to identify the carbamate forming residues. To investigate which amines these signals correspond to, I introduced site-directed Lys to Arg mutations to eliminate the carbamates stepwise. Using readily available protein, I assayed the double lysine mutant Ub K48R/K6R which yielded partial reductions in the intensities for the two signals at 164.77 and 164.96 ppm, but not for the third signal at 163.25 ppm (Figure 17a). In a duplicate experiment with uniformly ^{15}N-labeled Ub K48R/K6R these remaining signals were split into doublets; the observed splittings of ~18.5 Hz were consistent with a one-bond ^{15}N-^{13}C coupling, indicating that the ^{13}C atoms responsible for these signals are directly bonded to ^{15}N atoms and confirms that these are carbamate signals (Figure 17b).

To unambiguously identify the carbamate signals we performed consecutive lysine to arginine Ub mutations starting with K48R and adding K33R, K6R, and K63R (Figure 18a, b, c, d, respectively), which revealed that the signals at 164.77 and 164.96 ppm were constituted of overlapping carbamate signals from K6/K33 and K48/K63, respectively. The ^{13}C-NMR spectrum of the quadruple mutant (K6R/K33R/K48R/K63R) shown in Figure 18d matched the spectrum of a lysine-free Ub (K0) as shown in Figure 18e. This demonstrates that carbamylation was not evident

at K11, K27 or K29. Replacement of all Ub lysines with arginine in Ub K0 did not affect the carbamate resonance at 163.25 ppm.



The only potential nucleophilic nitrogen atoms remaining in Ub K0 are the *N*-terminal amino group and, possibly, the imidazole of H68. *N*-acetyl-L-histidine cannot form adducts with CO_2 but is identified in the binding pockets of proteins that can interact with CO_2. To examine if H68, the single histidine residue in Ub, could be responsible for this remaining signal, a H68A mutation was introduced into Ub. This mutation altered the population of carbamylated K6 (Figure 19a), likely reflecting an altered pK_a of K6 due to the proximity of its ε-amino group to the imidazole of H68 but did not affect the resonance at 163.25 ppm.

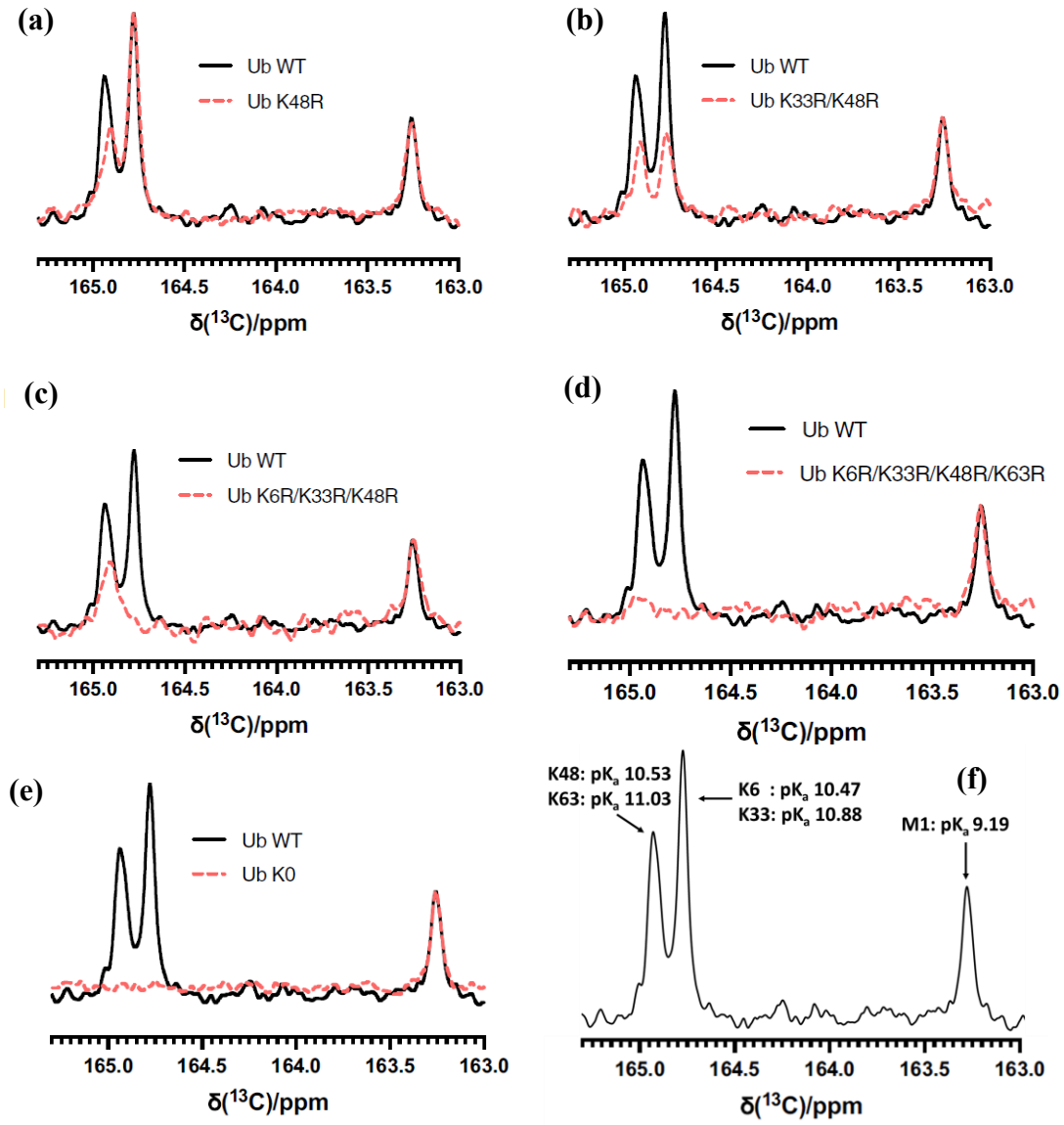


Figure 17. Carbamate resonances disappear with Lys to Arg mutagenesis. The Lys carbamate resonances in ubiquitin were identified by a reduction in intensity from stepwise K48R **(a)**, K33R **(b)**, K6R **(c)**, and K63R **(d)** mutations. **(e)** The spectrum for K6R/K33R/K48R/K63R matched "K0" ubiquitin where all Lys have been mutated to Arg. **(f)** Ubiquitin carbamate ^{13}C resonance assignments.

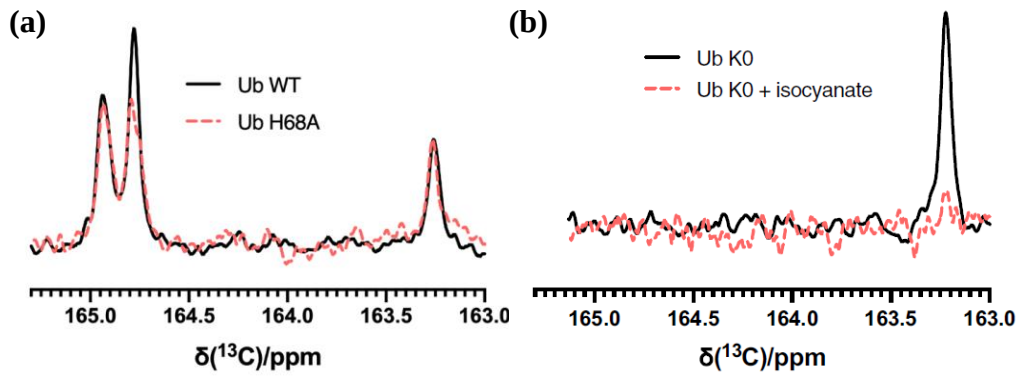


Figure 18. The upfield carbamate resonance comes from the N-terminal amine. (a) Mutating H68 in ubiquitin reduces the intensity of the K6 carbamate but not the upfield resonance. **(b)** The final carbamate resonance disappears after carbamoylation with isocyanic acid.

We reacted Ub K0 with one equivalent of sodium cyanate which ablated the carbamate resonance at 163.25 ppm (Figure 19b) identifying it as originating from carbamylation of the *N*-terminal amine. Based on these results we concluded that the Ub amines of M1, K6, K33, K48, and K63 form carbamates *in vitro*. Surprisingly the peak intensities for the five carbamylated amines were roughly equivalent, and the observed population of carbamates on N-terminal amine (pKa of 9.19) were similar to K63 (pKa of 11.03) despite the former having 70x the uncharged and reactive form at pH 7.4. These observations additionally conflicted with MS-MS measurements which only detected carbamates on two lysines, K33 and K48. In an effort to reconcile these observations and to tease out the carbamylation kinetics we needed a system which mimics the formation of CO_2 carbamates while being an irreversible modification (Figure 21a). Isocyanic acid has a similar structure to CO_2 and binding mode to amines.[386,387] Due to these properties isocyanic acid was classically utilized to determine the carbamate site and population in hemoglobin. We initially tested the

validity of the reaction of isocyanic acid with Ub amines using 1D ^{15}N NMR experiments.

2.4 The rates of Ub amine carbamylation can be measured using 15N-detected NMR experiments

Despite the limitations of sensitivity and resolution, ^{15}N-detected NMR spectroscopy can be used to directly observe the ionization state, dynamics, and reactions occurring with nitrogen groups undergoing rapid proton exchange with water. The limited chemical shift dispersion of solvent-exposed amine resonances often limits the ability to determine site-specific properties of these groups within proteins. However, 1D ^{15}N direct detection experiments have been used successfully in special circumstances to determine the amine pKas in a small peptide [380] or to study atypical lysines engineered into the hydrophobic core of staphylococcal nuclease [379]. In the case of Ub the seven lysine sidechains and the N-terminal amine are positioned in unique environments owing to electrostatic, steric, and hydrogen bonding interactions. These environmental factors influence unique sidechain dynamics for the Ub lysines [385] and were expected to confer a wide range of sidechain pK_as and spread out the ^{15}N resonances at elevated pH. At pH values below 9 and above 10 many of the Ub lysine sidechain ^{15}N resonances can be distinctly observed by 1D ^{15}N NMR experiments but not all can be resolved into isolated peaks. At the intermediate of these pH values,

around 9.3, all the lysines and N-terminal amine have unique chemical shifts that could be assigned with N_ζ chemical shifts recorded from H2CN experiments (Figure 20).

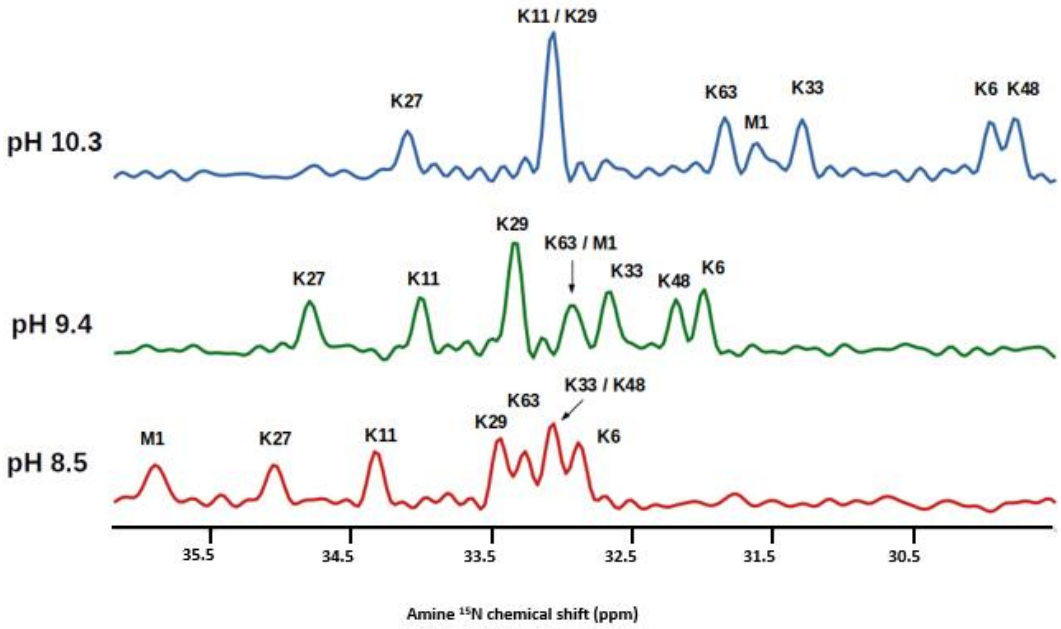


Figure 19. Aligned representative 1D 15N spectra recorded at different pH. The 15N resonances for ubiquitin amines are labeled on these spectra and show that signals overlap in each spectrum.

The carbamylation reactions of Ub were ascertained by ^{15}N direct-detection spectroscopy at pH 9.2 by adding 24 molar equivalents of sodium cyanate (3x per amine) to 3 mM U-^{15}N Ub, and the reaction was followed for two days by monitoring the ablation of the amine signal intensities (Figure 21b). Three distinct reactivity groups were gleaned for the Ub amines: the N-terminal amine resonance disappeared within 10 hours, K6, K33, and K48 signals decayed to ~20% after two days, and the remaining lysines K11, K27, K29, and K63 demonstrated limited reactivity, with K27 and K29 being markedly unreactive (Figure 22). Two resonances were initially observed for K63, and the intensity of the downfield signal appeared to increase with a rate proportional to carbamylation at the N-terminal amine. This suggests that the chemical

shift for the sidechain amine of K63 is affected by the ionization or carbamylation state of the N-terminal amine. The reactivities of the sidechain amines align with the pKas. M1, with a pK_a near the experimental pH of 9.2, reacted the quickest, followed by K6, K33, and K48 as the group of lysines with sidechain pK_a values below 11. The reaction rates for the other amines are significantly depressed in comparison but remain distributed with respect to pK_a values, aside from the anomalously low reactivity for the substantially buried sidechain of K27. While demonstrative, the limited pH window required for resolution of the ^{15}N resonances limited the utility of these experiments to broadly characterizing these reactions.

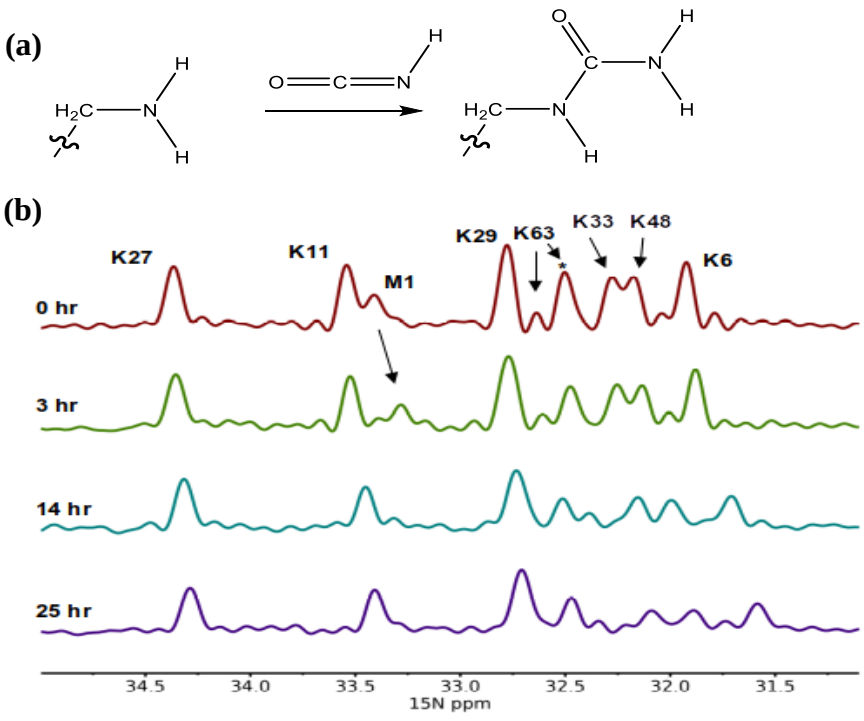


Figure 20. Ubiquitin amines are carbamylated by isocyanic acid. (a) reaction scheme for isocyanic acid carbamylation of amines. (b) 1D ^{15}N spectra of ubiquitin amines at several timepoints after the addition of sodium cyanate at pH 9.2. A slight upfield shift was observed for all of the amines and is reflective of an increase in pH, presumed to be due to the degradation of cyanate. The chemical shift for K63 presented as two signals, denoted as K63 and K63*, which coalesced over time into the downfield chemical shift.

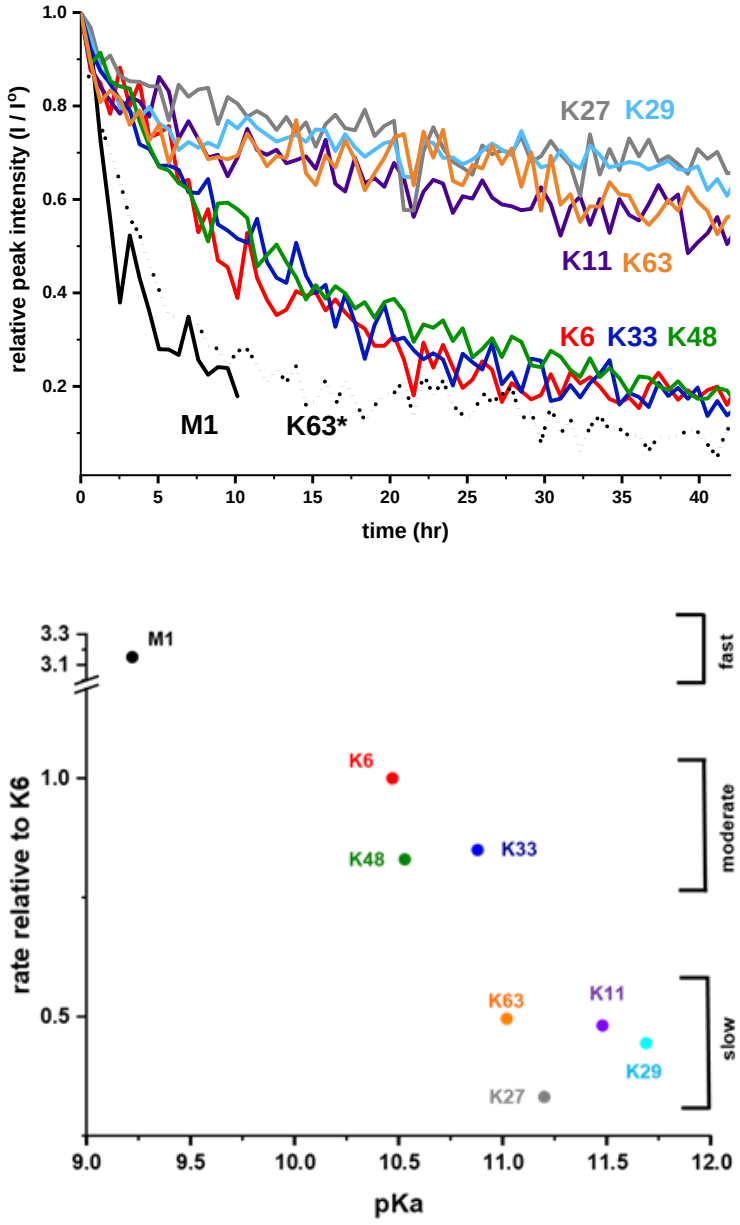


Figure 21. Isocyanic acid carbamylation rates are proportional to the sidechain pKa. The top panel shows the decrease in peak intensity for the indicated amines during the reaction timecourse. The reaction rates relative to pK_a are grouped in the bottom panel.

2.5 H2CN experiments were used to measure the rates of ubiquitin carbamylation

To circumvent the sensitivity and resolution issues with ^{15}N direct detection we proposed that these carbamylation reactions could be followed by H2CN experiment series. The H2CN experiment provided sufficient Ub amine chemical shift dispersions at all pH values to allow characterization of residue-specific carbamylation kinetics at physiological pH and to assess how these reaction rates are affected by environmental factors. The reaction with cyanate was initially probed at pH 7.5, 27.0°C, and 15 molar equivalents of sodium cyanate per amine, and the pseudo first-order reaction kinetics were observed over a 12-hour period (Figure 23a).

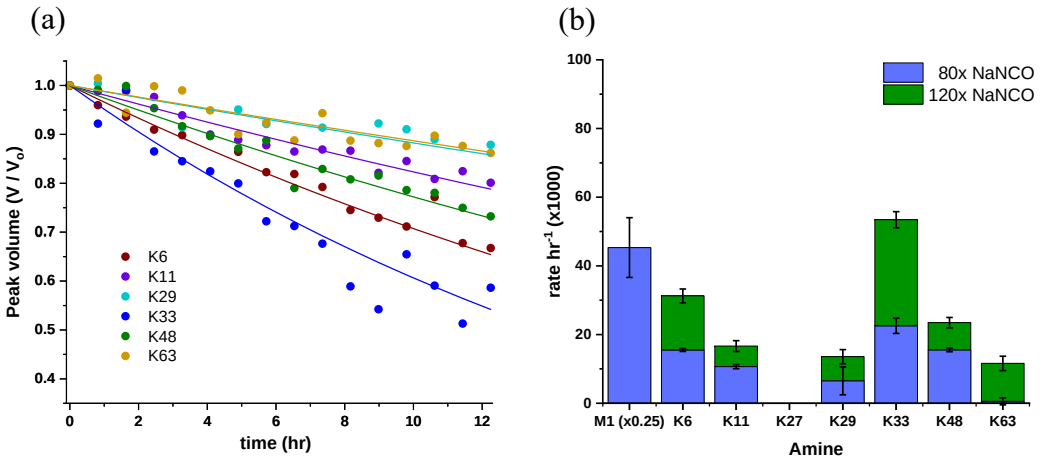


Figure 22. The carbamylation of Ub amines was measured by H2CN experiments. (a) The decrease in signal intensity for each visible amine due to carbamylation are overlayed with an exponential fit. (b) The carbamylation rates are shown for two sodium cyanate concentrations at 27.0°C. The error bars shown in the graph correspond to the standard error of the mean obtained from the linear regression model.

In contrast to our 1D ^{15}N data, the amine reactivities under these conditions did not fall into distinct reactivity groups but generally reflected the spread of sidechain pK_a. K33 exhibits an exceptionally quick carbamylation rate relative to its pK_a value,

only reacting slower than the N-terminal amine which reacted too quickly for any signal to be observed. K63 (pK_a 11.03) was also carbamylated anomalously slower than K11 (pK_a 11.47) and K29 (pK_a 11.68) with respect to pK_a values. These results were reproduced at a lower sodium cyanate concentration of 10 molar equivalents, however the quick reaction rate for the N-terminal amine could be measured and K63 was unreactive to carbamylation (Figure 23b). We expected a linear relationship between isocyanic acid concentration and the reaction rates, but some of the lysines exhibited a greater sensitivity. For a 1.5x increase in sodium cyanate concentration, K33 reacted 2.4x faster, K29 and K6 2.0x faster, K48 and K11 showed a rate increase consistent with the increase in isocyanic acid concentration of 1.5-fold, and appreciable carbamylation was observed at K63.

We attempted to assess the reaction enthalpies for the observable lysines by running temperature series. The rate of carbamylation was measured at three temperatures, 27.0°C, 32.0°C, and 37.0°C, for both 10 and 15 molar amine-equivalents of sodium cyanate (Figure 24a,b). Under both conditions the reaction rates for the low pK_a amines of K6, K33, and K48 increased in a stepwise manner with increasing temperature. This is contrasted with the reaction rates for the other observable lysines which either increase at 32.0°C and then level off (K11) or increase at 32.0°C followed by a decline in reaction rate at higher temperature (K29, K63). From 27.0°C to 37.0°C the rates increased roughly 4-fold for K6, K33, and K48 in contrast to modest overall increases for the others. The temperature dependent carbamylation rates for these three amines, at both sodium cyanate concentrations, were fit to the Eyring equation to determine the enthalpies of activation (Chapter 7.3.1). The average enthalpies of 23

kcal mol^{-1} (10x cyanate) and 17 kcal mol^{-1} (15x cyanate) are similar to the reported activation energies for Ub lysine acetylation of 22.6 kcal mol^{-1}.

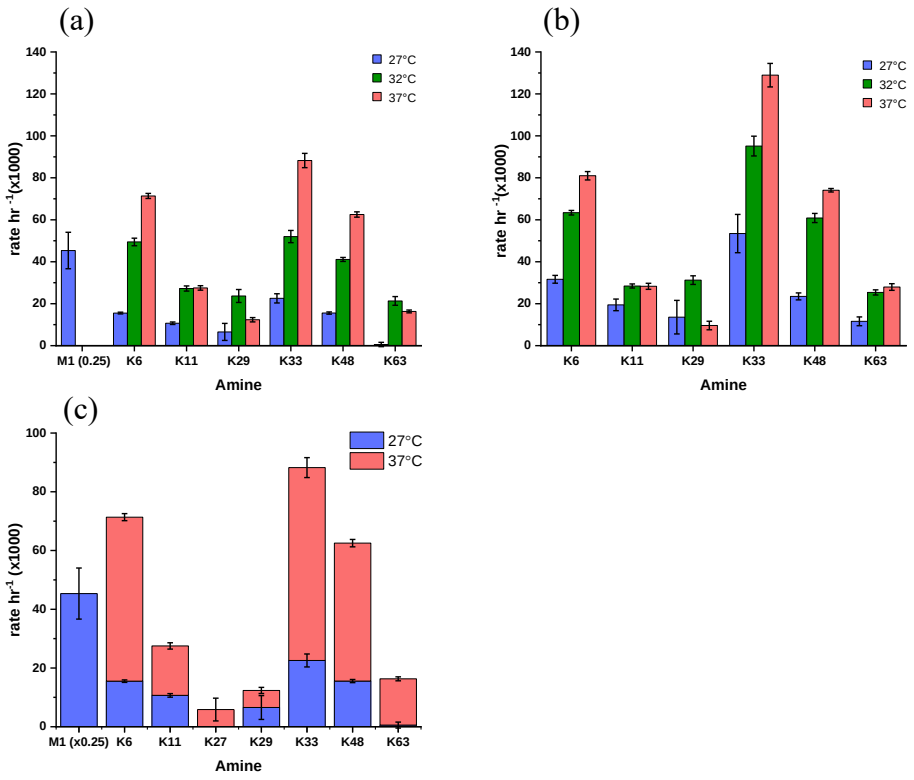


Figure 23. Carbamylation rates determined from H2CN series at variable temperature. The carbamylation rates were measured with (a) 10 or (b) 15 equivalents of sodium cyanate at pH 7.5. (c) the rates for 37°C (red) are overlayed on 27°C (blue) at 80x sodium cyanate. The error bars shown in the graph correspond to the standard error of the mean obtained from the linear regression model.

The pH dependence of the reaction rates was characterized at five additional points above pH 7.5 at 27.0°C. Amine carbamylation by isocyanic acid is known to proceed with the neutral forms of both components and is reported to be a second-order reaction. As the pK_a of isocyanic acid is ~3.5 under standard conditions, the reaction rates are theoretically constant in the pH range above 5.0 and ~1.5 units below the amine pK_as before decreasing as the pH approaches the pK_a. Mostly consistent with

these expectations, for all the Ub amines the reaction rates increased at each pH step approaching the respective pK_as, leveled off near the pK_as, and significantly decreased at higher pH as shown in Figure 25. The reaction occurring at the N-terminal amine could be followed at neutral and high pH but was too fast to be observed at intermediate pH values and likely adheres to a similar trend. At pH values below 10.3 the K33 sidechain (pK_a 10.88) is particularly susceptible to carbamylation and reacts faster than the lower pK_a sidechains of K6 (pK_a 10.47) and K48 (pK_a 10.53). As the pH approaches these sidechain pK_as, this disparity vanishes, and the reaction rates become distributed reflective of their pK_as. This is coupled with an anomalously high reactivity rate observed for K29 at this pH, and possibly indicates a pH-induced rearrangement of the contacts between the β2 strand and the α-helix.

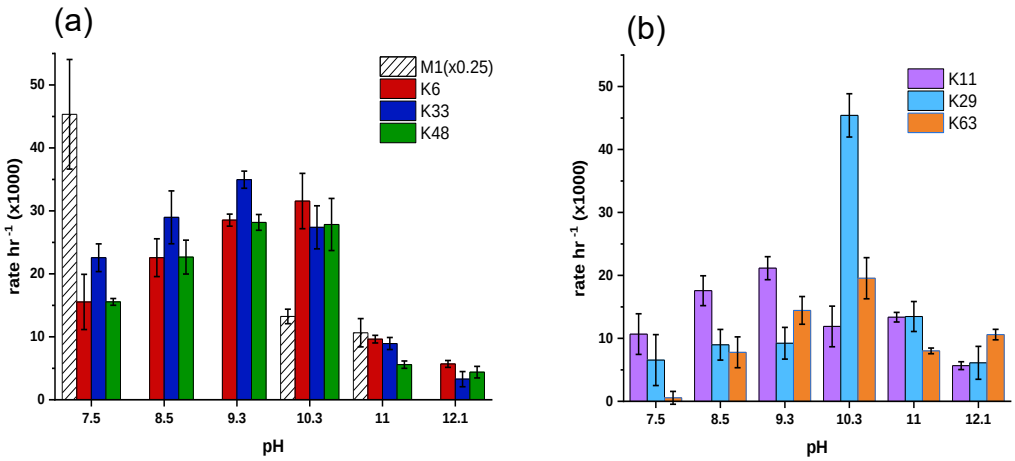


Figure 24. Carbamylation rates are affected by pH. H2CN reaction series were performed at 27 °C and at the indicated pH. (a) data for the four low pKa amines (M1, K6, K33, K48). (b) data for the three pKa amines with high pKas. The crosspeak volume for K27 could not be measured at higher pH and was excluded from analysis. The error bars shown in the graph correspond to the standard error of the mean obtained from the linear regression model.

Ub is particularly resilient to structural destabilization by mutation or PTMs, however complete shielding the sidechain amines by hydrophobic Alloc or Boc

protecting groups has been shown to make Ub insoluble in water and decrease the overall structural stability. During some carbamylation studies we did observe visible precipitation during extended reaction times (36+ hours) when the number of carbamylated amines per Ub exceeded 5-6 on average, as determined by electrospray ionization mass spectrometry shown in Figure 26. Within our 12.5-hour assay, where the average degree of carbamylation was similarly measured to be approximately 2, the concentration of soluble Ub was confirmed by 1D ^{1}H series where the total amide (6.55ppm – 9.70ppm) and hydrophobic methyl (-0.10ppm – -0.60 ppm) signal integration was monitored after the addition of sodium cyanate.

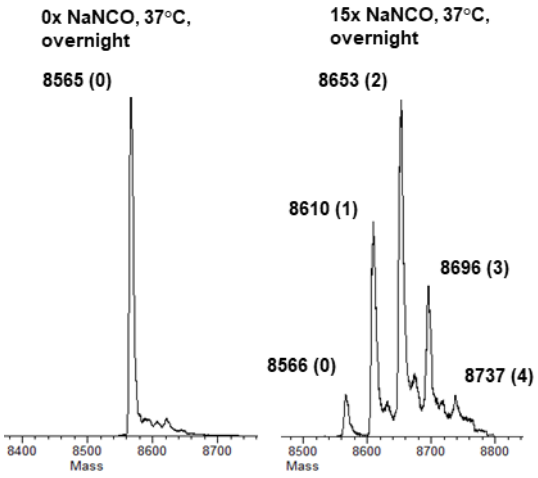


Figure 25. Deconvoluted mass spectra for ubiquitin incubated overnight with or without sodium cyanate. Peaks are labeled with the respective observed mass and the number of carbamylated amines.

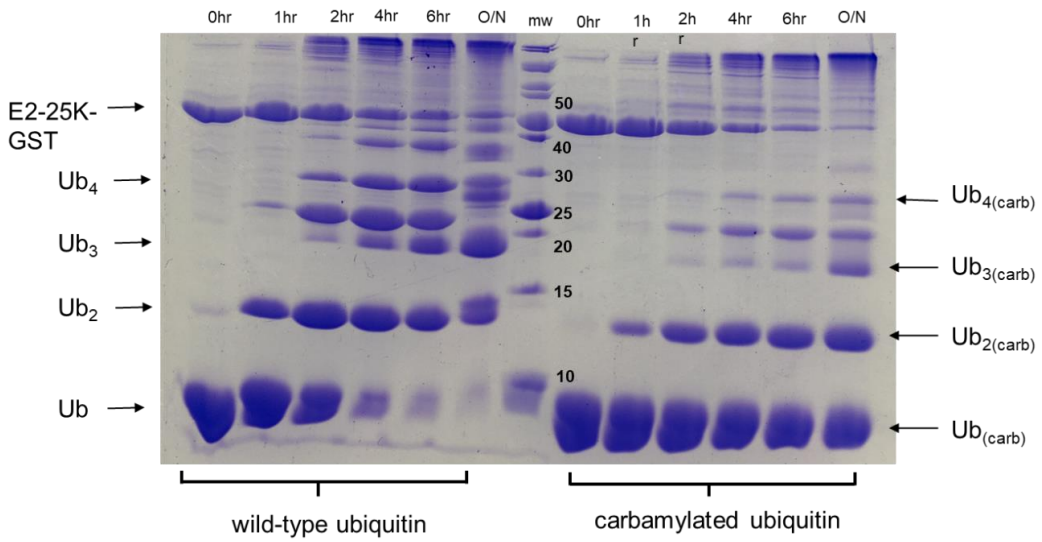


Figure 26. SDS-PAGE gel showing that Ub which has been incubated overnight with sodium cyanate at 37°C can be recognized by E1 and E2 (E2-25K) enzymes and assembled into polyubiquitin chains. After an overnight conjugation reaction the majority of non-carbamylated monomeric ubiquitin is depleted due to the construction of polymeric Ub chains of varying lengths. Significant amounts of carbamylated monomeric Ub can still be observed over the same timeframe, and this is coupled with less intense polymeric Ub bands. The residual monomeric Ub either has a carbamate on K48, directly blocking chain elongation, or some other carbamylated amine which interferes with enzyme recognition or processing.

Under conditions of elevated temperature or at an intermediate pH, where maximal carbamylation rates were observed, no decrease in bulk signal intensity was evident. To confirm the viability of Ub after carbamylation, a conjugation assay with the E1 and E2 enzymes that assemble K48-linked Ub chains [388] was performed after an overnight incubation of Ub with 15 molar equivalents of isocyanic acid at 37°C. Aside from the decreased population of higher order chains due to the blocked conjugation for Ub $K48_{carb}$, the unmodified Ub was effectively polymerized into chains.

This observation confirms that carbamylation of Ub, presumedly at the sidechain amines of K48 but could include others, can inhibit the formation of K48 polyUb chains and can affect the rates or propensity of protein degradation.

2.6 Conclusion: Four ubiquitin amines are carbamylated *in vitro*

From the NMR data in totality, I was able to corroborate with MS-MS analyses that Ub lysines K33 and K48 of Ub form carbamates with CO_2 under physiological conditions. The sidechain amines of these two residues have pKas within the normal range for hydrogen- or non-bonded amines, respectively, and are also within the same range as K6 and K63. In the presence $^{13}CO_2$ (aq) we have demonstrated that all four of these lysines and M1, with a significantly lower pKa common to N-terminal amines, form carbamates. To quantify carbamylation kinetics for these amines I performed experiments with the CO_2 mimic isocyanic acid at variable pH and temperatures. Similarly to the qualitative $^{13}CO_2$ results, M1, K6, K33, and K48 were all carbamylated by isocyanic acid at physiological pH. Relative to the other lysine sidechains K33 appears to be particularly susceptible to carbamylation at physiological pH and temperature, whereas K63 was resistant to carbamylation by isocyanic acid.

The group of Dr. Martin Cann performed *in vitro* and *in cellulo* studies to complement these results and demonstrated how physiologically relevant concentrations of CO_2 can affect Ub signaling. Ub conjugation rates by the E1 and E2/E3 combinations specific for building K48 or K63 polyUb chains were measured from 0.0 mM (hypocapnia) to 3.0 mM CO_2 (hypercapnia). From hypo- to hypercapnia K48 polyUb formation was decreased by 12% whereas K63 polyUb formation was unaffected. These results point to the K48 carbamate having a long enough lifetime to inhibit the E2 binding to $K48_{carb}$ whereas the K63 carbamate, which was not detected by TEO trapping, is too labile to be physiologically relevant. The K48 polyUb E2/E3 were not detected to form carbamates by MS-MS, however alternative carbamate sites

at the other Ub amines could contribute to the decrease in K48 polyUb formation. Additionally, they have shown that NF-κB signaling is altered by hypercapnia in HEK 293 cells. In replicate experiments where addition WT, K48R, or K63R Ub were endogenously expressed only the K48R experiments were insensitive to elevated CO_2.

The specificity for carbamylation at the four Ub amines with the lowest pK_as proposes a mechanism for intensity of specific polyUb signals to be modulated by environmental pH and the flux of CO_2. Evidence has been presented that hyper- or hypocapnic conditions can modulate K48 polyUb signaling *in cellulo*, and thereby affect the process of protein turnover by proteasomal degradation which is necessary for cellular homeostasis. By this same token, the innately low pK_as and enhanced susceptibility to carbamylation for the amines of M1, K6, and K33 would render these polyUb chains subject to regulation by passive carbamylation within different organelles, cell types, and localization within the body due to divergencies in pH and CO_2 concentrations. This effect would manifest as a modulation of the population and length of homogeneous-linkage Ub chains, where the direct consequences of these chain types are well known, in addition to regulating the generation of mixed-linkage unbranched and branched Ub chains that have enigmatic physiological responses.

In this work I have showcased the utility of using NMR-based assays to observe post-translational modifications occurring at sidechain functional groups. While I have focused on the carbamylation of lysine sidechain amines these same experimental principles are applicable to other reactions and functional groups. Carbamylation is only one potential Ub lysine modifier, and the effects of other non-enzymatic PTMs

remain elusive. I hope that this work and methodology can seed future endeavors to understand how Ub signaling pathways are modified by biomolecules.

2.7 Contributions and acknowledgements

The initial spark for this collaboration was facilitated by Dr. George Lorimer and was based on the discovery from the Cann lab that Ub appeared to bind to carbon dioxide. The ^{15}N 1D-experiments included in this work were acquired by Dr. Christos Kougentakis at Johns Hopkins University. He and Dr. Ananya Majumdar advised me on the measurement of pK_a values. Dr. Majumdar additionally provided ^{15}N 1D and H2CN pulse sequences and provided advice on their setup. Dr. Clark Ridge from the Food and Drug Administration tested the applicability of using ^{15}N-detected NMR experiments with a Varian spectrometer at University of Maryland.

I performed all of the NMR data acquisition, processing, and analysis with advice from Dr. David Fushman. I grew most of the protein material used in these NMR studies, although other material was provided by Dr. Apurva Chaturvedi.

Chapter 3. The kinetics of ubiquitin-like protein activation by E1 enzymes.

3.1 The kinetics of UBL activation have only been measured indirectly

The NEDD8 E1 contains an evolutionary stopgap which prevents Ub from binding at the adenylation site and entering the NEDD8 signaling pathway. The residue located four positions away from the C-terminus is an Arg in Ub (R72) and an Ala in NEDD8 (A72), and the cationic Arg is specifically repelled in the active site. The sequence of Ub is highly conserved from yeast to human differing only by four residues, whereas the yeast equivalent to NEDD8 (RUB1) differs by 27 residues. This includes position 72, which is the site of the critical Arg in Ub, which is a polar but uncharged Thr in RUB1. There is a sparse understanding of how these three similar, but unique, proteins are otherwise differentiated by E1, E2, E3, and DUB enzymes from both the Ub or NEDD8 families.

The activation of Ubls by E1s and the subsequent transfer to E2s is commonly detected indirectly, either by radiolabeled [γ-^{32}P]pyrophosphate from the hydrolysis of ATP in the activation step[240,389] or by performing competition assays with the UBLs conjugated to a fluorescent tag.[390] Our lab has previously developed a mass spectrometry (MS) based assay which detected the E1-induced C-terminal thioesterification of UBLs directly, and we were able to directly demonstrate how different UBLs and mutations to their binding domains dictate the thioesterification efficiencies of the UAE and NAE. Quantitation by this method involved performing the reaction with a ^{15}N-labeled UBL, UAE or NAE, and the thiol-containing small molecule Mesna to transthiolate the E1 – UBL thioester and yield a free UBL – Mesna thioester of higher mass. These reactions were quenched with EDTA at desired

timepoints and spiked with the same UBL without isotopic labeling for detection by electron-spray ionization MS.[391] While demonstrative, the limited time resolution and reliance on the careful addition of an external standard for quantitation prompted us to develop an alternative method to follow UBL activation reactions with higher fidelity.

NMR spectroscopy is uniquely able to observe kinetic processes with site specific resolution but can be limited by signal overlap and a poor signal to noise ratio. With the advent of isotopic labeling schemes and quick-repetition $^{1}H - ^{15}N$ or $^{1}H - ^{13}C$ correlation experiments 2D spectra of proteins with significantly enhanced signal intensities can be acquired in less than a minute. This ability to measure spectra with high kinetic resolution and continuous monitoring *in situ* has allowed for investigations into kinetic processes otherwise invisible to other forms of spectroscopy.[392–394] It is well established that in Ub dimers (Ub2) the Ub C-terminal amide involved in the isopeptide linkage (Distal G76) has unique ^{1}H and ^{15}N chemical shifts compared to the free C-terminus (Proximal G76).[395] We proposed that upon thioesterification with Mesna a similar change in chemical shift would be observed, and the disappearance for the signal corresponding to UBL C-termini, G76 in Ub and RUB1 / NEDD8, could be directly related to activation by E1s (Figure 27). In this work we will demonstrate the ability of 2D NMR experiments to monitor the enzymatic thioesterification of UBLs and show how single-site mutations to residue 72 dramatically perturb the processing by UAE or NAE. We will further expand upon these results by determining the thioesterification Km values for Ub, RUB1, and the residue 72 A, T, and R constructs with UAE and NAE to highlight how the sidechain of residue 72 dictates these E1 processing efficiencies.

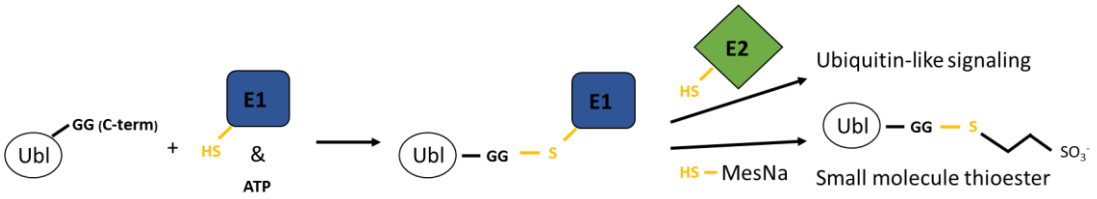


Figure 27. Mechanism for the small molecule thioesterification of UBLs.

3.2 2D NMR measurements of thioesterification reactions

We initially observed the thioesterification product with WT Ub and Ub E1, where 100 μM Ub was mixed with 4 mM ATP, 4 mM MgCl2, and 40mM Mesna buffered with sodium phosphate pH 8, and the reaction was started by the addition of Ub E1 to a final concentration of 300 nM. The 1H - 15N crosspeak for the canonical C-terminal GG motif found in UBLs is readily distinguished as a strong signal with a proton chemical shift at ~7.9 ppm and a nitrogen chemical shift at ~115 ppm at physiological pH and temperature. Throughout the course of the thioesterification reaction this peak begins to lose intensity concomitant with a new peak appearing at 1H chemical shift of 8.45 ppm and 15N chemical shift of 109 ppm (Figure 28). G75 and L73 similarly exhibit a significant change in chemical shift as the reaction progresses, however R72 and R74 have a minimal chemical shift perturbation.

A similar change in chemical shifts was observed during Ub2 formation with a K48-specific E2 UBE2K, and this change in chemical shift could be reversed upon the addition of OTUB1, a deubiquitinase which cleaves K48 polyUb chains at the isopeptide linkage. No other significant chemical shift or intensity changes were observed, and the average intensities for the other backbone amides were subsequently used as an internal standard. The rate of thioesterification scaled linearly with the

concentration of E1 and Mesna and showed no response to saturating ATP, and pseudo first-order rate constants were used to describe the reaction rates.

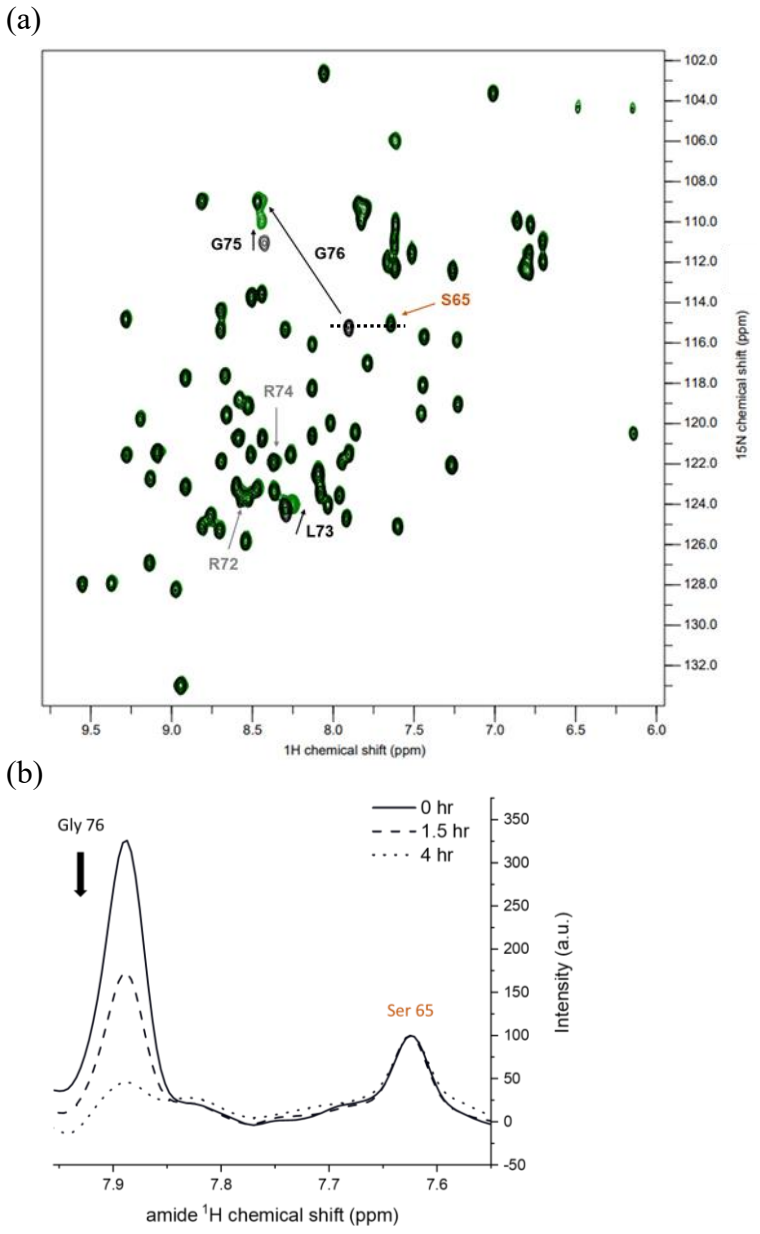


Figure 28. Formation of the ubiquitin thioester changes the chemical shift of G76. The rate of ubiquitin activation is recorded through the decrease in signal intensity for G76 as ubiquitin is thioesterified. (a) Overlay of ^{1}H-^{15}N HMQC experiments showing ubiquitin WT (black) and ubiquitin thioester (green). Major backbone amide chemical shifts are indicated with black arrows while minor shifts are indicated with grey arrows. (b) Three 1D-projections are shown from the ^{15}N chemical shift of 115.2 ppm at three timepoints after the reaction was started.

The same procedure was repeated using UAE and NAE with Ub and Rub1 where residue 72 was present as an R (WT Ub), T (WT Rub1), or A (WT NEDD8) sidechain and with WT NEDD8. The speed of thioesterification by NAE is shown to be principally determined by the residue at position 72 for all of the constructs we tested (Figure 29 and Table 3). NEDD8 72A(WT), as the primary substrate, was processed faster than both RUB1 or Ub 72A mutants which were reacted at roughly the same rate. RUB1 and Ub 72T (RUB1 WT) constructs were less reactive but still processed efficiently and with the same speed, and the 72R constructs (Ub WT) were completely non-reactive, as has been previously reported. The picture of UAE affinities is more muddled and, although modulated by the residue at position 72, is affected by other structural factors (Figure 29 and Table 3). RUB1 and Ub 72R (Ub WT) react the quickest, although surprisingly the rate of RUB1 72R thioesterification is nearly 2x quicker than WT Ub. NEDD8 was processed about 1/2 the rate of WT Ub, Rub1 72A and Ub 72T were processed at ~1/3rd the rate, Ub 72A at ~1/5th the speed, and Rub1 72T (Rub1 WT) reacted the slowest at ~1/10th the rate of WT Ub.

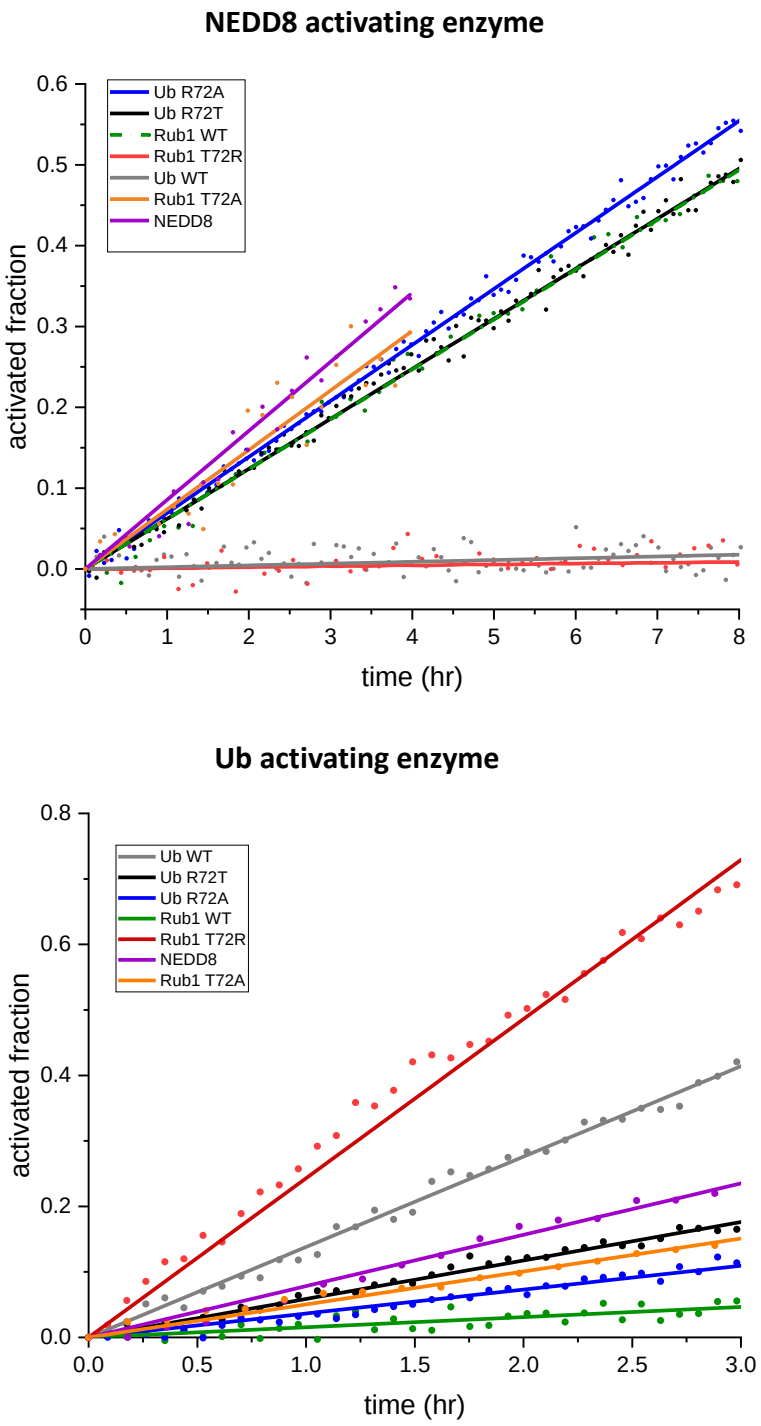


Figure 29. Kinetics of thioesterification reactions with UAE and NAE. The timecourse for thioesterification reactions with 100 µM substrates are shown.

Table 3. Activation rates and calculated K_M for Ub, RUB1, and NEDD8 constructs with UAE and NAE.

Construct	UAE activation rate (% / hr)	UAE K_M (μM)	NAE activation rate (% / hr)	NAE K_M (uM)
Ub 72R	13.8 ± 0.1	4.2 ± 0.5	-	-
RUB1 T72R	24.3 ± 0.3	8 ± 2	-	-
Ub R72T	5.87 ± 0.07	36 ± 2	6.20 ± 0.02	41 ± 3
RUB1 72T	1.6 ± 0.1	205 ± 13	6.22 ± 0.04	37 ± 6
Ub R72A	3.65 ± 0.07	26 ± 3	6.93 ± 0.02	16 ± 3
RUB1 T72A	5.0 ± 0.1	59 ± 4	7.4 ± 0.3	14 ± 1
NEDD8 72A	7.8 ± 0.1	143 ± 11	8.5 ± 0.2	4.1 ± 0.4

3.3 1D measurements of thioesterification reactions

At protein concentrations below ~20 μM the relatively small signal to noise ratio of NMR spectroscopy begins to make signal intensity quantitation more prone to errors and less accurate. To measure the thioesterification reaction with low protein concentrations we proposed that the strong signal for G76, with limited 1H chemical shift overlap, could be measured by 1D ^{1}H experiments. Employing the same fast-repetition ^{1}H – ^{15}N experiments but without discrimination of ^{15}N frequencies yielded 1D ^{1}H spectra with greater signal to noise than conventional 1D or 2D experiments (Figure 30a). The calculated rates at 20 μM were consistent between 2D and 1D experiments and allowed for these measurements to be performed with protein concentrations down to ~5 μM (Table 4). For experiments where Ub was thioesterified by UAE, as the concentration of Ub approached 5 μM the thioesterification rates began to decrease in line with approaching the Michaelis constant. These measurements were

repeated spanning 7 – 45 μM and the concentration-dependent rates were fit to Michaelis-Menten kinetics which is shown in Figure 30b,c.

Table 4. Thioesterification rates of ubiquitin by UAE were non-linear at low concentrations

Concentration	**Initial Velocity**	**Rate equivalents**	**Concentration weighted equivalent**
20 μM (2D)	0.33 ± 0.01	1.03	
20 μM	0.32 ± 0.06	1.00 / 1.0	100%
15 μM	0.42 ± 0.05	1.31 / 1.5	87%
10 μM	0.53 ± 0.08	1.66 / 2.0	83%
5 μM	0.9 ± 0.2	2.69 / 4.0	70%

For the thioesterification of Ub by transthiolation with UAE and Mesna we observed an apparent Km of 4.5 ± 0.7 μM, compared to a Km of ~1 μM observed for cleavage with DTT and recorded by radiolabeled ATP or fluorescent tags[390]. A similar procedure was performed with NEDD8 and Ub and Rub1 72 A, T, and R constructs for both UAE and NAE and the apparent Km values are shown in Tables 2. With NAE the Mesna thioesterification affinities are highly correlated to the residue at position 72. 72R (WT Ub) constructs were completely unreactive, 72T (WT Rub1) have a K_M of ~38 μM, 72A (WT NEDD8) have a K_M of ~15 μM, and NEDD8 itself has a Km below the detection limit. Similarly for UAE the thioesterification affinities are related to the position 72 residue. Ub and RUB1 72R had low K_M values of approximately ~6 μM

matching the significantly quicker thioesterification rates. Ub and RUB1 72A were both shown to have smaller K_M values than 72T constructs, although in both cases the Ub constructs had a greater thioesterification affinity than RUB1.

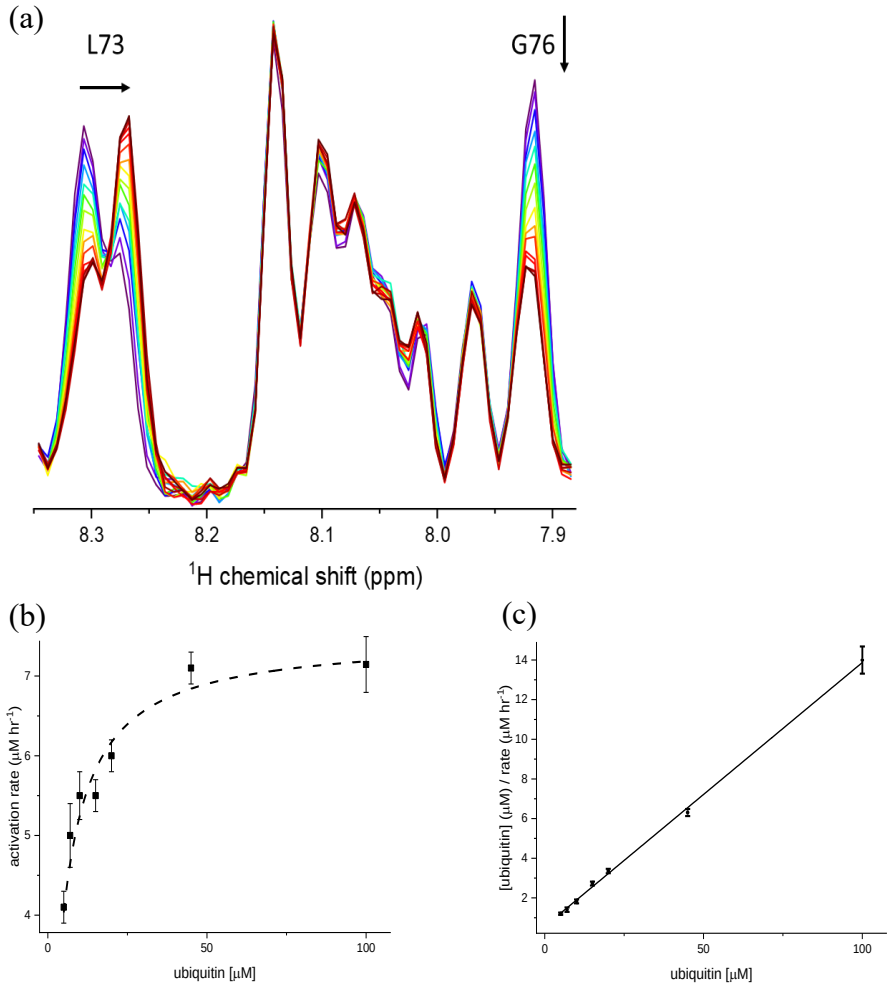


Figure 30. The activation rates were measured at low substrate concentrations by a 1D variants of SOFAST HMQC. (a) A complete reaction timecourse for ubiquitin with UAE is depicted with cool to warm colors. The intensity for the peak containing G76 decreases while L73 shifts upfield. The concentration-dependent reaction rates were used to fit for Michaelis-Menten kinetic parameters by least squares **(b)** or by linearization with a Hanes-Woolf plot **(c)**. The error bars shown in the graph correspond to the standard error of the mean obtained from the linear regression model.

3.4 Conclusion: NMR can be used to directly measure UBL activation rates

In this work we have shown that the rate of UBL thioester formation by either UAE or NAE could be measured by detecting the cleaved UBL product as the UBL with a small-molecule C-terminal thioester. It was demonstrated that the C-terminal Gly of Ub and RUB1 has a strong and isolated signal observed in 2D ^{1}H-^{15}N NMR spectroscopy experiments, and this signal decreases in intensity proportional to the amount of thioesterified product. Using this methodology we have shown that the NAE is highly selective for the Ub/RUB1 C-terminus, and the reaction velocities were modulated mainly by the variable residue at position 72. Regardless of whether Ub or RUB1 was used as the test scaffold, the A or T variants were thioesterified at the same rate with a slight preference shown for the smaller A sidechain. The UAE thioesterification velocities are enigmatic, which is most likely a consequence of the concomitant two-site binding mechanisms. The UBLs which contained the native R as residue 72 were readily thioesterified, with the RUB1 variant surprisingly reacting quicker than Ub. The A and T constructs had similar thioesterification rates otherwise, aside from the poor reactivity of WT RUB1.

Applying a similar methodology while enhancing signal intensity at the expense of resolution, reasonable thioesterification kinetics could be recorded at low protein concentrations which would otherwise be difficult to measure with NMR spectroscopy. The concentration-dependent reaction velocities recorded in this manner were fit to Michaelis-Menten kinetics to determine the Michaelis constants for both E1s. Unsurprisingly the K_M values with NAE are highly correlated to position 72 sidechain and consistent between Ub and RUB1 constructs, similarly to the reaction velocities.

K_M values for the UAE also exhibit a correlation to the residue at position 72, and for Ub and RUB1 these are ordered as R < A < T. In all cases with UAE the Ub scaffold is processed at least 2x more efficiently than RUB1. All together these results point to a mechanism for allowing Ub and RUB1 signaling to coexist despite the significant similarity between the UBLs.

The RUB1 pathway clearly prohibits Ub at the level of E1 activation by designing a cationic residue into the active site pocket which excludes the Ub C-terminus from binding. This binding pocket appears to be relatively unchanged from early eukaryotic (RUB1) to mammalian (NEDD8) E1s as hydrophobic residues in the variable position 72 in the C-terminus are processed with roughly the same efficiency. Under basal conditions the concentrations of Ub and NEDD8 have been reported to be approximately equal at ~27 μM and Ub E1 would be expected to bind and process Ub nearly exclusively based on the respective K_Ms for the UAE.[238] When this concentration of free Ub is decreased significantly by cellular stressors, or in conditions where the ratio of free Ub to RUB1 is otherwise decreased, RUB1 could become a favorable substrate for the UAE.

Some have proposed that this could result in unique signaling events where the RUB1-bound UAE interacts with E2s in novel ways.[237,396] However UAE has been shown to process and transthiolate RUB1 to Ub E2 enzymes.[397] These charged E2 enzymes have been shown to effectively cap growing Ub chains with RUB1 but lack the capability to form additional chains off RUB1.[224] The principal role of polyUb signaling is to mark proteins for proteasomal degradation, which is most efficient by generating K48-linked polyUb chains with a minimum of four Ub units.[21] When

polyUb formation is enhanced due to cellular stressors, and the amount of free Ub decreases as it is polymerized, RUB1 may play a key role in limiting extraneous polyUb chain formation and conserving the dwindling concentration of Ub.

In this work I have demonstrated how an NMR-based assay can be used to follow UBL activation pathways. I have applied this methodology to study the intertwining of Ub and NEDD8 signaling pathways, but it is readily applicable to other UBL systems. Additionally, I only scratched the surface of observing both Ub dimer formation or deconjugation through DUB activity. I hope that this methodology will be used to elucidate further properties of UBL activation and conjugation. Ultimately teasing out these intricacies will lead to a better understanding of the interplay of UBL signaling and how these systems have developed to function together.

3.5 Contributions and acknowledgements

I performed all of the NMR data acquisition, processing, and analysis with advice from Dr. David Fushman. I prepared all the material necessary for this work, aside from bulk Ub E1 used in the Fushman lab.

Chapter 4. Structure and ubiquitin-binding properties of the hDDI1 ULD

4.1 Structure determination of hDDI1 ULD

The ULD of both hDDIs share a 73% sequence identity and 80% sequence similarity, whereas they both share only a 35% percent identity with yDDI1 ULD (Figure 34). We hypothesized that the multiple functions of yDDI1 might be divided between the two hDDI isoforms, and that polyUb binding and substrate cleavage could be conserved in hDDI1. I approached this project by first determining the solution structure of the hDDI1 ULD (hULD) by NMR experiments. I initially assigned the ^{1}H-^{15}N chemical shifts for the backbone amides using a combination of ^{1}H, ^{13}C, and ^{15}N triple resonance experiments which correlate the ^{13}C Cα and ^{13}C carbonyl (CO) chemical shifts from the i and i-1 residues. As prolines lack a protonated amide they create natural breaks in the sequence and could result in an ambiguous assignment. The residues were unambiguously assigned using a combination of total correlated spectroscopy (TOCSY), which link chemical shifts through bonds, and nuclear Overhauser effect (NOESY) experiments which link the chemical shifts of nuclei that are in spatial proximity. Unambiguous sidechain assignments were primarily achieved through H(CC)CONH and (H)CCCONH TOCSY experiments which report most of the ^{1}H and ^{13}C resonances of the i-1 sidechain, respectively, and are characteristic of the sidechain type.

To generate the solution structures of proteins by Xplor-NIH[398] numerous distance restraints between ^{1}H resonances are required, necessitating a complete sidechain assignment. Some of the ^{1}H and ^{13}C chemical shifts are undetectable from the amide linked TOCSY experiments due to the distance from the amide bond and the

efficiency of the TOCSY transfers. H(C)CH and (H)CCH TOCSY experiments allowed for a comprehensive assignment of aliphatic sidechain resonances. Aromatic ^{1}H resonances, which overlap with amide protons, were assigned by ^{1}H-^{1}H TOCSY/NOESY experiments after the ULD was lyophilized and then solvated in the same buffer where the H_2O was replaced with D_2O. All in all I was able to assign 86% of the chemical shifts.

After 24 hours in this D_2O buffer, the remaining ^{1}H-^{15}N crosspeaks were from amides which are locked into strong hydrogen bonds or are extensively buried and cannot undergo exchange with the bulk D2O; those that could be preliminarily matched to hydrogen bonds (H-bonds) were used as restraints for structure generation and are shown in Table 5. ^{15}N Residual dipolar couplings (RDCs) were determined by measuring the scalar coupling constants of the amide ^{1}H-^{15}N chemical shifts in isotropic and anisotropic media, the latter consisting of a liquid-crystalline matrix of PEG-hexanol.[399–401]

From the complete signal assignment distance restraints were generated from ^{15}N-^{1}H- and ^{13}C-^{1}H-detected NOESY experiments. The backbone ^{1}H, ^{15}N, $^{13}C\alpha$, and ^{13}CO chemical shifts were used to generate dihedral angle restraints using the program TALOS-N.[402] Distance and dihedral angle restraints were used to generate a preliminary ensemble of structures using the ARIA program[403] interfacing with Xplor-NIH. The lowest energy structure and refined distance restraints were subsequently employed in a second ARIA run, which included the addition of H-bonding and RDC restraints (Figure 31), resulting in an ensemble of 30 structures with a Cα RMSD of 0.37 Å and shown in Figure 32.

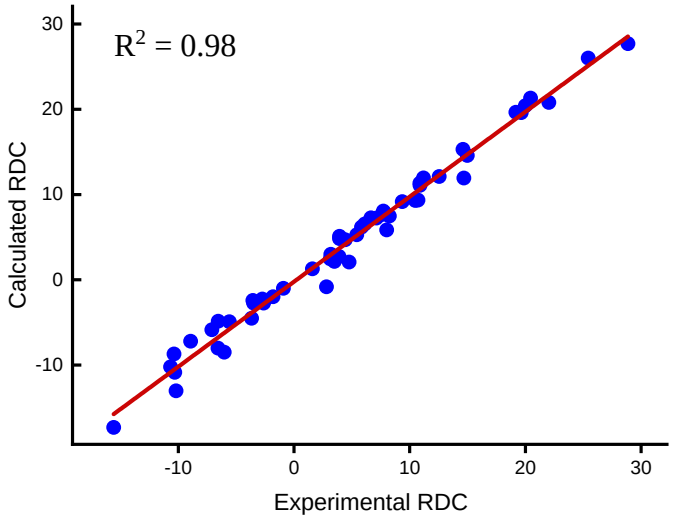


Figure 31. Comparison of experimental and calculated RDCs.

The final distance and H-bonding restraints are included in Table 5. In accordance with the sequence similarities between the ULD of hDDIs (80% sequence similarity) the structures of the ULDs overlay well with an all-atom RMSD of 2.06 Å. The hDDI1 ULD fold resembles Ub and the yDDI1 ULD as presented in Figure 33.

Table 5. Final NOE distance restraints and H-bonds. Hydrogen bonds are listed as donor amide NH to acceptor carbonyl (NH:O).

Total NOEs	1837
Intra-residue	666
Sequential	498
Short (2 - 3 residues)	258
Medium (4 - 5 residues)	74
Long (> 5 residues)	341
Hydrogen-bonds	36:32, 37:33, 39:35, 38:34, 40:36, 41:37, 42:38

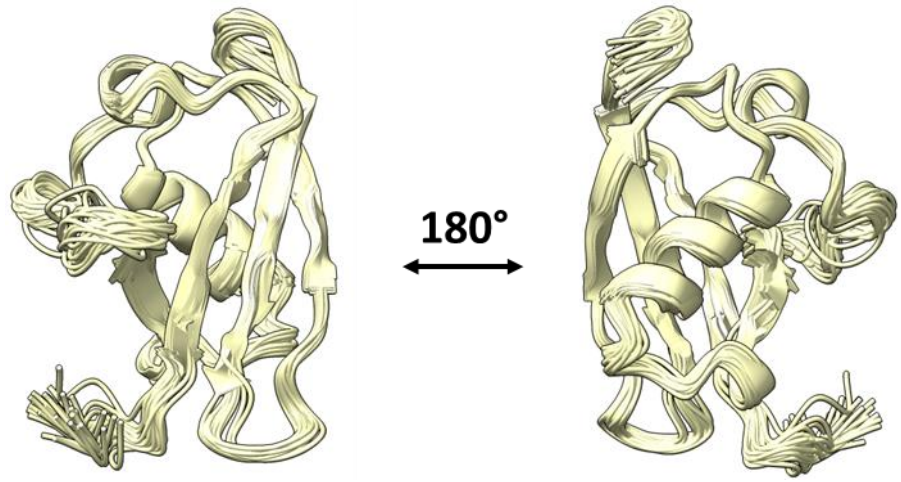


Figure 32. The solution-state ensemble of the hDDI1 ULD.

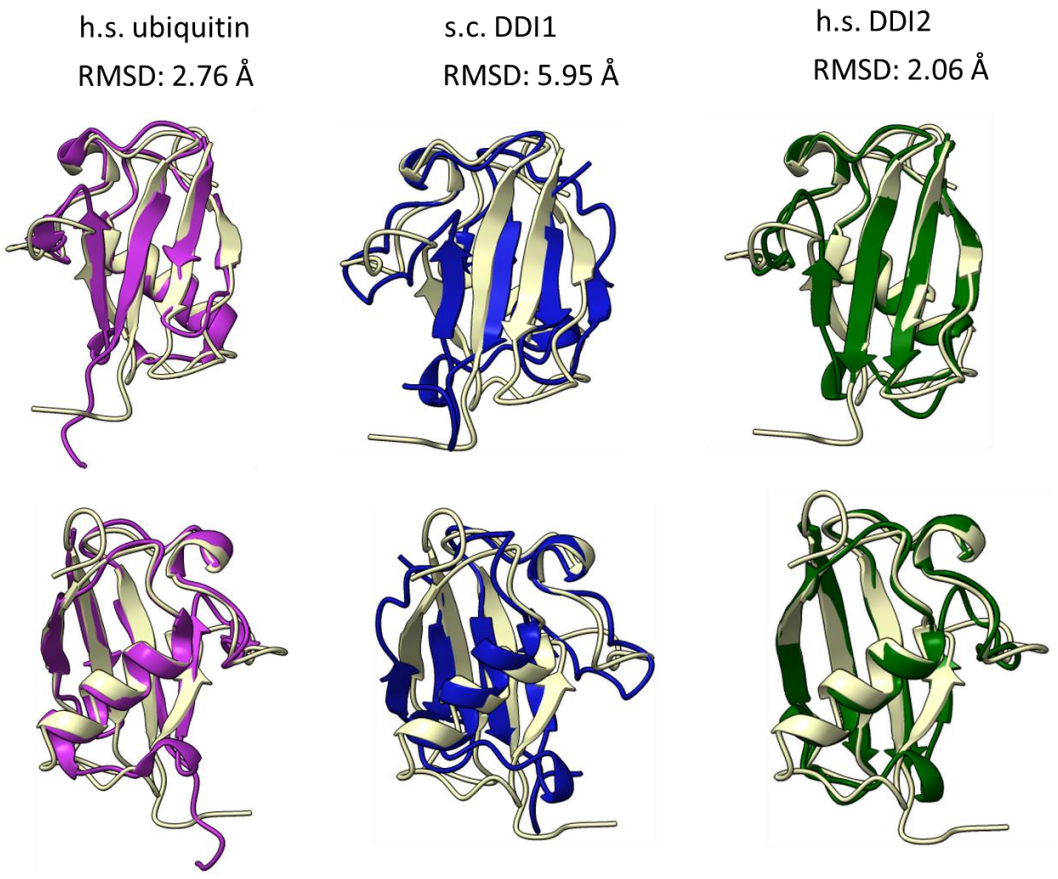


Figure 33. Structure comparison of h.s. DDI1 ULD to ubiquitin (1D3Z, purple), s.c. DDI1 ULD (2MRP, blue), and h.s. DDI2 ULD (27ND, green). The indicated RMSD are from hDDI1 ULD to the overlayed protein excluding the tail residues.

4.2 Ubiquitin and polyubiquitin binding properties of the hDDI1 ULD

To evaluate the ULD's capacity to bind Ub, we conducted NMR titrations where either the hULD or Ub was ^{15}N-labeled, with the unlabeled complementary component added incrementally. Concentration-dependent chemical shift perturbations (CSPs) were used to determine the ratio of bound to unbound species and to calculate the dissociation constant (K_D). Like yDDI1, hULD exhibited a relatively weak binding affinity for Ub with a K_D of 27 ± 6 μM compared to 45 ± 6 μM (Figure 35) and Table 6). Furthermore, we explored whether the linkage-specific chain geometries of K48- or K63-Ub2 affected the binding affinity using a similar approach. The extended topology of K63-Ub2 did not enhance binding capability, and the observed K_D of 17 ± 4 μM represents the two independent Ub binding sites exerting only cumulative binding affinity . In contrast, titrations with K48-Ub2 resulted in significant signal attenuation and larger CSPs, suggesting a slower rate of exchange and a distinct binding mode (Figure 35 and Table 6).

As anticipated, the significant CSPs observed for monoUb upon binding to ULD correspond to the hydrophobic binding patch consisting of L8, I44, and V70, (Figure 35) akin to findings reported with yDDI1. The CSPs for the hULD are focused around beta-strands 1, 2, and 5, forming a contiguous binding interface (Figures 34 and 35, Table 6).

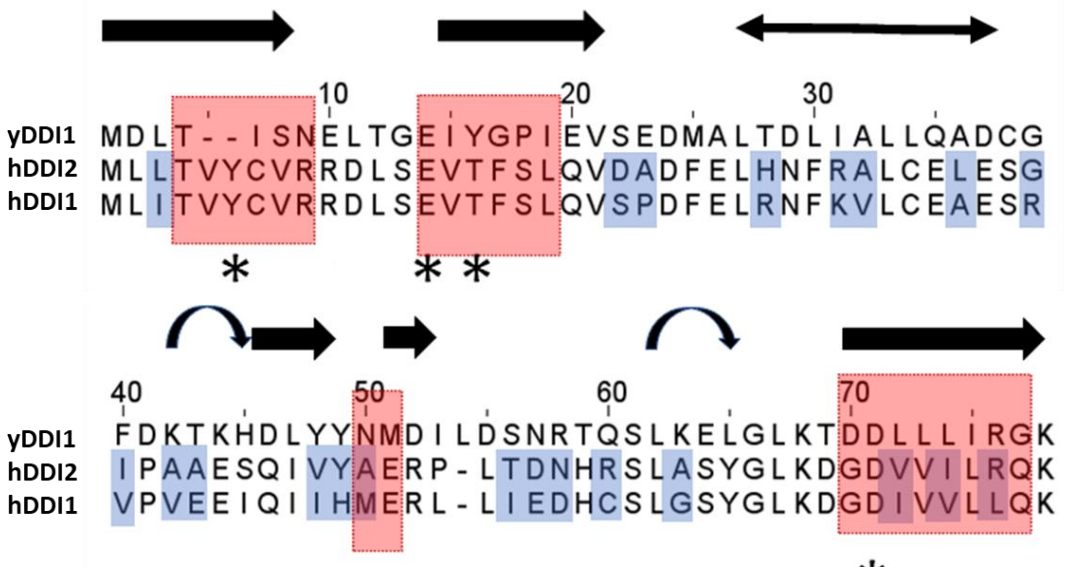


Figure 34. Sequence alignment of hDDIs and yDDI. Secondary structure elements are represented above the sequences. Thick arrows are β-strands, double-sided arrows are α-helix, and arching arrows are turns. The areas with significant CSPs upon binding to Ub are highlighted in red, whereas locations where hDDI1 differs from hDDI2 is highlighted in blue. Asterisks are placed on residues with the largest CSPs.

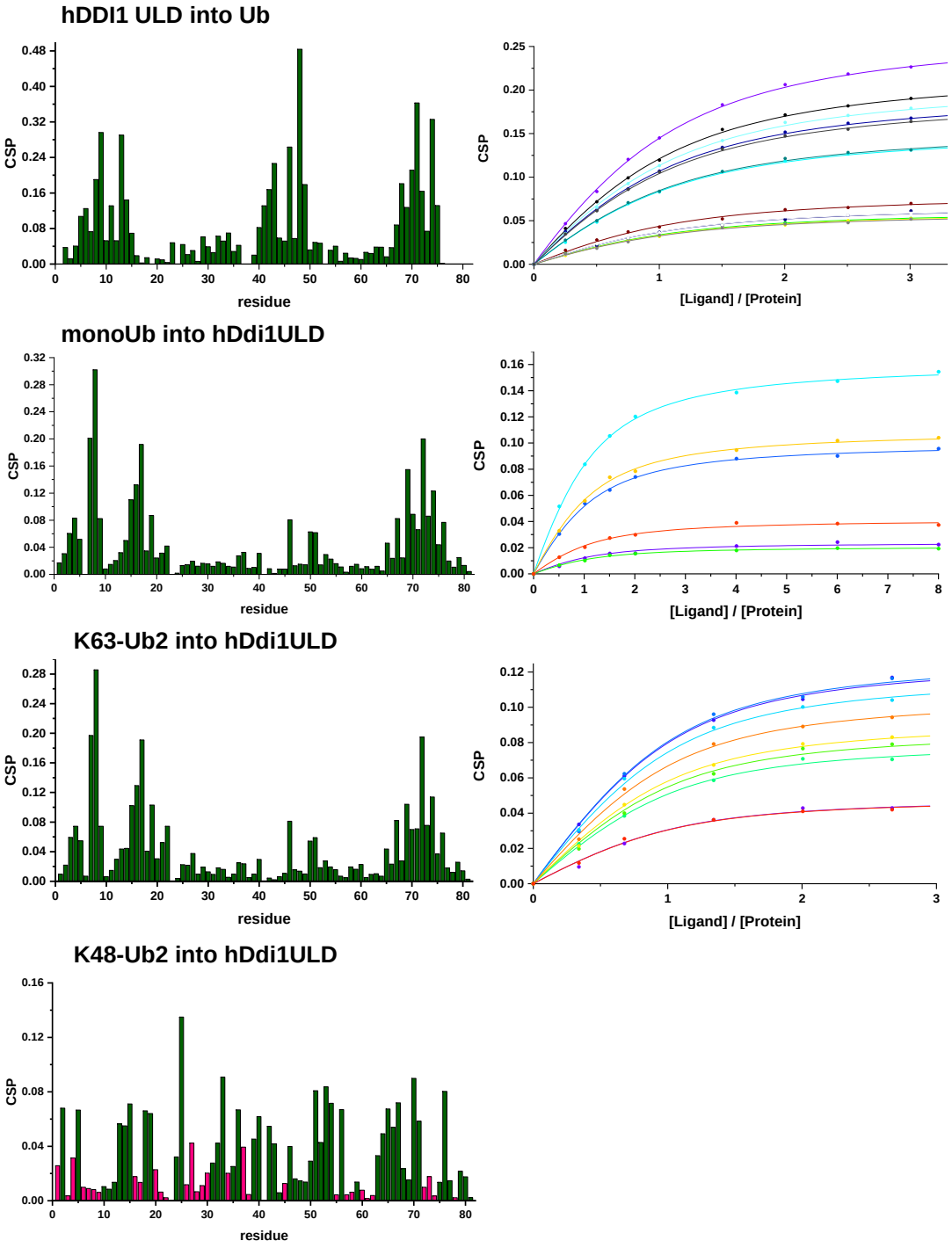


Figure 35. CSPs at titration endpoints and dissociation constant fitting. CSPs are shown in green, and in the case of intermediate exchange (K48 Ub2) the residues are shown in fuchsia with the CSP before they exchanged. The CSPs were fit to the ratio of ligand to protein concentration through the course of the titration to calculate the dissociation constants.

Residues exhibiting the largest CSPs in both Ub and the hULD were designated as "active residues" (Table 7) and input into the HADDOCK webserver[404] to generate a predicted binding mode by docking the solution structures of hULD and Ub as shown in Figure 36a. The best cluster, as designated by thermodynamic minimization, consisted of 23 structures with a RMSD of 1.3 ± 1.0 Å, and the lowest energy structure was used for further analysis. The resulting model suggests that, in addition to binding energies driven by the hydrophobic effect (Figure 36b), several other interactions contribute to the orientation of the bound state.

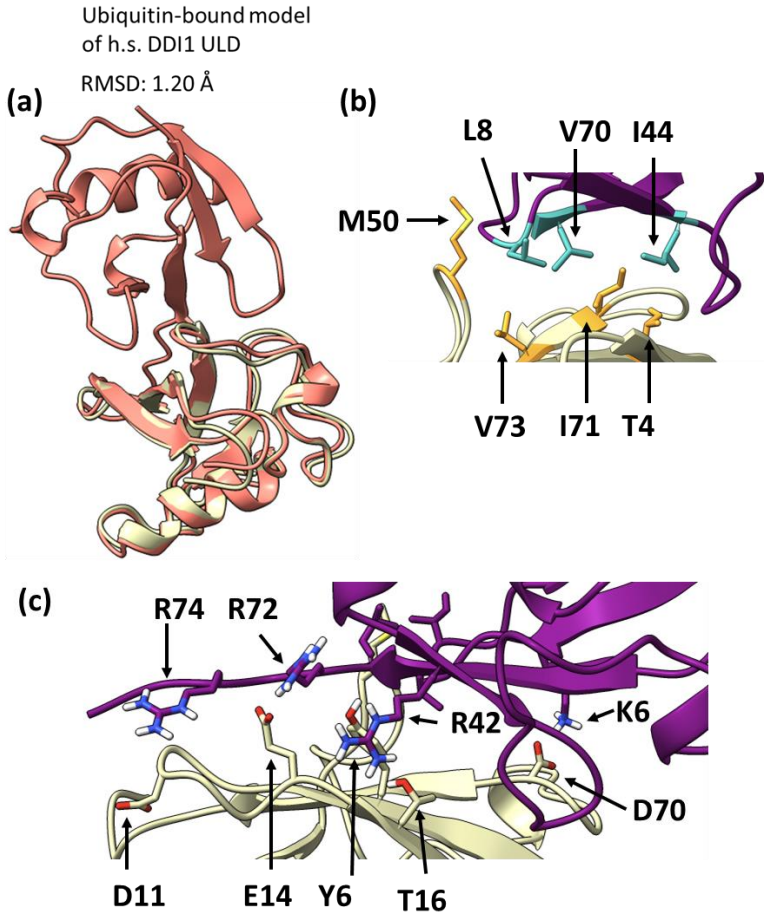


Figure 36. HADDOCK model of Ub binding to h.s. DDI1 ULD. (a) overlay of h.s. DDI1 ULD (white) overlayed on the ULD domain from the HADDOCK model (pink). (b) Contacts from the ubiquitin hydrophobic patch to hULD. (c) Acidic groups on the β-sheet of hULD complement cationic groups of ubiquitin.

These include charge-charge interactions from R72/R74 of Ub and D11/E14 in the hULD, a salt bridge from Ub K6 to hULD D70, and hydrogen bonding between Ub R42 and Y6/T16 in the hULD help to orient the bound state (Figure 36c). The Ub-bound state of hULD aligns well with yULD and occurs by the same binding mode and intermolecular contacts and are compared in Figure 37.

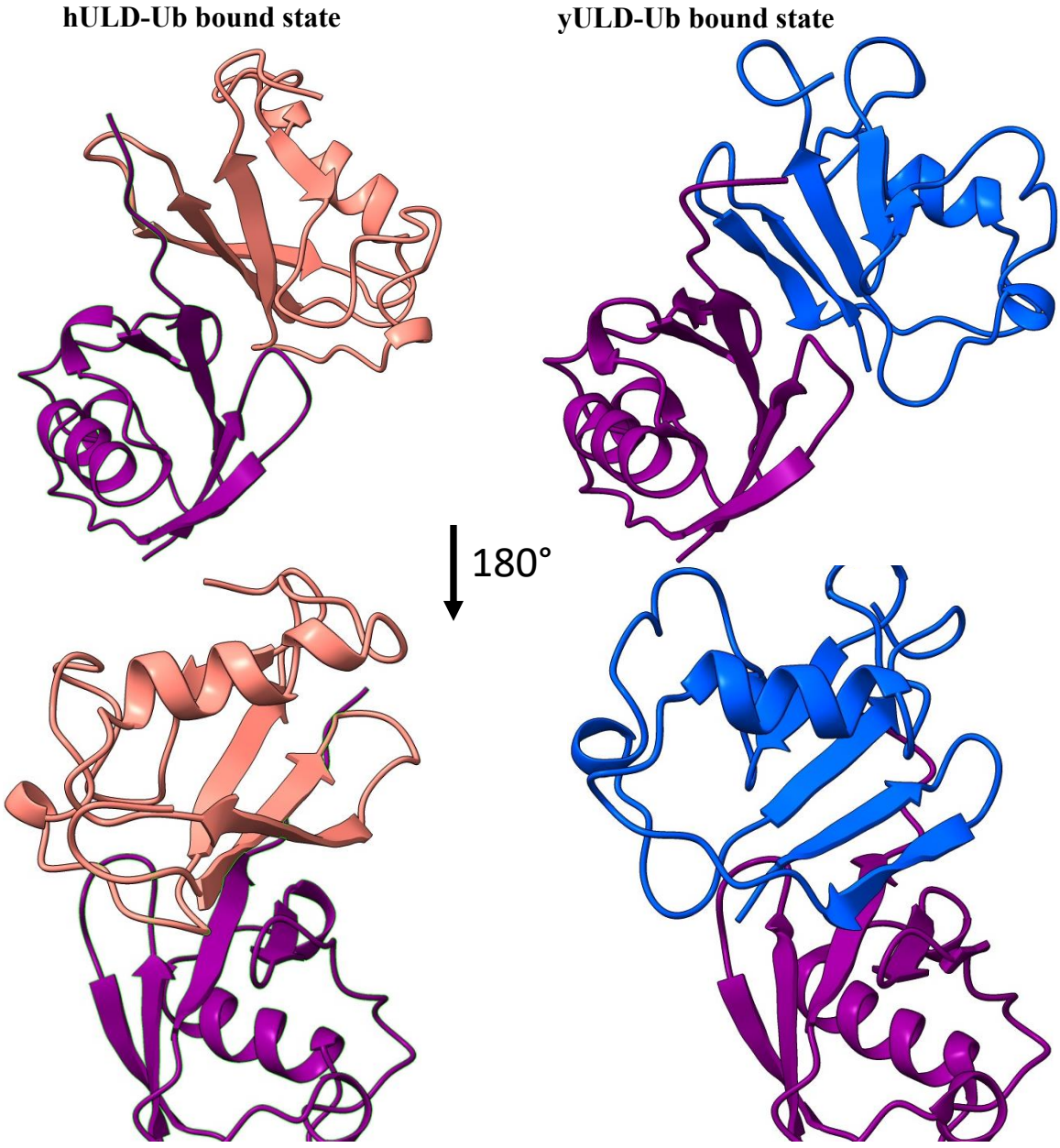


Figure 37. Comparison of yULD and hULD HADDOCK structures with ubiquitin. HADDOCK structures are shown for hDDI1 ULD (pink) and yDDI1 ULD (blue) bound to ubiquitin (purple).

Table 6. Titration data used to calculate dissociation constants. * Due to significant signal attenuations immediately on adding the K48 Ub2 titrant the K_D could not be calculated. The K_D for yeast DDI1 ULD was reported previously. ** K_D values were calculated using the in house KDFit program (7.3.2) and the errors are derived from the standard deviation within the residue-specific K_Ds.

	Residues	K_D (μM)**
^{15}N Ub		
hDDI1 ULD	8,10,11,12,29,34,41,42,43,44,45,49,72	40 ± 3
yDDI1 ULD[168]		45 ± 7
^{15}N hDDI1 ULD		
monoUb	20,26,55,71,72,79	27 ± 6
K48-Ub2	*	*
K63-Ub2	4,8,9,14,19,26,55,81,84	17 ± 4
^{15}N hDDI1 ULD VL_AK		
monoUb	43, 45, 55, 57, 81, 84, 85	40 ± 4
K48 Ub2	10, 33, 64, 71, 73, 84, 85	19 ± 5

Table 7. Active residues selected for HADDOCK.

Ambiguous CSP-derived interaction restraints	
Ubiquitin	8, 9, 11 ,14 ,42 ,46 ,49 ,68 ,70 ,71 ,72 ,74 ,75
hDDI1 ULD	7, 10, 12, 16, 33, 34, 36, 37, 38, 39, 40, 41, 44, 45, 47, 48, 50, 51, 52, 59, 61, 64, 66, 73, 76

The marked difference in reported Ub binding abilities between hDDIs is not easily discernible from the bound state model, electrostatic surface map, or a primary structure comparison (Figures 34 and 38). Most distinctions between hDDI ULDs are localized on the backside of protein, surrounding the alpha helix, and are entirely separate from the primary binding site. Most non-equivalent residues situated near those exhibiting substantial CSPs in hDDI1 ULD are hydrophobic to hydrophobic mutations in the hDDI2 analogue as highlighted in Figure 34. In the case of L76, present as R76 in hDDI2, the equivalent R74 is conserved in yDDI1. Although minor differences in hydrophobic contacts might collectively impair binding, the hDDIs exhibit notable divergence at residues 49 and 50. Specifically, M50 is conserved with scDDI1 (M48), and despite only having hydrophobic contacts, it may play a crucial role in determining Ub binding in hDDI1 and scDDI1.

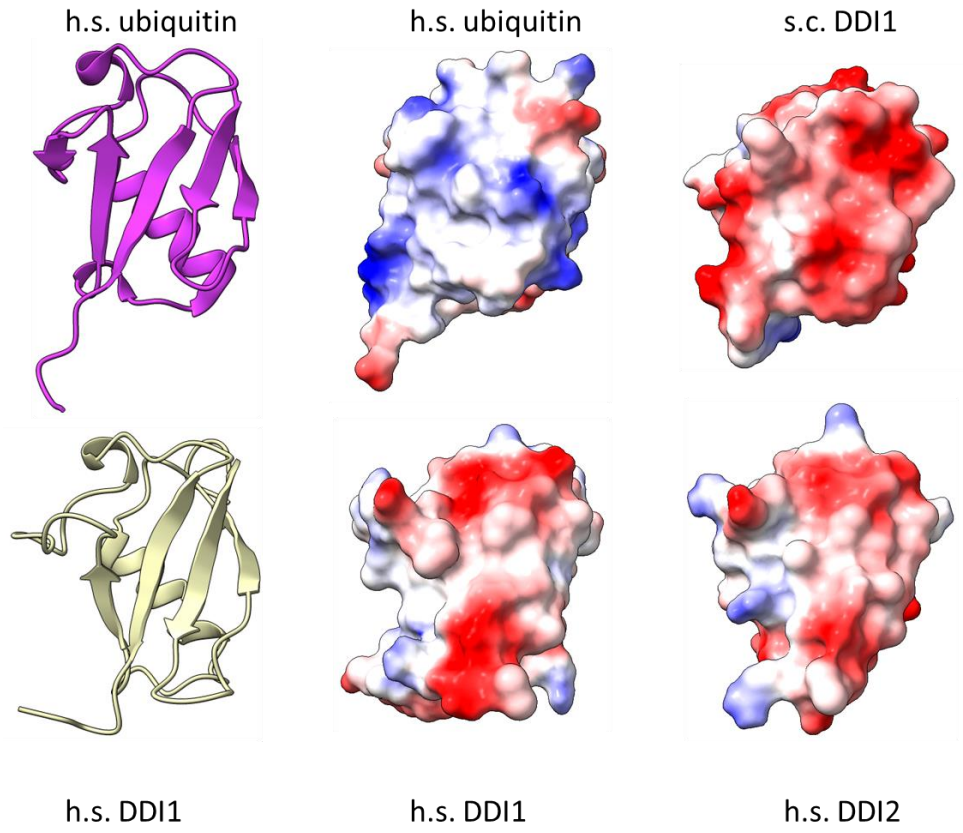


Figure 38. Surface electrostatic map of Ub and DDI1 ULDs. The Coulombic electrostatic potential is shown on a sliding scale where red represents a negative potential and blue a positive potential.

4.3 DDI1 contains a second Ub binding site

In a previous study by Dr. Urszula Nowicka from the Fushman lab, it was demonstrated that K48-linked Ub dimers (K48-Ub2) bind to the s.c.DDI1 ULD (scULD) with a significantly higher affinity than K63-Ub2 or monoUb. Upon the addition of K48-Ub2 to ^{15}N-labeled yULD, large CSPs were observed throughout the structure (Figure 35). Many residues also entered an intermediate exchange regime due to tighter binding. The extensive CSPs and stronger binding suggest a distinctive binding mode for K48-Ub2, which also involves interaction with the ULD α-helix.

Both yULD and hULD possess a similar arrangement of hydrophobic and acidic sidechains at the solvent-exposed portion of the helix. This arrangement features a hydrophobic surface consisting of V32, L33, and A36, located near E37 (A30, L31, A34 and D35 in yULD) on the C-terminus of the helix (Figure 39). In yULD, it was demonstrated that mutating L31 to a Lys, which corresponds to the residue at this position in Ub (hDDI1 ULD VL_AK), effectively abolished Ub binding to this site. Figure 40 shows that the CSPs for monoUb and K48 Ub2 show an identical binding mode. Similarly, introducing a comparable mutation into hULD yielded the same outcome, and K48-Ub2 is shown to bind to this mutant at twice the affinity of monoUb and recapitulate the binding mode.

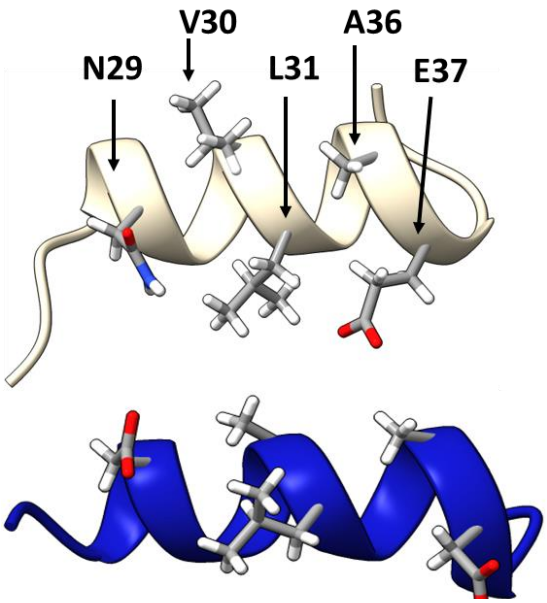


Figure 39. A secondary binding site specific for K48-linked polyUb is conserved from yDDI1 to hDDI1 ULD. V30, L31, and A36 are highlighted for hDDI1 ULD (White), and the synonymous residues are drawn on yDDI1 (blue).

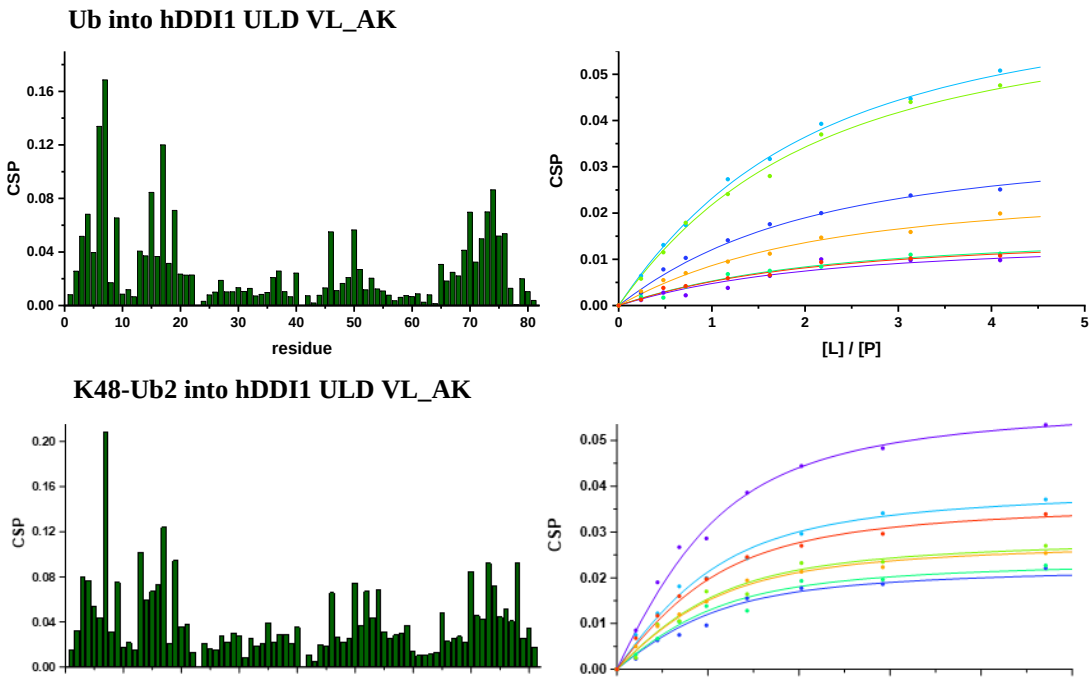


Figure 40. CSPs at the titration endpoints and dissociation constant fitting. CSPs are shown in green for the indicated ubiquitin mutants interacting with the hDDI1 ULD VL_AK mutant. The CSPs were fit to a ratio of ligand to protein concentration through the course of the titration to calculate the dissociation constants.

4.4 Conclusions: hDDI1, in contrast to hDDI2, is a ubiquitin binding protein

Our results demonstrate that the ULD of hDDI1 retains the Ub binding functionality observed in yDDI1 and shares the ability to interact with the UPS. A recent report revealed that yDDI1 acts as a Ub-dependent protease, apparently requiring long polyUb chains indicative of malfunctioning or overloaded proteasomes. This finding challenged the prevailing theory that yDDI1 only serves as a proteasomal shuttle in the UPS.[185] The proteolytic activity provides an explanation for the highly conserved RVP domain found in all DDIs, as well as the significance of the functional protease domain in the protein degradation phenotype associated with DDIs.

In all DDIs, the C-terminal domains, previously thought to interact with Ub through motifs analogous to UBA and UIM domains, have been determined to only

have weak binding affinities for Ub[181] or are superseded by a stronger Ub affinity of the ULD.[168] In yDDI1, the binding affinities to polyUb chains were unaffected by deletion of the UBA domain, slightly reduced by removal of the HDD domain, but decreased by four orders of magnitude when the ULD was truncated[185]. However, enzymatic activity was only partially reduced by removal of the ULD.

The ULD and C-terminal portion of hDDI2 was shown to be ineffective at binding to Ub.[181] Despite this, DDI2 knockout cells exhibit a hyperpolyubiquitination phenotype, and it has been demonstrated that DDI2 can proteolytically process Nrf1. RAD23 was shown to enhance the rate of DDI2 cleavage of polyUb substrates, suggesting that DDI2, and possibly DDI1, can interact with the UPS without direct connections with Ub.[172] Our findings reveal that the ULD of hDDI1 retains the ability to associate with Ub and binds tightly to K48-linked polyUb. This preference for K48 polyUb chains would likely increase DDI1 localization to K48-polyUb chains, resulting in enhanced DDI1-induced proteolysis rates of proteins marked for degradation with K48-linked polyUb chains which would aid in protein turnover during cell cycle transitions or when proteasomes are under stress. However, limited information is currently available on the interplay between the hDDI1 and hDDI2 paralogues and how the tempered Ub-binding ability of hDDI2 influences DNA repair processes and interactions with the UPS.

The observation that the ULD of hDDI1 can interact with Ub prompts further research into understanding how the full length hDDI1 interacts with polyUb chains. Specifically, the study of DDI proteins in higher organisms typically discounts DDI1 as simply a gene duplication. However my data suggests that hDDI1 shares similar Ub

binding properties with yDDI1, and we are only just beginning to understand the full extent of the yDDI1 endopeptidic activity upon binding to long polyUb chains. While the individual domains of hDDI2 have been analyzed in detail, my work suggests that it is erroneous to assume that their functionality is completely conserved in hDDI1. Further study is necessary to form the context of DDI1 activity, especially in concert with the functionalities of DDI2.

4.5 Contributions and acknowledgements

This project emerged through a joint effort of Dr. Apurva Chaturvedi, Lillian Hallmark, and me. Much of the original protein work and NMR titrations were performed by Dr. Chaturvedi, with assistance from Ms. Hallmark. I performed all of the experiments involved with structure elucidation and all of the work with the VL_AK mutant of hULD.

The original protein construct was designed by Dr. Michal Chojnacki from the lab of Dr. Michael Glickman.

Chapter 5. Identification of chain-specific ubiquitin-binding domains by photo-crosslinking

Chapter 5 was adapted from the following manuscript:

Braxton, C. N., Quartner, E., Pawloski, W., Fushman, D. & Cropp, T. A. Ubiquitin Chains Bearing Genetically Encoded Photo-Cross-Linkers Enable Efficient Covalent Capture of (Poly)ubiquitin-Binding Domains. *Biochemistry* **58**, 883–886 (2019)

5.1 Introduction to photo-crosslinking in ubiquitin signaling

UBDs for monoUb and polyUb typically exhibit moderate to low affinity binding, ranging from ca. 3 μM to 2 mM. To enhance the understanding of specific polyUb recognition by UBDs and capture low-affinity binding, photo-activatable crosslinkers have been incorporated into polyUb chains. This has classically been achieved with photo-leucine containing Ub monomers which were produced using linear total chemical synthesis.[405] Total chemical synthesis has also been used to incorporate a diazirine photo-crosslinker into K48-linked and K63-linked Ub2 chains.[406] As an alternative strategy, genetically encoded crosslinking amino acids *p*-benzoylphenylalanine (BPA) and *p*-azidophenylalanine (AzF) have been incorporated into SUMO.[407] Upon activation with low energy UV light ca. 330-365 nm these functional groups form a 1,2-diradical or nitrene, respectively. The 1,2-diradical specifically forms crosslinks with unsaturated C-H bonds[408–410], whereas nitrenes can react with a variety of chemical groups[411]. These studies used photo-crosslinkers to covalently bind Ub to known UBDs, peptide sequences, and proteins from cellular lysates.

To investigate whether genetically encoded crosslinkers are a viable strategy for capturing weak or transient polyUb-target interactions, we chose to work with the UBA2 domain of hHR23a[155,412] and the tandem UIM (tUIM) of human Rap80[135] which have been shown to have binding specificity for K48- or K63-linked Ub chains, respectively. We chose to incorporate the chemically stable *p*-benzoyl-$_L$-phenylalanine (BPA) photo-crosslinker into Ub at selected sites using bacterial expression (Figure 41a). Dr. Courtney Braxton from the Cropp lab performed the genetic encoding of BPA using an orthogonal suppressor tyrosyl tRNA ($tRNA^{Tyr}$) and *Methanococcus jannaschii* tyrosyl-tRNA synthetase (MjTyRS) pair.[410] The codons corresponding to the desired positions in Ub were mutated to the amber stop codon, TAG, for BPA incorporation. Positions for mutation to BPA were selected based on the three-dimensional structure of Ub and the proximity to the hydrophobic patch residues.[413] These residues included L8, T9, G10, K11, T12, S20, R42, A46, K48, Q49, E51, D52, H68, and V70. Residues S20 and D52 were chosen as negative controls as they are not located close to the potential binding interface, and therefore are positions in which crosslinking should not occur (Figure 41b).

5.2 Mono- and polyubiquitin crosslinking studies

To identify positions that undergo crosslinking, we tested monomeric Ub variants containing single-site BPA mutations with UBA2. The proteins were mixed on ice and half of the sample reaction subjected to irradiation with a stationary UV lamp (365 nm) for 30 min. Samples were analyzed initially by Coomassie-stained SDS-PAGE, however, no significant crosslinking was observed. The binding affinity of UBA2 for mono-Ub is weak, with a K_D of 400-500 μM,[87,414,415] and we therefore

thought there may be crosslinking but at yields below the detectable limits of a Coomassie stain. These samples were further analyzed using silver-stained gels which revealed potential crosslinked proteins formed, albeit at very low abundance (<1%).

Interestingly, the band intensity of the crosslinked protein varied depending on the site of BPA incorporation, with the strongest band intensity occurring in the instances where the BPA residue was located at positions 8, 9, 42, 48, 49, or 52 (Figure 42a). Similar observations were made for crosslinking of Ub and the UBA domain of human ubiquilin-1 protein (UBAy) and are shown in Figure 42b.

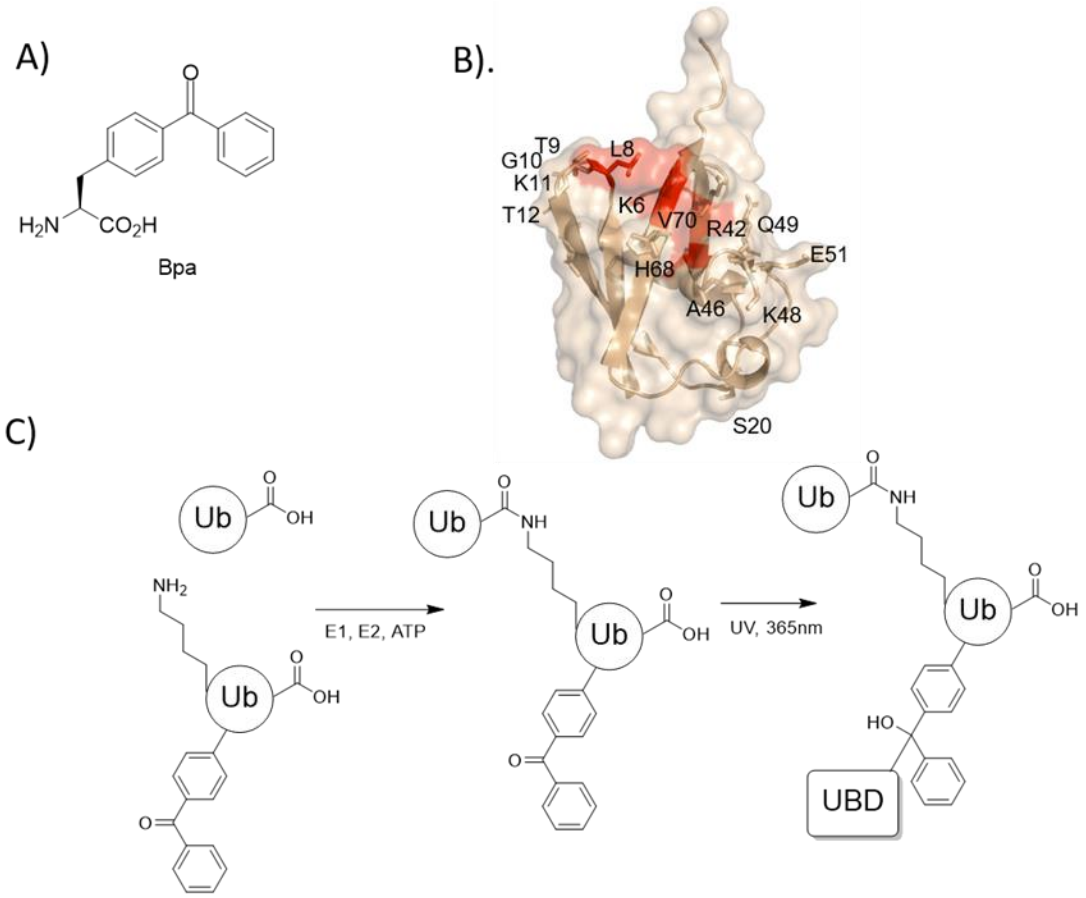


Figure 41. Photo-crosslinking methodology and structures. (a) Structure of p-benzoyl-L-phenylalanine (Bpa). (B) Structure of Ub showing hydrophobic patch (red) and sites chosen for Bpa incorporation. (C) Schematic of the enzymatic assembly of Ub2 with the photo-cross-linker, Bpa in the proximal Ub unit, and capture of a UBD upon photoactivation.

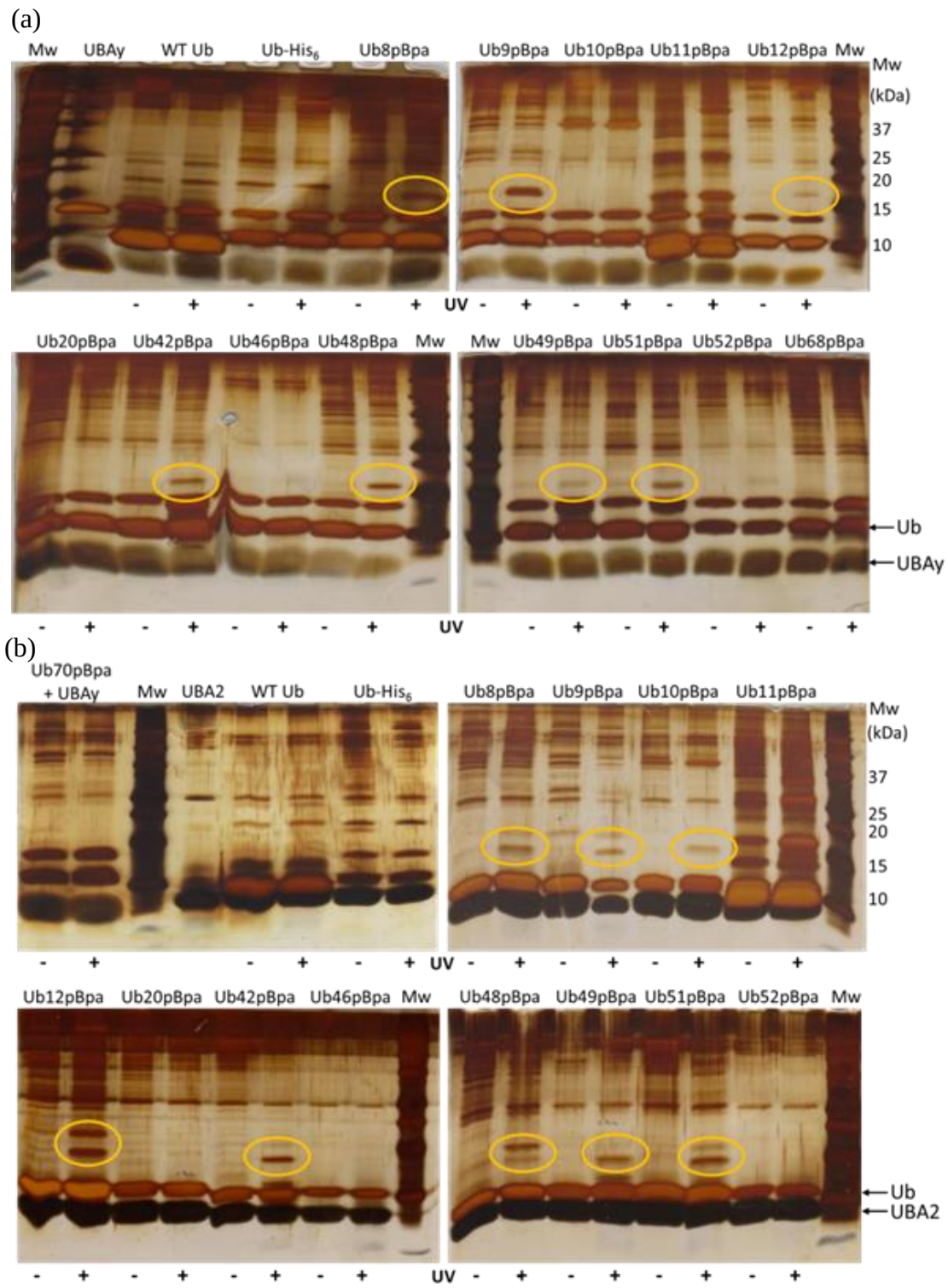


Figure 42. Photo-crosslinking of BPA-**containing monoUb variants. (a)** photo-crosslinking reactions with the UBA domain of human ubiquilin-1 (UBAy)**. (b)** Photo-crosslinking reactions with UBAy (lanes 1/2) and the UBA2 domain of hHr23A (all other lanes). The bands corresponding to crosslinked products are circled (orange).

These results indicate that crosslinking efficiency depends on the site of BPA incorporation and suggest that the above-mentioned sites are the most reactive positions for capturing transient inter-molecular interactions. We decided to investigate covalent capture further by preparing Ub2 chains that contain a BPA at position 9 in the proximal Ub unit. Of note, all the Ub variants we tested were efficiently processed into Ub2 chains, which indicates that a single BPA mutation does not affect the substrate recognition characteristics by E2 enzymes. After preparation, we tested the ability of these chains to cross-link with UBA2. UBA2 has a greater affinity for K48-linked Ub2 chains (K_D = 8 μM) than for K63-linked Ub2 (K_D ~ 200 μM) or monoUb (K_D ~ 400-500 μM). Upon analysis of a Coomassie stained SDS-PAGE gel, a cross-linked band was observed for K48-linked dimers bearing a BPA mutation at position 9, while no cross-linking was observed for Ub2 linked via K63 (Figure 43a,b). These cross-linking results agree with *in vitro* data on UBA2's Ub chains binding preferences. Importantly, we detected no self-crosslinking in reactions that contained Ub2 alone (Figure 44a), which supports the notion that crosslinking will only occur for proteins interacting with Ub near the BPA residue. In a complementary experiment, Rap80 tUIM, which was shown to have binding preference for K63-linked Ub2 chains was used as the binding partner. Rap80 has a seven-fold difference in affinity for the two Ub2 linkages, with a K_D = 22 μM for K63-Ub2, and K_D = 157 μM for K48-Ub2 (Figure 43c,d).[135]

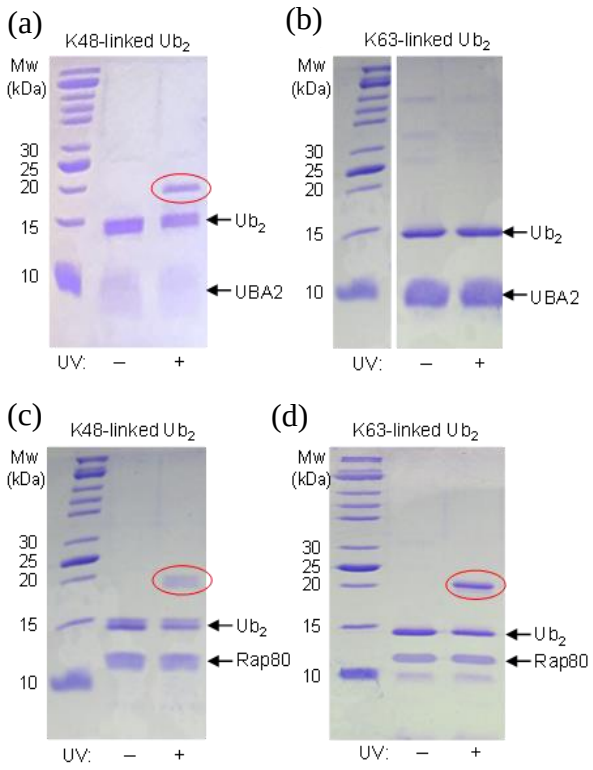


Figure 44. SDS-PAGE analysis of photo-crosslinking of Ub2 **chains containing K48 and K63 linkages when irradiated at 365 nm.** a) K48-linked Ub2 and b) K63-linked Ub2 with UBA2 domain. c) K48-linked Ub2 and d) K63-linked Ub2 with Rap80 tUIM. BPA was inserted in the proximal Ub at position 9. The gels were Coomassie stained; the bands corresponding to crosslinked product are circled

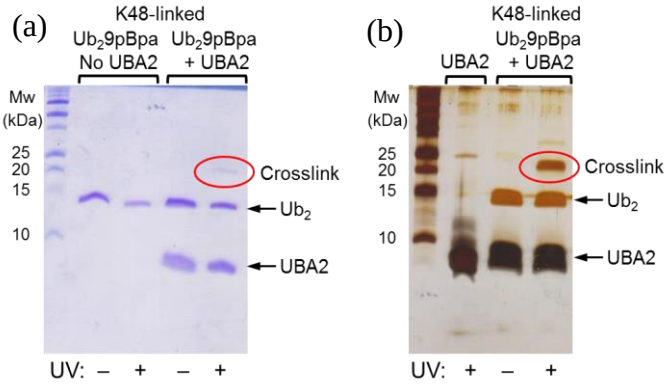


Figure 43. SDS-PAGE gels of photo-crosslinking reaction of K48-linked $Ub_2$9pBPA with UBA2. (a) $Ub_2$9p BPA does not crosslink to itself. (b) The observed crosslinking band is only observed with BPA-containing Ub. The bands corresponding to crosslinked products are circled (orange)

An analogous experiment using the same chains revealed a cross-linked band corresponding to Rap80-Ub2. In contrast to UBA2, cross-linking was observed for both K48- and K63-Ub2. We performed similar experiments using Ub2 chains with BPA inserted in the proximal Ub unit at positions 11, 46, 48, or 49, all of which are summarized in Appendix 1.3. As in the case of monoUb, BPA at positions 9 and 49 were most reactive.

The Ub2 experiments described above all correspond to capturing interactions when the reactive moiety is in the proximal Ub unit in the chain. To confirm that UBD-capturing capability is not limited to the proximal Ub, we assembled Ub dimers bearing BPA in the other, distal Ub unit. For this we generated Ub mutants that contain either K48R or K63R chain-terminating mutations in addition to a BPA mutation. Interestingly, despite the BPA at position 9 being in a different Ub unit/orientation relative to the binding partner, we see evidence of crosslinking (~5%) of both K48- and K63-linked Ub dimers to both binding partners, indicating again, that this position is very reactive (Figure 45).

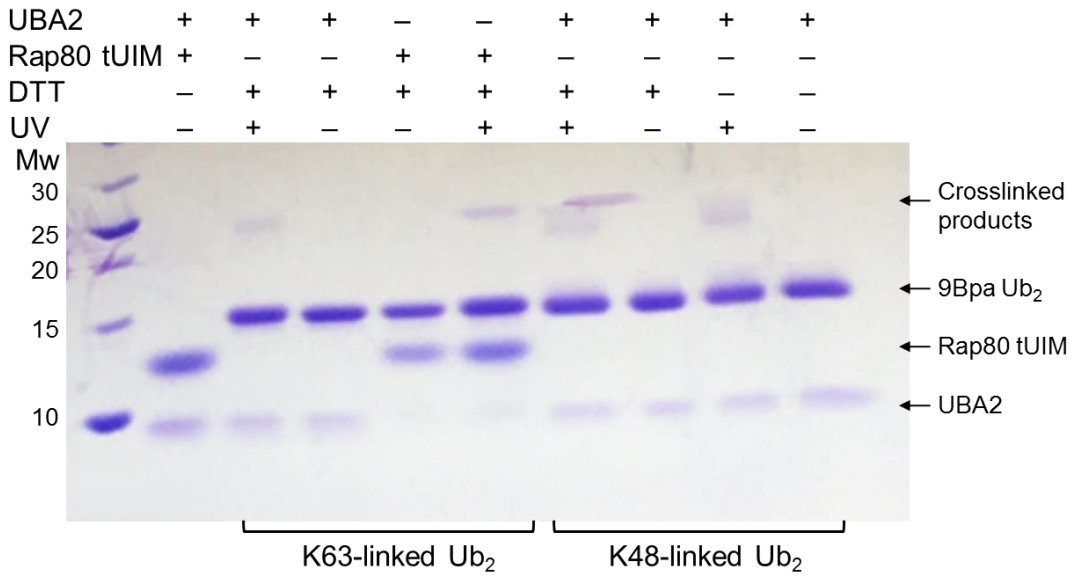


Figure 45. K48- and K63-Ub2 with 9Bpa in distal Ub photo-crosslinking reactions with hHR23A UBA2 domain and Rap80 tUIM.

Finally, to test if BPA-containing Ub2s can capture significantly larger binding partners, we used a 252 amino acid fragment (residues 391-642) of the Rpn1 ($Rpn1^{391-642}$) subunit of the proteasome which recently has been shown to serve as a proteasomal receptor for polyUb.[416] As is evident from Figure 46, both K48- and K63-linked Ub2s having BPA at position 49 in the proximal Ub were able to crosslink with $Rpn1^{391-642}$. This is fully consistent with published NMR data[405] that showed that K48- and K63-Ub2 have comparable binding affinities (K_D ~ 110 μM) for $Rpn1^{391-642}$. In addition, the presence of two crosslinking bands on the gel likely reflects the fact that $Rpn1^{(391–642)}$

contains two Ub-binding surfaces/sites,[417] hence possibly resulting in two distinct cross-linked configurations.

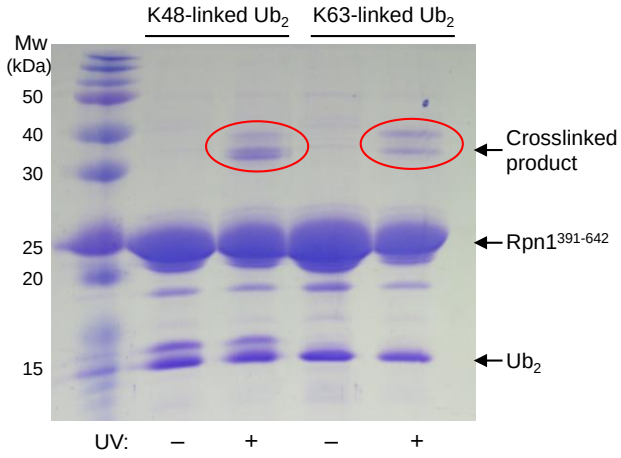


Figure 46. **Photo-crosslinking of K48- and K63-linked Ub2 chains having Bpa at position 49 in the proximal Ub with Rpn1391-642, visualized using Coomassie-stained SDS-PAGE gel of the reaction.**

5.3 Conclusions: Genetic incorporation of photo-crosslinkers into polyubiquitin efficiently identifies linkage-specific binding partners

Our current findings demonstrate that Ub chains can be converted into sensitive photo-affinity probes using unnatural amino acid mutagenesis. These probes can capture weak or transient protein-protein interactions which are often critical in Ub-mediated signaling. Importantly, we have shown that ubiquitination cascades catalyzed by E1 and E2 enzymes can accommodate Ub with unnatural amino acid labels to produce chains containing photo-affinity probes in the distal or proximal units. These dimers are functional and can form covalent linkages with known UBDs. This strategy can readily be extended to longer Ub chains, with the photo-crosslinker incorporated at any desired Ub unit (or multiple units) in the chain. This methodology can also be applied to make Ub chains containing other linkages using enzymatic or non-enzymatic

chain assembly.[418] Furthermore, genetic encoding of photo-affinity probes into Ub chains assembled from recombinant Ub monomers provides a viable alternative to total chemical synthesis as it can be achieved using tools readily available in a biochemical laboratory setting. Finally, we have identified hotspot mutations in the context of monomeric or dimeric Ub that are more reactive towards capturing intermolecular interactions. We also envision that it might be possible to fine tune crosslinking specificity for different binding partners by choosing different positions for probe(s) incorporation into a polyUb chain. This technology now sets the stage for finding and interrogation of unknown binding partners of polyUb chains using combinations of crosslinking in a cellular context followed by proteomics identification.

5.4 Contributions and acknowledgements

The development and incorporation of the unnatural amino acid with a BPA sidechain into Ub was performed by the Cropp lab, primarily by Dr. Courtney Braxton. Mr. Evan Quartner and I split the responsibilities for producing the BPA-containing material and running the SDS-PAGE assays. I performed the NMR experiments which validated the structural integrity and binding affinity of BPA-containing Ub.

Chapter 6. Concluding remarks

The Ub-signaling system is expansive and interwoven with Enzymes, UBDs, and PTMs working in concert to direct the flux of Ub, polyUb, and polyUb proteins throughout the cellular milieu. From the initial discovery that Ub signals proteins for proteasomal degradation, significant research effort was sparked into characterizing the entirety of the Ub-signaling system. The current breadth of knowledge has elucidated many of the most important components and their primary functions, but the intricate details and mechanisms of this complex network are still being uncovered.

The Ub E1 catalyzes the entirety of Ub signaling by activating Ub with a charged thioester that is then relayed by E2 and E3 enzymes to finally being conjugated to a substrate protein. Seminal research into the functionality of E1 characterized the entirety of the thioesterification mechanism using radiolabeled pyrophosphate, tritium-labeled ATP, and I^{125}-labeled Ub.[11] Only recently was any attempt made to revisit the kinetics of Ub activation, but these similarly used a fluorescently tagged Ub moiety and back-calculated the K_M from competition assays with WT Ub. We reasoned that NMR spectroscopy could easily, efficiently, and directly measure these reaction kinetics and allow for UBL or E1 screening reactions. I have presented that for UBL proteins, which contain a strong and isolated ^{1}H-^{15}N crosspeak for the C-terminal Gly, obtaining relative activation rates is trivial and easily allows for mapping the activation rate landscape of E1s. Substrate K_M values can readily be acquired by slightly adapting the assay setup, but the measurements remain straightforward. This method could easily be extended to DUB kinetics or E2 handoff assays to characterize additional systems.

UBDs transduce the signal from Ub or polyUb chains to downstream effectors. In this work I have commonly referenced UBA domains which have been shown to have affinities for K48- or K63-linked polyUb chains, but there are undoubtedly promiscuous or highly-specific UBDs for non-canonical polyUb chains which have yet to be identified. An interesting aspect of UBD identification, briefly touched on in the outline of DDI1 (1.1.3), is that the functionality of some Ub-interacting proteins may require long polyUb chains or specific types of branching for activation. Although we have a good understanding of the functions of canonical Ub linkages, often in form of dimeric or tetrameric Ub polymers, a dearth of data exists for other linkages.

UBDs are designed to have weak affinities for polyUb to ensure that the system is robust and dynamic. As a result, UBD-containing proteins are localized to long polyUb chains or areas with a high degree of E3 ligase activity. UBDs possess binding interfaces which complement those in Ub, often including the hydrophobic patch. However, these binding interfaces could be extended to encompass the binding interfaces of two or more Ub units, allowing for the selection of specific linkage types or branching.

The inherently weak interactions between UBDs and ubiquitin can be challenging to isolate using pull-down assays, as they may be overshadowed by stronger affinity complexes. As demonstrated in my work, this limitation can be overcome by using a photo-crosslinking reagent to convert these transient interactions into covalent attachments. By engineering a non-canonical BPA side chain into the hydrophobic patch of Ub, we were able to show that when this Ub monomer was mixed with UBAy, a well-known Ub-binding protein, and irradiated with UV light sufficient

to generate the BPA radical, the resulting covalent complex could be detected by SDS-PAGE. We expanded this to include studies with K48 Ub2 or K63 Ub2, prepared by conventional enzymatic synthesis, where the photo-activatable crosslinker could be designed into either the distal or proximal Ub. Reactions of either Ub dimer were shown to recapitulate the known binding affinities for the cognate binding partner, and we were able to show there were no instances of spurious crosslinking. This methodology is readily applicable to *in vitro* pulldown experiments using an immobilized Ub as bait, and could be extended to include polymeric Ub in any desired orientation and length by enzymatic assembly.

The functionality of DDI1 as an endopeptidase went undiscovered until only recently due to the hidden nature of its true UBD, the N-terminal ULD. yDDI1 has a negative phenotype with the proteasome, and through bioinformatics screens was identified to contain the putative ULD and UBA domains necessary to act as a proteasomal shuttle. During the evolution from yeast to human, DDI1 shed the UBA domain, presumedly along with the Ub-binding affinity. After the ULD of yDDI1 was shown to bind Ub, it was quickly reported that hDDI2 did not share the same affinity for Ub and it was surmised that the Ub-binding of yDDI is a historic functionality which is no longer necessary in vertebrates.

We have shown that the ULD of DDI1 in humans not only retains the ability to bind to Ub, but also contains a second binding site which dramatically increases the affinity for K48-linked polyUb chains. Protein paralogues frequently evolve to regulate specific functions through mechanisms such as competition between the more or less fit homologues, a diminished activity in the newer derivative, or a functional

specialization between the two. The interplay between hDDI1 and hDDI2 remains to be clarified. Both of these proteins have been shown to cleave long polyUb when the proteasome is stressed, but how DDI2 is recruited to polyUb chains remains a mystery. In contrast, the ULD of hDDI1, a UBD in this context, has an abnormally high affinity for K48-linked polyUb chains. The precise function of this elevated affinity in relation to peptidase activity remains enigmatic.

PTMs have been shown to interact with the Ub signaling system at multiple levels, including substrates, Ub conjugation enzymes, and Ub itself. The most well-studied PTM of Ub is phosphorylation by PINK1, and this phosphoUb is specifically recognized by a unique UBD in Parkin. Additionally, Ub has been shown to be acetylated at three lysines *in vivo* but there is currently no identified acetylase which works with Ub. Interestingly, a recent report showed that Ub lysines could react with Acetyl-CoA *in vitro* to generate N-acetyl amines. The lysine sidechains with the lowest pK_as were preferentially acetylated, and the *in vitro* profile aligned well with the acetylation profile detected *in vivo*.

In conjunction, these results evidence a potential for the sidechain amines of Ub to broadly react with electrophilic biomolecules. If these reactions occurred frequently on the intracellular pool of free Ub, then these amines would be blocked for conjugation reactions and Ub signaling could be modulated. This mechanism would suggest that environmental conditions, including the types and populations of electrophilic biomolecules, could passively regulate of Ub signaling.

As I reported, we were able to demonstrate that CO_2 is one such biomolecule which can interact with Ub amines. CO_2 can react with amines to form readily

reversible carbamate modifications, however only two Ub lysines were demonstrated to have this affinity *in vivo.* I was able to show that these two lysine residues possessed relatively low pKa values, albeit not exclusively. Using NMR assays at physiological pH, I was able to show that all four amines with the lowest pKas react reversibly with CO_2, and these same sidechains were preferentially carbamylated by isocyanic acid. Carbamylation is an often-overlooked PTM due to its reversible nature, however our results show that polyUb formation and E3 activity can be affected by the amount of environmental CO_2.

All in all my work highlights incremental steps to tease out some intricacies in Ub signaling pathways. I chose to tackle most of these questions from an NMR perspective to achieve site-specific information. This allowed me to characterize the kinetics of Ub activation, pK_as for each Ub sidechain amine and their carbamylation susceptibility, the binding mode and affinities for an unusual Ub-binding protein, and verify the integrity of a mutated Ub. I hope that this body of work will seed future discussions and lay the foundation for further advances in the Ub field.

Chapter 7. Materials and methods

7.1 Protein expression and purification

All buffers were prepared with 0.03 % NaN_3

7.1.1 Cell growth and lysis

Bacterial cultures were grown in LB to an optical density of 0.6 – 0.8 at 600 nm followed by the addition of IPTG to a final concentration of 1 mM and were shaken overnight at 18 °C for 16 hr. In all cases the protein yield under these conditions was similar to a 3 hr induction at 37 °C. After induction cells were harvested by centrifugation at 6000 rpm for 15 minutes. Cells were resuspended in the primary purification buffer to a volume of 30 mL with a spatula-tip of phenylmethylsulfonyl fluoride and 0.01% Triton X-100. Cells were lysed by sonication for 6 minutes in 30 s on 30 s off intervals. The cellular debris were pelleted by centrifugation at 25,000 rpm for 30 min and the lysate was syringe-filtered through a 0.45 μm filter immediately prior to purification.

For uniformly ^{15}N- or $^{13}C/^{15}N$-labeled material the procedure was slightly modified. After reaching an OD of 0.8 the cells were pelleted by centrifugation followed by resuspension into M9 minimal media without nutrients. The volume of minimal media was ¼ of the LB. The cells were acclimated to the new media for 45 min at 37°C. 1 g / L ^{15}N-ammonium chloride and 8 g/L dextrose or 4 g/L ^{13}C-dextrose, depending on carbon labeling requirements, were added and the cells were again shaken at 37°C for 30 minutes. After cooling to room temperature the cells were similarly induced by IPTG and for the same period of time.

7.1.2 Ub, UBL, and ULD purification

Ub monomers from H. sapiens were expressed in E. coli BL21(DE3) pJY2 competent cells. BPA-containing Ub monomers from H. sapiens were expressed in E. coli BL21(DE3) competent cells. Ub derivatives were purified as described previously.21 The ULD of DDI1 was expressed from BL21(DE3) Rosetta cells with a N-terminal Glutathione S-transferase (GST) linked by a Precision Protease cleavage site: Leu-Glu-Val-Leu-Phe-Gln/Gly-Pro which is followed by a sequence cloning artifact of Gly-Leu that does not interact with the ULD. Residues 1 – 81 of hDD1 were used to define the ULD. GST affinity purification by GSTrap FF column was performed in a buffer containing 20 mM sodium phosphate (NaP) pH 7.6 and 150 mM NaCl. The same buffer but with the addition of 15mM reduced glutathione was used to elute the tagged construct. 5 units of Precision protease were added to the elution and this solution was dialyzed overnight against the buffer without glutathione. The solution was passed through the GST affinity column a second time and the flowthrough was collected and concentrated to a final volume of 1.5 mL followed by size-exclusion chromatography (SEC) in the same buffer. The fractions containing the hDDI1 ULD were concentrated and diluted to 20 mL with 20 mM tris(hydroxymethyl)aminomethane (tris) pH 7. This was applied to a Q FF anion column and was eluted by a linear gradient from 0 – 80% of the same buffer containing 1 M NaCl. The ULD eluted at 150 mM NaCl and was exchanged into 20 mM Na Phos pH 6.8 and frozen at -20 °C. The VL_AK mutant was purified similarly but all buffers contained 20% glycerol.

NEDD8 was expressed from BL21(DE3) Rosetta cells with an N-terminal GST affinity tag and linked by a thrombin cleavage site: Leu-Val-Pro-Arg/Gly-Ser. This was

followed by an additional Gly-Ser cloning artifact which does not interact with NEDD8. The GST-fused construct was purified similarly to the hDDI1 ULD. After elution the protein was exchanged once into a buffer consisting of 20 mM NaP pH 8 and 150 mM NaCl and concentrated to 1.5 mL. 5 units of thrombin were added to the solution and was mixed at room temperature for 1.5 hr. After centrifugation to remove any precipitated protein pure NEDD8 was obtained from SEC utilizing a buffer consisting of 20 mM NaP pH 8 and 130 mM NaCl. The protein was exchanged once into a buffer consisting of 20 mM NaP pH 8 and used immediately. NEDD8 cannot be frozen in these conditions.

RUB1 and its derivatives were expressed from BL21(DE3) with a C-terminal chitin binding domain and purified by the NEB IMPACT Kit protocol. The RUB1 derivatives were expressed with a C-terminal intein / chitin binding domain motif and immobilized on a chitin resin after purification with the recommended buffer. Cleavage was induced by swapping the buffer to one containing dithiothreitol (DTT). After incubation for ~36 hours at r.t. the protein was eluted and replaced with fresh DTT-containing buffer. After another ~36 hours the protein was eluted and combined with the previous elution. Pure RUB1 was obtained after SEC with 20 mM NaP pH 8 and 130 mM NaCl, and was exchanged into the same buffer without salt before freezing at -20 °C.

7.1.3 Enzyme purification

Ub E1, E2-25K, and OTUB1 were purified as described previously[388]. Precision protease was expressed with an N-terminal GST affinity tag and purified similarly to the UBLs, and was bound to GSTrap FF column in a buffer containing 20 mM NaP and 150 mM NaCl at pH 7.5. After elution with glutathione the protein was exchanged into a buffer consisting of 20 mM NaP pH 7.5 and 150 mM NaCl prepared

with 30% glycerol (v/v). The heterodimer of APP/UBA3, without removal of the GST tag on APP, was purified as described previously.[239]

7.2 Enzymatic Ub chain assembly and disassembly

Chains were assembled enzymatically using a chain-terminator Ub building block carrying a K48R or K63R mutation on the distal Ub and a C-terminal extension, a hexa-histidine tag or single aspartate residue as D77, or truncation to Arg 74 on the proximal Ub. This allows full control of the E2-catalyzed conjugation reaction, and enables domain-specific incorporation of mutations or isotope labeling. The conjugating E2 enzymes used for the K48-linked and K63-linked dimers formation were E2-25K-GST and Ubc13-GST/Mms2-His_6, respectively.[419,420] These reactions were generally performed with 300 nM E1, 10 μM E2, 10 mM ATP, and 10 mM $MgCl_2$ in 50 mM tris at pH 7.5 and at 302 K.

7.3 NMR experiments

Unless specified, all experiments were acquired at the University of Maryland biochemical NMR facility with magnetic fields corresponding to a ^{1}H resonance frequency of 600 MHz or 800 MHz equipped with cryogenically-cooled probes. All 2D spectra were processed with NMRPipe[421] and analyzed in CCPNMR V2[422] and all 1D spectra were processed and analyzed using Mestrelab MNova within an NMRBox[423] environment. The acquisition time was typically limited to 75-80 ms for ^{15}N decoupling with a ^{1}H spectral width of 16 ppm at both 18.8 or 14.1 T magnets.

The distance of CSP were calculated from Equation 3 where the ^{15}N chemical shift is reduced fivefold to match the ^{1}H and ^{15}N chemical shift range for backbone amides.

$$\Delta\delta = \sqrt{(\Delta\delta_H)^2 + \left(\frac{\Delta\delta_N}{5}\right)^2} \quad (2)$$

7.3.1 1D ^{15}N NMR experiments

1D ^{15}N spectra for Ub at various pH conditions were acquired on a 500 MHz Varian Inova spectrometer with a room temperature broad band observe (BBO) probe at Johns Hopkins University. The reaction of Ub with isocyanic acid was monitored using 1D ^{15}N spectra recorded on a Bruker Avance-III 600 MHz spectrometer equipped with a room temperature BBO probe. Ub was exchanged into a buffer of 10 mM each of CHES and CAPS at pH 9.2 with 10% D_2O. The reaction was started by the addition of a 500 mM NaNCO stock solution to final concentrations of 3 mM Ub and 72 mM NaNCO. Experiments were acquired with 1024 scans and a 220 ppm spectral width centered at 30 ppm, 150 ms acquisition time, continuous wave proton decoupling for NOE enhancement, and a 2 s recycle delay.

7.2.2 Ubiquitin amine pK_a measurements

^{13}C/^{15}N labeled Ub was exchanged into a buffer containing 5 mM each of CHES and CAPS at a pH of 10.5 and separated into two 500 µL 1 mM aliquots that were run as two divergent samples for lower or higher pH measurements. After each experiment these samples were consecutively exchanged into the same buffer with either increasing or decreasing pH by the addition of 0.1 M HCl and 0.1 M KOH in approximately 0.3

(H2CN) or 0.4 (HSQC) pH unit steps. D_2O was added to each sample to a final concentration of 5% (v/v). The pK_a values were obtained by fitting the observed signal shifts to a Henderson-Hasselbalch equation modified with or without the Hill coefficient. The errors in reported pK_a values for each residue were assessed by fitting 2000 Monte Carlo generated synthetic data sets in which Gaussian noise was added to the chemical shifts according to the resolution in the respective H2CN or ^{1}H-^{13}C HSQC spectra. Both H2CN and ^{1}H-^{13}C HSQC spectra were acquired on a Bruker Avance-III 600 MHz spectrometer equipped with a TCI cryoprobe at 23°C. Signals from the H2CN experiments were assigned by a (H2C)N(CC)H-TOCSY experiment[424] .

7.2.3 Carbamylation measurements using H2CN spectral series

600 μM $^{13}C/^{15}N$ Ub was prepared in buffers containing either 20 mM potassium phosphate (pH 7.5) or 10 mM each of CHES and CAPS (High pH) containing 10% D_2O, 48 mM NaNCO, and at 23°C unless otherwise specified. $^{1}H/^{15}N$ H2CN spectra were acquired sequentially over 12.5 hrs with a 75 ms acquisition time and 2 s recycle delay. The reaction rates were obtained by fitting the percentage decrease in peak volumes to a first order exponential decay. Peak volumes for the magnetically inequivalent epsilon protons of K33 and K29 were combined before analysis.

7.2.4 Following the intensity of Ub G76 using 1D or 2D SOFAST series.

All reactions were performed with 20 mM NaP pH 8. Ub dimer or thioester formation was performed with 100 μM of the UBL, 4 mM ATP, 4 mM $MgCl_2$, and 40 mM Mesna or E2 to a final concentration of 10 μM. The reactions were started by the addition of E1 to a final concentration of 300 nM. The signal for the backbone amide

of G76 is an isolated crosspeak in 2D $^{1}H - ^{15}N$ correlation spectra when the experiments are recorded with a 35 ppm chemical shift window centered at 115 ppm in the ^{15}N dimension. Quantitative spectra were acquired with a ^{15}N spectral width of 8 ppm. The spectra were acquired in 5- or 10-minute increments and processed as a serries with uniform processing parameters. The spectra were recorded at 302 K unless otherwise specified. The position of G76 was identified in the first spectra and not adjusted. The decrease in intensity for this crosspeak was fit to a linear equation without defining the y-intercept. The reverse reaction was followed by digestion with OTUB1, wherein the chemical shift for G76 was assessed at the final spectrum.

1D SOFAST spectra were acquired with 1420 points over a spectral width of 16 ppm. The spectra were processed with exponential apodization of 5 Hz. The reaction series were acquired with 5- or 10-minute increments and the chemical shift of G76 was standardized to the maximum peak intensity near a ^{1}H chemical shift of 7.89 ppm. The peak maxima for Ub, RUB1, and NEDD8 varies slightly.

7.2.5 hULD NMR experiments

600 µM $^{13}C/^{15}N$ uniformly labeled hULD was exchanged into a buffer consisting of 20 mM NaP and a pH of 6.8 for $^{1}H/^{13}C/^{15}N$ triple resonance experiments. All experiments previously referenced (Chapter 4) were acquired at 298 K. Binding experiments were performed with both the hULD or Ub derivatives dissolved in this buffer and at this temperature. The experiments with the hULD VL_AK mutant were performed in buffers containing 10% (v/v) glycerol.

7.3 Data fitting

7.3.1 Determination of the activation enthalpies and entropies for carbamylation reaction

The temperature-dependent rates, *k*, of carbamylation for Ub lysines K6, K33, and K48 were fit (using linear regression) to the following Eyring equation to determine the enthalpy ($\Delta H^{\ddagger}$) and entropy ($\Delta S^{\ddagger}$) of activation:

$$\ln\frac{k}{T} = \frac{-\Delta H^{\ddagger}}{RT} + \ln\frac{\kappa k_B}{h} + \frac{\Delta S^{\ddagger}}{R}, \quad (3)$$

where *R* is the gas constant, *h* is the Planck constant, k_B is the Boltzmann constant; the transmission coefficient κ was set to 1.

7.3.2 Determination of the dissociation constant by NMR experiments

The dissociation constants were determined by fitting the experimental CSPs from various titration points to a model of single-site binding where Δδ is the experimental CSP, $\Delta\delta_{max}$ is the fitted max CSP, K_D is the calculated dissociation constant, and [Pt] and [Lt] were the protein and ligand concentrations. These data were fit to a single-site binding model using an in-house Matlab program Kdfit.[420] K_D was calculated for all residues simultaneously:

$$\Delta\delta = \Delta\delta_{max}\frac{[Pt]+[Lt]+K_D-\sqrt{([Pt]+[Lt]+K_D)^2-4[Pt][Lt]}}{2[Pt]} \quad (4)$$

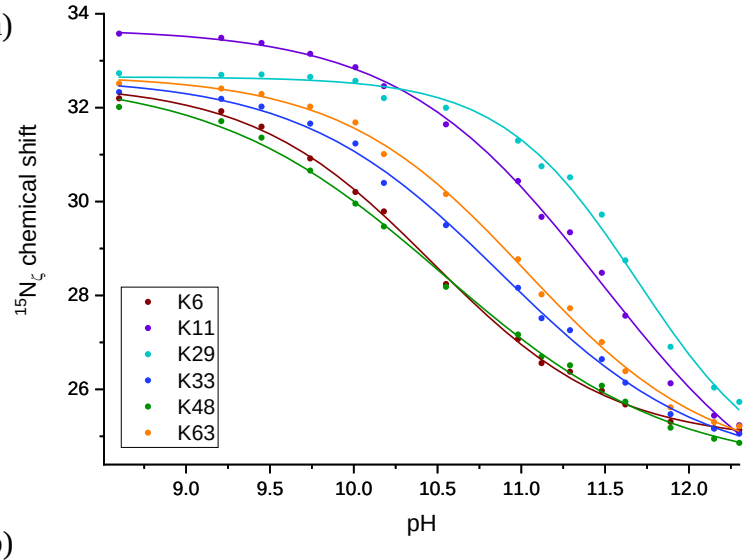
Appendices

Additional information:

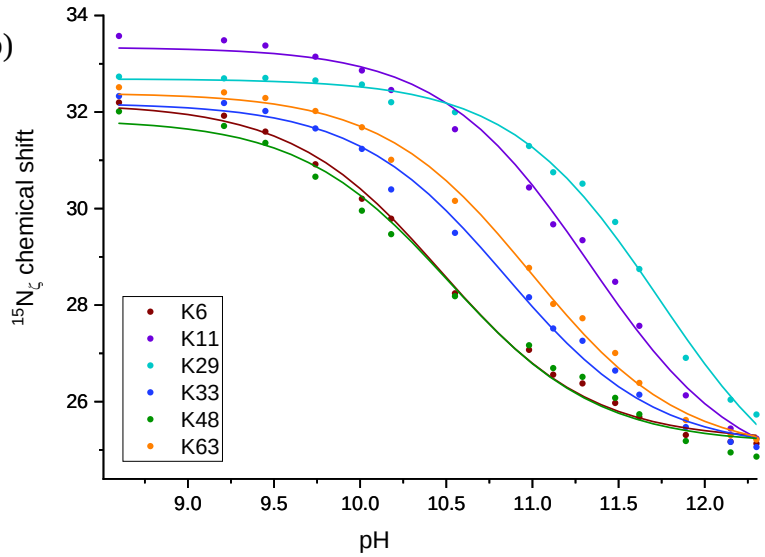
Appendix 1.1: pKa with and without Hill coefficient

The ^{15}N chemical shifts are shown below with (a) and without (b) the Hill coefficient. Although the inflection points are similar for the fits with either equation, the data are better represented with shallower slopes for all lysines except the partially buried sidechain of K29 where the Hill coefficient is near 1.

(a)



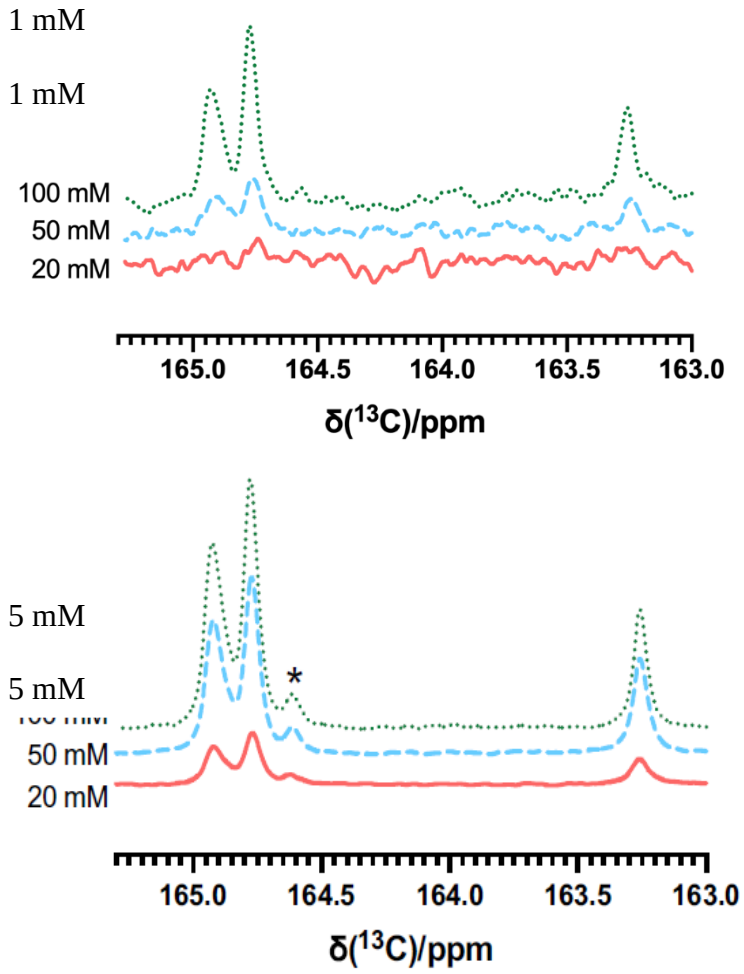
(b)



Appendix Table 1.1: Ubiquitin lysine pK_a s calculated with and without the Hill coefficient. *P** is the probability that the improvement in the fit when including the Hill coefficient has occurred by chance.

Amine	$^{15}N_\zeta$ pK_a	$^{15}N_\zeta$ pK_a (Hill)	*P**
K6	10.47 ± 0.02	10.47 ± 0.02	3.40E-12
K48	10.51 ± 0.02	10.53 ± 0.02	1.26E-13
K33	10.85 ± 0.02	10.88 ± 0.03	2.84E-12
K63	10.99 ± 0.02	11.03 ± 0.02	2.06E-11
K11	11.32 ± 0.01	11.47 ± 0.03	2.67E-11
K29	11.73 ± 0.02	11.68 ± 0.03	1.80E-09

Appendix 1.2: Ubiquitin carbamate signals are shown for 1 mM and 5 mM Ub.



Appendix figure 1.2. After adding ^{13}C-enriched bicarbonate new signals corresponding to carbon dioxide, bicarbonate/carbonate, and ubiquitin carbamates are evident. The same carbamate signals are seen at 1 mM and 5 mM ubiquitin. An additional low-intensity carbamate is potentially observed with 5 mM Ub and is indicated with an asterisk.

Appendix 1.3: Additional UBD interactions captured by photocrosslinking.

Appendix Table 1.3a: Crosslinking summary of Bpa-containing monoUbiquitin variants with ubiquilin-1 UBA and hHR23A UBA2 domains.

Residue position of BPA	UBDs	
	Ubiquilin-1 UBA	hHR23A UBA2
Residue 8	+	+
Residue 9	+	+
Residue 10	–	+
Residue 11	–	-
Residue 12	+	+
Residue 20	–	-
Residue 42	+	+
Residue 46	–	-
Residue 48	+	+
Residue 49	+	+
Residue 51	+	+
Residue 52	–	–
Residue 68	–	+
Residue 70	–	–
+ crosslinking; – no crosslinking		

Appendix Table 1.3b: Crosslinking summary of Bpa-containing monoUbiquitin variants with ubiquilin-1 UBA and hHR23A UBA2 domains.

Domain and residue position of BPA	UBDs	
	hHR23A UBA2	Rap80 tUIM
Control		
K48-linked Ub_2 (no- BPA control)	–	–
K48-linked Ub2iquitin		
Distal Ub, residue 9	+	+**
Proximal Ub, residue 9	+	+**
Proximal Ub, residue 11	–	–
Proximal Ub, residue 46	–	+**
Proximal Ub, residue 49	+	+**
K63-linked Ub2iquitin		
Distal Ub, residue 9	+**	+
Proximal Ub, residue 9	–	+
Proximal Ub, residue 11	–	–
Proximal Ub, residue 46	–	–
Proximal Ub, residue 48	–	–
Proximal Ub, residue 49	–	+
+ crosslinking; +** unexpected; – no crosslinking		

Bibliography

1. Pickart, C. M. & Fushman, D. Polyubiquitin chains: polymeric protein signals. *Current Opinion in Chemical Biology* **8**, 610–616 (2004).
2. Kliza, K. & Husnjak, K. Resolving the Complexity of Ubiquitin Networks. *Front. Mol. Biosci.* **7**, 21 (2020).
3. Dikic, I. & Schulman, B. A. An expanded lexicon for the ubiquitin code. *Nat Rev Mol Cell Biol* **24**, 273–287 (2023).
4. Wang, Y.-S., Wu, K.-P., Jiang, H.-K., Kurkute, P. & Chen, R.-H. Branched Ubiquitination: Detection Methods, Biological Functions and Chemical Synthesis. *Molecules* **25**, 5200 (2020).
5. Yau, R. & Rape, M. The increasing complexity of the ubiquitin code. *Nat Cell Biol* **18**, 579–586 (2016).
6. 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.* **77**, 1365–1368 (1980).
7. Ciechanover, A., Heller, H., Katz-Etzion, R. & Hershko, A. Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 761–765 (1981).
8. Ciechanover, A., Elias, S., Heller, H. & Hershko, A. 'Covalent affinity' purification of ubiquitin-activating enzyme. *Journal of Biological Chemistry* **257**, 2537–2542 (1982).
9. Hershko, A., Ciechanover, A., Heller, H., Haas, A. L. & Rose, I. A. Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1783–1786 (1980).
10. Hershko, A., Ciechanover, A. & Rose, I. A. Identification of the active amino acid residue of the polypeptide of ATP-dependent protein breakdown. *Journal of Biological Chemistry* **256**, 1525–1528 (1981).
11. Haas, A. L., Warms, J. V., Hershko, A. & Rose, I. A. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *Journal of Biological Chemistry* **257**, 2543–2548 (1982).
12. Ciechanover, A. The ubiquitin-proteasome proteolytic pathway. *Cell* **79**, 13–21 (1994).
13. Pickart, C. M. Targeting of substrates to the 26S proteasome. *FASEB j.* **11**, 1055–1066 (1997).
14. Hershko, A. The ubiquitin system for protein degradation and some of its roles in the control of the cell division cycle. *Cell Death & Differentiation* **12**, 1191–1197 (2005).
15. 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* **36**, 10526–10537 (1997).

16. Thrower, J. S. Recognition of the polyubiquitin proteolytic signal. *The EMBO Journal* **19**, 94–102 (2000).
17. Hoege, C., Pfander, B., Moldovan, G.-L., Pyrowolakis, G. & Jentsch, S. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* **419**, 135–141 (2002).
18. Randles, L. Ubiquitin and its binding domains. *Front Biosci* **17**, 2140 (2012).
19. Vijay-Kumar, S., Bugg, C. E. & Cook, W. J. Structure of ubiquitin refined at 1.8Åresolution. *Journal of Molecular Biology* **194**, 531–544 (1987).
20. Pickart, C. M. & Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1695**, 55–72 (2004).
21. Varadan, R., Walker, O., Pickart, C. & Fushman, D. Structural Properties of Polyubiquitin Chains in Solution. *Journal of Molecular Biology* **324**, 637–647 (2002).
22. Hurley, J. H., Lee, S. & Prag, G. Ubiquitin-binding domains. *Biochemical Journal* **399**, 361–372 (2006).
23. Winget, J. M. & Mayor, T. The Diversity of Ubiquitin Recognition: Hot Spots and Varied Specificity. *Molecular Cell* **38**, 627–635 (2010).
24. Eddins, M. J., Varadan, R., Fushman, D., Pickart, C. M. & Wolberger, C. Crystal Structure and Solution NMR Studies of Lys48-linked Tetraubiquitin at Neutral pH. *Journal of Molecular Biology* **367**, 204–211 (2007).
25. Haririnia, A., D'Onofrio, M. & Fushman, D. Mapping the Interactions between Lys48 and Lys63-Linked Di-ubiquitins and a Ubiquitin-Interacting Motif of S5a. *Journal of Molecular Biology* **368**, 753–766 (2007).
26. Nakasone, M. A., Livnat-Levanon, N., Glickman, M. H., Cohen, R. E. & Fushman, D. Mixed-Linkage Ubiquitin Chains Send Mixed Messages. *Structure* **21**, 727–740 (2013).
27. Nathan, J. A., Tae Kim, H., Ting, L., Gygi, S. P. & Goldberg, A. L. Why do cellular proteins linked to K63-polyubiquitin chains not associate with proteasomes? *EMBO J* **32**, 552–565 (2013).
28. Alfano, C., Faggiano, S. & Pastore, A. The Ball and Chain of Polyubiquitin Structures. *Trends in Biochemical Sciences* **41**, 371–385 (2016).
29. Haakonsen, D. L. & Rape, M. Branching Out: Improved Signaling by Heterotypic Ubiquitin Chains. *Trends in Cell Biology* **29**, 704–716 (2019).
30. Ohtake, F. Branched ubiquitin code: from basic biology to targeted protein degradation. *The Journal of Biochemistry* **171**, 361–366 (2022).
31. Lambert-Smith, I. A., Saunders, D. N. & Yerbury, J. J. The pivotal role of ubiquitin-activating enzyme E1 (UBA1) in neuronal health and neurodegeneration. *The International Journal of Biochemistry & Cell Biology* **123**, 105746 (2020).
32. Groen, E. J. N. & Gillingwater, T. H. UBA1: At the Crossroads of Ubiquitin Homeostasis and Neurodegeneration. *Trends in Molecular Medicine* **21**, 622–632 (2015).
33. Groettrup, M., Pelzer, C., Schmidtke, G. & Hofmann, K. Activating the ubiquitin family: UBA6 challenges the field. *Trends in Biochemical Sciences* **33**, 230–237 (2008).

34. Wijk, S. J. L. & Timmers, H. T. M. The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. *FASEB j.* **24**, 981–993 (2010).
35. Liu, W. *et al.* The Ubiquitin Conjugating Enzyme: An Important Ubiquitin Transfer Platform in Ubiquitin-Proteasome System. *IJMS* **21**, 2894 (2020).
36. Hosseini, S. M. *et al.* E2 ubiquitin-conjugating enzymes in cancer: Implications for immunotherapeutic interventions. *Clinica Chimica Acta* **498**, 126–134 (2019).
37. Bui, Q. T., Hong, J. H., Kwak, M., Lee, J. Y. & Lee, P. C.-W. Ubiquitin-Conjugating Enzymes in Cancer. *Cells* **10**, 1383 (2021).
38. Haas, A. L. & Siepmann, T. J. Pathways of ubiquitin conjugation. *FASEB j.* **11**, 1257–1268 (1997).
39. Pickart, C. M. Mechanisms Underlying Ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).
40. Deshaies, R. J. & Joazeiro, C. A. P. RING Domain E3 Ubiquitin Ligases. *Annu. Rev. Biochem.* **78**, 399–434 (2009).
41. Berndsen, C. E. & Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nat Struct Mol Biol* **21**, 301–307 (2014).
42. Buetow, L. & Huang, D. T. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nat Rev Mol Cell Biol* **17**, 626–642 (2016).
43. Weber, J., Polo, S. & Maspero, E. HECT E3 Ligases: A Tale With Multiple Facets. *Front. Physiol.* **10**, 370 (2019).
44. Yang, Q., Zhao, J., Chen, D. & Wang, Y. E3 ubiquitin ligases: styles, structures and functions. *Mol Biomed* **2**, 23 (2021).
45. Toma-Fukai, S. & Shimizu, T. Structural Diversity of Ubiquitin E3 Ligase. *Molecules* **26**, 6682 (2021).
46. Ye, Y. & Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* **10**, 755–764 (2009).
47. Stewart, M. D., Ritterhoff, T., Klevit, R. E. & Brzovic, P. S. E2 enzymes: more than just middle men. *Cell Res* **26**, 423–440 (2016).
48. Du, X., Song, H., Shen, N., Hua, R. & Yang, G. The Molecular Basis of Ubiquitin-Conjugating Enzymes (E2s) as a Potential Target for Cancer Therapy. *IJMS* **22**, 3440 (2021).
49. Chang, S.-C., Zhang, B.-X. & Ding, J. L. E2-E3 ubiquitin enzyme pairing - partnership in provoking or mitigating cancers. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1877**, 188679 (2022).
50. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).
51. Kumar, M. & Wiener, R. Closing the gap of ubiquitin activation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 15319–15321 (2019).
52. Komander, D., Clague, M. J. & Urbé, S. Breaking the chains: structure and function of the deubiquitinases. *Nat Rev Mol Cell Biol* **10**, 550–563 (2009).
53. Neutzner, M. & Neutzner, A. Enzymes of ubiquitination and deubiquitination. *Essays in Biochemistry* **52**, 37–50 (2012).
54. Huang, O. W. & Cochran, A. G. Regulation of deubiquitinase proteolytic activity. *Current Opinion in Structural Biology* **23**, 806–811 (2013).
55. Mevissen, T. E. T. & Komander, D. Mechanisms of Deubiquitinase Specificity and Regulation. *Annu. Rev. Biochem.* **86**, 159–192 (2017).

56. Mann, M. & Jensen, O. N. Proteomic analysis of post-translational modifications. *nature biotechnology* **21**, (2003).
57. Kirkpatrick, D. S., Denison, C. & Gygi, S. P. Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics. *Nat Cell Biol* **7**, 750–757 (2005).
58. Denis, N. J., Vasilescu, J., Lambert, J.-P., Smith, J. C. & Figeys, D. Tryptic digestion of ubiquitin standards reveals an improved strategy for identifying ubiquitinated proteins by mass spectrometry. *Proteomics* **7**, 868–874 (2007).
59. Komander, D. The emerging complexity of protein ubiquitination. *Biochemical Society Transactions* **37**, 937–953 (2009).
60. Kulathu, Y. & Komander, D. Atypical ubiquitylation — the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. *Nat Rev Mol Cell Biol* **13**, 508–523 (2012).
61. Beaudette, P., Popp, O. & Dittmar, G. Proteomic techniques to probe the ubiquitin landscape. *Proteomics* **16**, 273–287 (2016).
62. Tsuchiya, H. *et al.* Ub-ProT reveals global length and composition of protein ubiquitylation in cells. *Nat Commun* **9**, 524 (2018).
63. Swatek, K. N. *et al.* Insights into ubiquitin chain architecture using Ub-clipping. *Nature* **572**, 533–537 (2019).
64. Porras-Yakushi, T. R. & Hess, S. Recent advances in defining the ubiquitylome. *Expert Review of Proteomics* **11**, 477–490 (2014).
65. Burke, M. C. *et al.* Unexpected Trypsin Cleavage at Ubiquitinated Lysines. *Anal. Chem.* **87**, 8144–8148 (2015).
66. Cannon, J., Nakasone, M., Fushman, D. & Fenselau, C. Proteomic Identification and Analysis of K63-Linked Ubiquitin Conjugates. *Anal. Chem.* **84**, 10121–10128 (2012).
67. Lee, A. E. *et al.* Preparing to read the ubiquitin code: top-down analysis of unanchored ubiquitin tetramers: Top-down analysis of Ub tetramers. *Journal of Mass Spectrometry* **51**, 629–637 (2016).
68. Gomes, F. *et al.* Top-down analysis of novel synthetic branched proteins. *J Mass Spectrom* **54**, 19–25 (2019).
69. Chen, D. *et al.* Top-Down Analysis of Branched Proteins Using Mass Spectrometry. *Anal. Chem.* **90**, 4032–4038 (2018).
70. Fulzele, A. & Bennett, E. J. Ubiquitin diGLY Proteomics as an Approach to Identify and Quantify the Ubiquitin-Modified Proteome. in *The Ubiquitin Proteasome System* (eds. Mayor, T. & Kleiger, G.) vol. 1844 363–384 (Springer New York, 2018).
71. Xu, G. & Jaffrey, S. R. Proteomic identification of protein ubiquitination events. *Biotechnology and Genetic Engineering Reviews* **29**, 73–109 (2013).
72. Vere, G., Kealy, R., Kessler, B. M. & Pinto-Fernandez, A. Ubiquitomics: An Overview and Future. *Biomolecules* **10**, 1453 (2020).
73. Hjerpe, R. *et al.* Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO Rep* **10**, 1250–1258 (2009).
74. Lopitz-Otsoa, F. *et al.* Integrative analysis of the ubiquitin proteome isolated using Tandem Ubiquitin Binding Entities (TUBEs). *Journal of Proteomics* **75**, 2998–3014 (2012).

75. Mattern, M., Sutherland, J., Kadimisetty, K., Barrio, R. & Rodriguez, M. S. Using Ubiquitin Binders to Decipher the Ubiquitin Code. *Trends in Biochemical Sciences* **44**, 599–615 (2019).
76. Mulder, M. P. C., Witting, K. F. & Ovaa, H. Cracking the Ubiquitin Code: The Ubiquitin Toolbox. *Current Issues in Molecular Biology* 1–20 (2020) doi:10.21775/cimb.037.001.
77. Sun, M. & Zhang, X. Current methodologies in protein ubiquitination characterization: from ubiquitinated protein to ubiquitin chain architecture. *Cell Biosci* **12**, 126 (2022).
78. Sigismund, S., Polo, S. & Di Fiore, P. P. Signaling Through Monoubiquitination. in *Signalling from Internalized Growth Factor Receptors* (ed. Madshus, I. H.) 149–185 (Springer, 2004). doi:10.1007/978-3-540-69494-6_6.
79. Hoeller, D. *et al.* Regulation of ubiquitin-binding proteins by monoubiquitination. *Nat Cell Biol* **8**, 163–169 (2006).
80. Sewduth, R. N., Baietti, M. F. & Sablina, A. A. Cracking the Monoubiquitin Code of Genetic Diseases. *IJMS* **21**, 3036 (2020).
81. Gavrilov, Y., Hagai, T. & Levy, Y. Nonspecific yet decisive: Ubiquitination can affect the native-state dynamics of the modified protein. *Protein Science* **24**, 1580–1592 (2015).
82. Hagai, T. & Levy, Y. Ubiquitin not only serves as a tag but also assists degradation by inducing protein unfolding. *Proceedings of the National Academy of Sciences* **107**, 2001–2006 (2010).
83. Carroll, E. C. & Marqusee, S. Site-specific ubiquitination: Deconstructing the degradation tag. *Current Opinion in Structural Biology* **73**, 102345 (2022).
84. Braten, O. *et al.* Numerous proteins with unique characteristics are degraded by the 26S proteasome following monoubiquitination. *Proceedings of the National Academy of Sciences* **113**, E4639–E4647 (2016).
85. Ronai, Z. A. Monoubiquitination in proteasomal degradation. *Proc Natl Acad Sci U S A* **113**, 8894–8896 (2016).
86. Livneh, I., Kravtsova-Ivantsiv, Y., Braten, O., Kwon, Y. T. & Ciechanover, A. Monoubiquitination joins polyubiquitination as an esteemed proteasomal targeting signal. *BioEssays* **39**, 1700027 (2017).
87. Varadan, R. Solution Conformation of Lys63-linked Di-ubiquitin Chain Provides Clues to Functional Diversity of Polyubiquitin Signaling. *Journal of Biological Chemistry* **279**, 7055–7063 (2003).
88. Dittmar, G. & Winklhofer, K. F. Linear Ubiquitin Chains: Cellular Functions and Strategies for Detection and Quantification. *Front. Chem.* **7**, 915 (2020).
89. Oikawa, D., Sato, Y., Ito, H. & Tokunaga, F. Linear Ubiquitin Code: Its Writer, Erasers, Decoders, Inhibitors, and Implications in Disorders. *IJMS* **21**, 3381 (2020).
90. Ikeda, F. *et al.* SHARPIN forms a linear ubiquitin ligase complex regulating NF-κB activity and apoptosis. *Nature* **471**, 637–641 (2011).
91. van Wijk, S. J. L. *et al.* Linear ubiquitination of cytosolic Salmonella Typhimurium activates NF-κB and restricts bacterial proliferation. *Nat Microbiol* **2**, 17066 (2017).

92. Rieser, E., Cordier, S. M. & Walczak, H. Linear ubiquitination: a newly discovered regulator of cell signalling. *Trends in Biochemical Sciences* **38**, 94–102 (2013).
93. Spit, M., Rieser, E. & Walczak, H. Linear ubiquitination at a glance. *Journal of Cell Science* **132**, jcs208512 (2019).
94. Tang, Y. *et al.* Linear ubiquitination of cFLIP induced by LUBAC contributes to TNFα-induced apoptosis. *J Biol Chem* **293**, 20062–20072 (2018).
95. Priem, D. *et al.* A20 protects cells from TNF-induced apoptosis through linear ubiquitin-dependent and -independent mechanisms. *Cell Death Dis* **10**, 1–16 (2019).
96. Wu-Baer, F., Lagrazon, K., Yuan, W. & Baer, R. The BRCA1/BARD1 Heterodimer Assembles Polyubiquitin Chains through an Unconventional Linkage Involving Lysine Residue K6 of Ubiquitin. *Journal of Biological Chemistry* **278**, 34743–34746 (2003).
97. 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. *Human Molecular Genetics* **13**, 807–817 (2004).
98. Ohta, T., Sato, K. & Wu, W. The BRCA1 ubiquitin ligase and homologous recombination repair. *FEBS Letters* **585**, 2836–2844 (2011).
99. Huang, Q. & Zhang, X. Emerging Roles and Research Tools of Atypical Ubiquitination. *Proteomics* **20**, 1900100 (2020).
100. Tracz, M. & Bialek, W. Beyond K48 and K63: non-canonical protein ubiquitination. *Cell Mol Biol Lett* **26**, 1 (2021).
101. Park, G. H., Park, J. H. & Chung, K. C. Precise control of mitophagy through ubiquitin proteasome system and deubiquitin proteases and their dysfunction in Parkinson's disease. *BMB Rep* **54**, 592–600 (2021).
102. Durcan, T. M. *et al.* USP8 regulates mitophagy by removing K6-linked ubiquitin conjugates from parkin. *The EMBO Journal* **33**, 2473–2491 (2014).
103. Matsumoto, M. L. *et al.* K11-Linked Polyubiquitination in Cell Cycle Control Revealed by a K11 Linkage-Specific Antibody. *Molecular Cell* **39**, 477–484 (2010).
104. Rape, M. Assembly of K11-Linked Ubiquitin Chains by the Anaphase-Promoting Complex. in *Conjugation and Deconjugation of Ubiquitin Family Modifiers: Subcellular Biochemistry* (ed. Groettrup, M.) 107–115 (Springer, 2010). doi:10.1007/978-1-4419-6676-6_9.
105. Wickliffe, K. E., Williamson, A., Meyer, H.-J., Kelly, A. & Rape, M. K11-linked ubiquitin chains as novel regulators of cell division. *Trends in Cell Biology* **21**, 656–663 (2011).
106. Kravtsova-Ivantsiv, Y. & Ciechanover, A. Non-canonical ubiquitin-based signals for proteasomal degradation. *Journal of Cell Science* **125**, 539–548 (2012).
107. Castañeda, C. A., Kashyap, T. R., Nakasone, M. A., Krueger, S. & Fushman, D. Unique Structural, Dynamical, and Functional Properties of K11-Linked Polyubiquitin Chains. *Structure* **21**, 1168–1181 (2013).
108. Bonacci, T. & Emanuele, M. J. Impressionist portraits of mitotic exit: APC/C, K11-linked ubiquitin chains and Cezanne. *Cell Cycle* **18**, 652–660 (2019).

109. Bonacci, T. *et al.* Cezanne/OTUD7B is a cell cycle-regulated deubiquitinase that antagonizes the degradation of APC/C substrates. *The EMBO Journal* **37**, e98701 (2018).
110. Bremm, A., Freund, S. M. V. & Komander, D. Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. *Nat Struct Mol Biol* **17**, 939–947 (2010).
111. Boughton, A. J., Krueger, S. & Fushman, D. Branching via K11 and K48 Bestows Ubiquitin Chains with a Unique Interdomain Interface and Enhanced Affinity for Proteasomal Subunit Rpn1. *Structure* **28**, 29-43.e6 (2020).
112. French, M. E., Koehler, C. F. & Hunter, T. Emerging functions of branched ubiquitin chains. *Cell Discov* **7**, 6 (2021).
113. Kolla, S., Ye, M., Mark, K. G. & Rapé, M. Assembly and function of branched ubiquitin chains. *Trends in Biochemical Sciences* **47**, 759–771 (2022).
114. Middleton, A. J. & Day, C. L. From seeds to trees: how E2 enzymes grow ubiquitin chains. *Biochemical Society Transactions* **51**, 353–362 (2023).
115. Grice, G. L. *et al.* The Proteasome Distinguishes between Heterotypic and Homotypic Lysine-11-Linked Polyubiquitin Chains. *Cell Reports* **12**, 545–553 (2015).
116. Castañeda, C. A. *et al.* Linkage via K27 Bestows Ubiquitin Chains with Unique Properties among Polyubiquitins. *Structure* **24**, 423–436 (2016).
117. Pan, M. *et al.* Chemical Protein Synthesis Enabled Mechanistic Studies on the Molecular Recognition of K27-linked Ubiquitin Chains. *Angew. Chem. Int. Ed.* **58**, 2627–2631 (2019).
118. Gatti, M. *et al.* RNF168 Promotes Noncanonical K27 Ubiquitination to Signal DNA Damage. *Cell Reports* **10**, 226–238 (2015).
119. Kelliher, J., Ghosal, G. & Leung, J. W. C. New answers to the old RIDDLE: RNF168 and the DNA damage response pathway. *The FEBS Journal* **289**, 2467–2480 (2022).
120. Kitevski-LeBlanc, J. *et al.* The RNF168 paralog RNF169 defines a new class of ubiquitylated histone reader involved in the response to DNA damage. *eLife* **6**, e23872 (2017).
121. Nucifora, F. C. *et al.* Ubiqutination via K27 and K29 chains signals aggregation and neuronal protection of LRRK2 by WSB1. *Nature Communications* **7**, 11792 (2016).
122. Zhou, Q. & Zhang, J. K27-linked noncanonic ubiquitination in immune regulation. *Journal of Leukocyte Biology* **111**, 223–235 (2021).
123. van Huizen, M. & Kikkert, M. The Role of Atypical Ubiquitin Chains in the Regulation of the Antiviral Innate Immune Response. *Front. Cell Dev. Biol.* **7**, 392 (2020).
124. Buneeva, O. & Medvedev, A. Atypical Ubiquitination and Parkinson's Disease. *IJMS* **23**, 3705 (2022).
125. Yuan, W.-C. *et al.* K33-Linked Polyubiquitination of Coronin 7 by Cul3-KLHL20 Ubiquitin E3 Ligase Regulates Protein Trafficking. *Molecular Cell* **54**, 586–600 (2014).
126. Huang, H. *et al.* K33-Linked Polyubiquitination of T Cell Receptor-ζ Regulates Proteolysis-Independent T Cell Signaling. *Immunity* **33**, 60–70 (2010).

127. Roscoe, B. P., Thayer, K. M., Zeldovich, K. B., Fushman, D. & Bolon, D. N. A. Analyses of the Effects of All Ubiquitin Point Mutants on Yeast Growth Rate. *Journal of Molecular Biology* **425**, 1363–1377 (2013).
128. Glickman, M. H. & Ciechanover, A. The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiological Reviews* **82**, 373–428 (2002).
129. Finley, D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* **78**, 477–513 (2009).
130. Grice, G. L. & Nathan, J. A. The recognition of ubiquitinated proteins by the proteasome. *Cell. Mol. Life Sci.* **73**, 3497–3506 (2016).
131. Saeki, Y. Ubiquitin recognition by the proteasome. *J Biochem* **161**, 113–124 (2017).
132. Liu, Z. *et al.* Structural basis for the recognition of K48-linked Ub chain by proteasomal receptor Rpn13. *Cell Discov* **5**, 19 (2019).
133. Davis, C., Spaller, B. L. & Matouschek, A. Mechanisms of substrate recognition by the 26S proteasome. *Current Opinion in Structural Biology* **67**, 161–169 (2021).
134. Osei-Amponsa, V. & Walters, K. J. Proteasome substrate receptors and their therapeutic potential. *Trends in Biochemical Sciences* **47**, 950–964 (2022).
135. Sims, J. J. & Cohen, R. E. Linkage-Specific Avidity Defines the Lysine 63-Linked Polyubiquitin-Binding Preference of Rap80. *Molecular Cell* **33**, 775–783 (2009).
136. Lim, K.-L. & Lim, G. G. Y. K63-linked ubiquitination and neurodegeneration. *Neurobiology of Disease* **43**, 9–16 (2011).
137. Chen, Z. J. Ubiquitination in signaling to and activation of IKK: Ubiquitin-mediated activation of IKK. *Immunological Reviews* **246**, 95–106 (2012).
138. Wang, G. *et al.* K63-Linked Ubiquitination in Kinase Activation and Cancer. *Front. Oncol.* **2**, (2012).
139. Erpapazoglou, Z., Walker, O. & Haguenauer-Tsapis, R. Versatile Roles of K63-Linked Ubiquitin Chains in Trafficking. *Cells* **3**, 1027–1088 (2014).
140. Iwai, K. Diverse roles of the ubiquitin system in NF-κB activation. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1843**, 129–136 (2014).
141. Cao, L., Liu, X., Zheng, B., Xing, C. & Liu, J. Role of K63-linked ubiquitination in cancer. *Cell Death Discov.* **8**, 410 (2022).
142. Buchberger, A. From UBA to UBX: new words in the ubiquitin vocabulary. *Trends in Cell Biology* **12**, 216–221 (2002).
143. Madura, K. The Ubiquitin-Associated (UBA) Domain: On the Path from Prudence to Prurience. *Cell Cycle* **1**, 233–242 (2002).
144. Grabbe, C. & Dikic, I. Functional Roles of Ubiquitin-Like Domain (ULD) and Ubiquitin-Binding Domain (UBD) Containing Proteins. *Chem. Rev.* **109**, 1481–1494 (2009).
145. Husnjak, K. & Dikic, I. Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu. Rev. Biochem.* **81**, 291–322 (2012).
146. Chen, L. & Madura, K. Rad23 Promotes the Targeting of Proteolytic Substrates to the Proteasome. *Mol Cell Biol* **22**, 4902–4913 (2002).
147. Zhang, D. *et al.* Together, Rpn10 and Dsk2 Can Serve as a Polyubiquitin Chain-Length Sensor. *Molecular Cell* **36**, 1018–1033 (2009).

148. Chen, X. *et al.* Structures of Rpn1 T1:Rad23 and hRpn13:hPLIC2 Reveal Distinct Binding Mechanisms between Substrate Receptors and Shuttle Factors of the Proteasome. *Structure* **24**, 1257–1270 (2016).
149. Pohl, C. & Dikic, I. Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science* **366**, 818–822 (2019).
150. Elsasser, S. & Finley, D. Delivery of ubiquitinated substrates to protein-unfolding machines. *Nat Cell Biol* **7**, 742–749 (2005).
151. Hicke, L., Schubert, H. L. & Hill, C. P. Ubiquitin-binding domains. *Nat. Rev. Mol. Cell Biol.* **6**, 610–621 (2005).
152. Fu, H., Lin, Y.-L. & Fatimababy, A. S. Proteasomal recognition of ubiquitylated substrates. *Trends in Plant Science* **15**, 375–386 (2010).
153. Zientara-Rytter, K. & Subramani, S. The Roles of Ubiquitin-Binding Protein Shuttles in the Degradative Fate of Ubiquitinated Proteins in the Ubiquitin-Proteasome System and Autophagy. *Cells* **8**, 40 (2019).
154. Johnston, H. E. & Samant, R. S. Alternative systems for misfolded protein clearance: life beyond the proteasome. *FEBS J* **288**, 4464–4487 (2021).
155. Raasi, S., Varadan, R., Fushman, D. & Pickart, C. M. Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nat Struct Mol Biol* **12**, 708–714 (2005).
156. Dantuma, N. P., Heinen, C. & Hoogstraten, D. The ubiquitin receptor Rad23: At the crossroads of nucleotide excision repair and proteasomal degradation. *DNA Repair* **8**, 449–460 (2009).
157. Lu, K., den Brave, F. & Jentsch, S. Receptor oligomerization guides pathway choice between proteasomal and autophagic degradation. *Nat Cell Biol* **19**, 732–739 (2017).
158. Zhang, D., Raasi, S. & Fushman, D. Affinity Makes the Difference: Nonselective Interaction of the UBA Domain of Ubiquilin-1 with Monomeric Ubiquitin and Polyubiquitin Chains. *Journal of Molecular Biology* **377**, 162–180 (2008).
159. Raasi, S. & Pickart, C. M. Rad23 Ubiquitin-associated Domains (UBA) Inhibit 26 S Proteasome-catalyzed Proteolysis by Sequestering Lysine 48-linked Polyubiquitin Chains. *Journal of Biological Chemistry* **278**, 8951–8959 (2003).
160. Ortolan, T. G. *et al.* The DNA repair protein Rad23 is a negative regulator of multi-ubiquitin chain assembly. *Nat Cell Biol* **2**, 601–608 (2000).
161. Matiuhin, Y. *et al.* Extraproteasomal Rpn10 Restricts Access of the Polyubiquitin-Binding Protein Dsk2 to Proteasome. *Molecular Cell* **32**, 415–425 (2008).
162. Yokoi, M. & Hanaoka, F. Two mammalian homologs of yeast Rad23, HR23A and HR23B, as multifunctional proteins. *Gene* **597**, 1–9 (2017).
163. Zhang, K. Y., Yang, S., Warraich, S. T. & Blair, I. P. Ubiquilin 2: A component of the ubiquitin–proteasome system with an emerging role in neurodegeneration. *The International Journal of Biochemistry & Cell Biology* **50**, 123–126 (2014).
164. Stieren, E. S. *et al.* Ubiquilin-1 Is a Molecular Chaperone for the Amyloid Precursor Protein. *Journal of Biological Chemistry* **286**, 35689–35698 (2011).
165. El Ayadi, A., Stieren, E. S., Barral, J. M. & Boehning, D. Ubiquilin-1 and protein quality control in Alzheimer disease. *Prion* **7**, 164–169 (2013).
166. Subudhi, I. & Shorter, J. Ubiquilin 2: Shuttling Clients Out of Phase? *Molecular Cell* **69**, 919–921 (2018).

167. Fry, M. Y., Saladi, S. M. & Clemons, W. M. The STI1-domain is a flexible alpha-helical fold with a hydrophobic groove. *Protein Science* **30**, 882–898 (2021).
168. Nowicka, U. *et al.* DNA-Damage-Inducible 1 Protein (Ddi1) Contains an Uncharacteristic Ubiquitin-like Domain that Binds Ubiquitin. *Structure* **23**, 542–557 (2015).
169. Kaplun, L. *et al.* The DNA Damage-Inducible UbL-UbA Protein Ddi1 Participates in Mec1-Mediated Degradation of Ho Endonuclease. *Mol Cell Biol* **25**, 5355–5362 (2005).
170. Gabriely, G., Kama, R., Gelin-Licht, R. & Gerst, J. E. Different Domains of the UBL-UBA Ubiquitin Receptor, Ddi1/Vsm1, Are Involved in Its Multiple Cellular Roles. *MBoC* **19**, 3625–3637 (2008).
171. Walters, K. J., Lech, P. J., Goh, A. M., Wang, Q. & Howley, P. M. DNA-repair protein hHR23a alters its protein structure upon binding proteasomal subunit S5a. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12694–12699 (2003).
172. Dirac-Svejstrup, A. B. *et al.* DDI2 Is a Ubiquitin-Directed Endoprotease Responsible for Cleavage of Transcription Factor NRF1. *Molecular Cell* **79**, 332-341.e7 (2020).
173. Trempe, J.-F. *et al.* Structural studies of the yeast DNA damage-inducible protein Ddi1 reveal domain architecture of this eukaryotic protein family. *Scientific Reports* **6**, (2016).
174. Siva, M. *et al. DDI2 protease activity controls embryonic development and inflammation via TCF11/NRF1*. http://biorxiv.org/lookup/doi/10.1101/2020.12.16.423023 (2020) doi:10.1101/2020.12.16.423023.
175. Asaithambi, K., Biswas, I. & Suguna, K. Structural and functional insights into the DNA damage-inducible protein 1 (Ddi1) from protozoa. *Current Research in Structural Biology* **4**, 175–191 (2022).
176. Liu, Z. *et al.* HIV protease inhibitor nelfinavir is a potent drug candidate against echinococcosis by targeting Ddi1-like protein. *eBioMedicine* **82**, 104177 (2022).
177. Op, M. *et al.* The aspartyl protease DDI2 drives adaptation to proteasome inhibition in multiple myeloma. *Cell Death Dis* **13**, 475 (2022).
178. Kumar, S. & Suguna, K. Crystal structure of the retroviral protease-like domain of a protozoal DNA damage-inducible 1 protein. *FEBS Open Bio* **8**, 1379–1394 (2018).
179. Lehrbach, N. J. & Ruvkun, G. Proteasome dysfunction triggers activation of SKN-1A/Nrf1 by the aspartic protease DDI-1. *eLife* **5**, e17721 (2016).
180. Kottemann, M. C., Conti, B. A., Lach, F. P. & Smogorzewska, A. Removal of RTF2 from Stalled Replisomes Promotes Maintenance of Genome Integrity. *Molecular Cell* **69**, 24-35.e5 (2018).
181. Sivá, M. *et al.* Human DNA-Damage-Inducible 2 Protein Is Structurally and Functionally Distinct from Its Yeast Ortholog. *Sci Rep* **6**, 30443 (2016).
182. Krylov, D. M. & Koonin, E. V. Correspondence: A novel family of predicted retroviral-like aspartyl proteases with a possible key role in eukaryotic cell cycle control. *Current Biology* **11**, R584–R587 (2001).
183. Ridewood, S. *et al. The Aspartyl Protease Ddi1 Is Essential for Erythrocyte Invasion by the Malaria Parasite*.

http://biorxiv.org/lookup/doi/10.1101/2021.05.11.443575 (2021) doi:10.1101/2021.05.11.443575.
184. Onchieku, N. M. *et al.* Artemisinin Binds and Inhibits the Activity of Plasmodium falciparum Ddi1, a Retroviral Aspartyl Protease. *Pathogens* **10**, 1465 (2021).
185. Yip, M. C. J., Bodnar, N. O. & Rapoport, T. A. Ddi1 is a ubiquitin-dependent protease. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 7776–7781 (2020).
186. Bouvier, L. A., Niemirowicz, G. T., Salas-Sarduy, E., Cazzulo, J. J. & Alvarez, V. E. DNA-damage inducible protein 1 is a conserved metacaspase substrate that is cleaved and further destabilized in yeast under specific metabolic conditions. *The FEBS Journal* **285**, 1097–1110 (2018).
187. Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chemical Reviews* **118**, 889–918 (2018).
188. Hochstrasser, M. Evolution and function of ubiquitin-like protein-conjugation systems. *Nat Cell Biol* **2**, E153–E157 (2000).
189. Huang, D. T., Walden, H., Duda, D. & Schulman, B. A. Ubiquitin-like protein activation. *Oncogene* **23**, 1958–1971 (2004).
190. Ronau, J. A., Beckmann, J. F. & Hochstrasser, M. Substrate specificity of the ubiquitin and Ubl proteases. *Cell Res* **26**, 441–456 (2016).
191. Nakatogawa, H., Ichimura, Y. & Ohsumi, Y. Atg8, a Ubiquitin-like Protein Required for Autophagosome Formation, Mediates Membrane Tethering and Hemifusion. *Cell* **130**, 165–178 (2007).
192. Geng, J. & Klionsky, D. J. The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. *EMBO Rep* **9**, 859–864 (2008).
193. Shpilka, T., Weidberg, H., Pietrokovski, S. & Elazar, Z. Atg8: an autophagy-related ubiquitin-like protein family. *Genome Biol* **12**, 226 (2011).
194. Kirkin, V. & Rogov, V. V. A Diversity of Selective Autophagy Receptors Determines the Specificity of the Autophagy Pathway. *Molecular Cell* **76**, 268–285 (2019).
195. Zhang, D. & Zhang, D.-E. Interferon-Stimulated Gene 15 and the Protein ISGylation System. *J Interferon Cytokine Res* **31**, 119–130 (2011).
196. Wang, Z., Li, T., Gong, Z. & Xie, J. Role of ISG15 post-translational modification in immunity against Mycobacterium tuberculosis infection. *Cellular Signalling* **94**, 110329 (2022).
197. Narasimhan, J. *et al.* Crystal Structure of the Interferon-induced Ubiquitin-like Protein ISG15. *Journal of Biological Chemistry* **280**, 27356–27365 (2005).
198. Dzimianski, J. V., Scholte, F. E. M., Bergeron, É. & Pegan, S. D. ISG15: It's Complicated. *Journal of Molecular Biology* **431**, 4203–4216 (2019).
199. Perng, Y.-C. & Lenschow, D. J. ISG15 in antiviral immunity and beyond. *Nat Rev Microbiol* **16**, 423–439 (2018).
200. Morales, D. J. & Lenschow, D. J. The Antiviral Activities of ISG15. *Journal of Molecular Biology* **425**, 4995–5008 (2013).
201. Freitas, B. T., Scholte, F. E. M., Bergeron, É. & Pegan, S. D. How ISG15 combats viral infection. *Virus Research* **286**, 198036 (2020).
202. Gold, I. M., Reis, N., Glaser, F. & Glickman, M. H. Coronaviral PLpro proteases and the immunomodulatory roles of conjugated versus free Interferon Stimulated Gene product-15 (ISG15). *Semin Cell Dev Biol* **132**, 16–26 (2022).

203. Zhang, Y. *et al.* The in vivo ISGylome links ISG15 to metabolic pathways and autophagy upon Listeria monocytogenes infection. *Nat Commun* **10**, 5383 (2019).
204. Durfee, L. A., Lyon, N., Seo, K. & Huibregtse, J. M. The ISG15 Conjugation System Broadly Targets Newly Synthesized Proteins: Implications for the Antiviral Function of ISG15. *Molecular Cell* **38**, 722–732 (2010).
205. Müller, S., Hoege, C., Pyrowolakis, G. & Jentsch, S. Sumo, ubiquitin's mysterious cousin. *Nat Rev Mol Cell Biol* **2**, 202–210 (2001).
206. Eisenhardt, N. *et al.* A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. *Nat Struct Mol Biol* **22**, 959–967 (2015).
207. Chang, H.-M. & Yeh, E. T. H. SUMO: From Bench to Bedside. *Physiological Reviews* **100**, 1599–1619 (2020).
208. Keiten-Schmitz, J., Schunck, K. & Müller, S. SUMO Chains Rule on Chromatin Occupancy. *Front. Cell Dev. Biol.* **7**, 343 (2020).
209. Han, Z.-J., Feng, Y.-H., Gu, B.-H., Li, Y.-M. & Chen, H. The post-translational modification, SUMOylation, and cancer (Review). *International Journal of Oncology* **52**, 1081–1094 (2018).
210. Lascorz, J., Codina-Fabra, J., Reverter, D. & Torres-Rosell, J. SUMO-SIM interactions: From structure to biological functions. *Seminars in Cell & Developmental Biology* **132**, 193–202 (2022).
211. Jansen, N. S. & Vertegaal, A. C. O. A Chain of Events: Regulating Target Proteins by SUMO Polymers. *Trends in Biochemical Sciences* **46**, 113–123 (2021).
212. Hay, R. T. SUMO. *Molecular Cell* **18**, 1–12 (2005).
213. Subramonian, D., Chen, T.-A., Paolini, N. & Zhang, X.-D. "David". Poly-SUMO-2/3 chain modification of Nuf2 facilitates CENP-E kinetochore localization and chromosome congression during mitosis. *Cell Cycle* **20**, 855–873 (2021).
214. Yau, T.-Y., Sander, W., Eidson, C. & Courey, A. J. SUMO Interacting Motifs: Structure and Function. *Cells* **10**, 2825 (2021).
215. Vertegaal, A. C. O. Signalling mechanisms and cellular functions of SUMO. *Nat Rev Mol Cell Biol* **23**, 715–731 (2022).
216. Abed, M., Bitman-Lotan, E. & Orian, A. The Biology of SUMO-Targeted Ubiquitin Ligases in Drosophila Development, Immunity, and Cancer. *JDB* **6**, 2 (2018).
217. Dhingra, N. & Zhao, X. A guide for targeted SUMO removal. *Genes Dev.* **31**, 719–720 (2017).
218. Adorisio, S. *et al.* SUMO proteins: Guardians of immune system. *Journal of Autoimmunity* **84**, 21–28 (2017).
219. K., S. T. *et al.* SUMO and SUMOylation Pathway at the Forefront of Host Immune Response. *Front. Cell Dev. Biol.* **9**, 681057 (2021).
220. Yates, G., Srivastava, A. K. & Sadanandom, A. SUMO proteases: uncovering the roles of deSUMOylation in plants. *EXBOTJ* **67**, 2541–2548 (2016).
221. Morris, J. R. & Garvin, A. J. SUMO in the DNA Double-Stranded Break Response: Similarities, Differences, and Cooperation with Ubiquitin. *Journal of Molecular Biology* **429**, 3376–3387 (2017).
222. Lecona, E. *et al.* USP7 is a SUMO deubiquitinase essential for DNA replication. *Nat Struct Mol Biol* **23**, 270–277 (2016).

223. Hochstrasser, M. There's the Rub: a novel ubiquitin-like modification linked to cell cycle regulation. *Genes & Development* **12**, 901–907 (1998).
224. Gurevich, S. Z. *et al.* Rub1/NEDD8, a ubiquitin-like modifier, is also a ubiquitin modifier. 2020.06.18.159145 Preprint at https://doi.org/10.1101/2020.06.18.159145 (2020).
225. Pan, Z.-Q., Kentsis, A., Dias, D. C., Yamoah, K. & Wu, K. Nedd8 on cullin: building an expressway to protein destruction. *Oncogene* **23**, 1985–1997 (2004).
226. Willems, A. R., Schwab, M. & Tyers, M. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1695**, 133–170 (2004).
227. Schwechheimer, C. NEDD8 — its role in the regulation of Cullin-RING ligases. *Current Opinion in Plant Biology* **45**, 112–119 (2018).
228. Schulze-Niemand, E. & Naumann, M. The COP9 signalosome: A versatile regulatory hub of Cullin-RING ligases. *Trends in Biochemical Sciences* **48**, 82–95 (2023).
229. Zhao, Y., Morgan, M. A. & Sun, Y. Targeting Neddylation Pathways to Inactivate Cullin-RING Ligases for Anticancer Therapy. *Antioxidants & Redox Signaling* **21**, 2383–2400 (2014).
230. Watson, I. R., Irwin, M. S. & Ohh, M. NEDD8 Pathways in Cancer, Sine Quibus Non. *Cancer Cell* **19**, 168–176 (2011).
231. Hua, Z. & Vierstra, R. D. The Cullin-RING Ubiquitin-Protein Ligases. *Annu. Rev. Plant Biol.* **62**, 299–334 (2011).
232. Santonico, E. Old and New Concepts in Ubiquitin and NEDD8 Recognition. *Biomolecules* **10**, 566 (2020).
233. Whitby, F. G., Xia, G., Pickart, C. M. & Hill, C. P. Crystal Structure of the Human Ubiquitin-like Protein NEDD8 and Interactions with Ubiquitin Pathway Enzymes. *Journal of Biological Chemistry* **273**, 34983–34991 (1998).
234. Singh, R. K. *et al.* Recognition and Cleavage of Related to Ubiquitin 1 (Rub1) and Rub1-Ubiquitin Chains by Components of the Ubiquitin-Proteasome System. *Molecular & Cellular Proteomics* **11**, 1595–1611 (2012).
235. Meszka, I., Polanowska, J. & Xirodimas, D. P. Mixed in chains: NEDD8 polymers in the Protein Quality Control system. *Seminars in Cell & Developmental Biology* **132**, 27–37 (2022).
236. Vijayasimha, K. & Dolan, B. P. The Many Potential Fates of Non-Canonical Protein Substrates Subject to NEDDylation. *Cells* **10**, 2660 (2021).
237. Leidecker, O., Matic, I., Mahata, B., Pion, E. & Xirodimas, D. P. The ubiquitin E1 enzyme Ube1 mediates NEDD8 activation under diverse stress conditions. *Cell Cycle* **11**, 1142–1150 (2012).
238. Hjerpe, R. *et al.* Changes in the ratio of free NEDD8 to ubiquitin triggers NEDDylation by ubiquitin enzymes. *Biochemical Journal* **441**, 927–939 (2012).
239. Huang, D. T. & Schulman, B. A. Expression, Purification, and Characterization of the E1 for Human NEDD8, the Heterodimeric APPBP1–UBA3 Complex. in *Methods in Enzymology* vol. 398 9–20 (Academic Press, 2005).
240. Bohnsack, R. N. & Haas, A. L. Conservation in the Mechanism of Nedd8 Activation by the Human AppBp1-Uba3 Heterodimer. *Journal of Biological Chemistry* **278**, 26823–26830 (2003).

241. Walden, H. *et al.* The Structure of the APPBP1-UBA3-NEDD8-ATP Complex Reveals the Basis for Selective Ubiquitin-like Protein Activation by an E1. *Molecular Cell* **12**, 1427–1437 (2003).
242. Lois, L. M. & Lima, C. D. Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. *EMBO J* **24**, 439–451 (2005).
243. Lee, I. & Schindelin, H. Structural Insights into E1-Catalyzed Ubiquitin Activation and Transfer to Conjugating Enzymes. *Cell* **134**, 268–278 (2008).
244. Lv, Z., Williams, K. M., Yuan, L., Atkison, J. H. & Olsen, S. K. Crystal structure of a human ubiquitin E1–ubiquitin complex reveals conserved functional elements essential for activity. *Journal of Biological Chemistry* **293**, 18337–18352 (2018).
245. Lehmann, C., Begley, T. P. & Ealick, S. E. Structure of the *Escherichia coli* ThiS−ThiF Complex, a Key Component of the Sulfur Transfer System in Thiamin Biosynthesis ·. *Biochemistry* **45**, 11–19 (2006).
246. Lake, M. W., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB–MoaD complex. *Nature* **414**, 325–329 (2001).
247. Hann, Z. S. *et al.* Structural basis for adenylation and thioester bond formation in the ubiquitin E1. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 15475–15484 (2019).
248. Rajagopalan, K. V. Biosynthesis and processing of the molybdenum cofactors. *Biochemical Society Transactions* **25**, 757–761 (1997).
249. Leimkühler, S. & Iobbi-Nivol, C. Bacterial molybdoenzymes: old enzymes for new purposes. *FEMS Microbiology Reviews* **40**, 1–18 (2016).
250. Burroughs, A. M., Iyer, L. M. & Aravind, L. Natural history of the E1-like superfamily: implication for adenylation, sulfur transfer and ubiquitin conjugation. *Proteins* **75**, 895–910 (2009).
251. Burch, T. J. & Haas, A. L. Site-directed mutagenesis of ubiquitin. Differential roles for arginine in the interaction with ubiquitin-activating enzyme. *Biochemistry* **33**, 7300–7308 (1994).
252. Schulman, B. A. & Wade Harper, J. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol* **10**, 319–331 (2009).
253. Walden, H., Podgorski, M. S. & Schulman, B. A. Insights into the ubiquitin transfer cascade from the structure of the activating enzyme for NEDD8. *Nature* **422**, 330–334 (2003).
254. Yuan, L., Lv, Z., Adams, M. J. & Olsen, S. K. Crystal structures of an E1–E2–ubiquitin thioester mimetic reveal molecular mechanisms of transthioesterification. *Nat Commun* **12**, 2370 (2021).
255. Huang, D. T. *et al.* Basis for a ubiquitin-like protein thioester switch toggling E1-E2 affinity. *Nature* **445**, 394–398 (2007).
256. Huang, D. T. *et al.* Structural Basis for Recruitment of Ubc12 by an E2 Binding Domain in NEDD8's E1. *Molecular Cell* **17**, 341–350 (2005).
257. Liu, B., Lois, L. M. & Reverter, D. Structural analysis and evolution of specificity of the SUMO UFD E1-E2 interactions. *Sci Rep* **7**, 41998 (2017).

258. Haas, A. L., Bright, P. M. & Jackson, V. E. Functional diversity among putative E2 isozymes in the mechanism of ubiquitin-histone ligation. *J Biol Chem* **263**, 13268–13275 (1988).
259. Witze, E. S., Old, W. M., Resing, K. A. & Ahn, N. G. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods* **4**, 798–806 (2007).
260. Zhao, Y. & Jensen, O. N. Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* **9**, 4632–4641 (2009).
261. Deribe, Y. L., Pawson, T. & Dikic, I. Post-translational modifications in signal integration. *Nat Struct Mol Biol* **17**, 666–672 (2010).
262. Li, Q. *et al.* Global Post-Translational Modification Discovery. *J. Proteome Res.* **16**, 1383–1390 (2017).
263. Macek, B. *et al.* Protein post-translational modifications in bacteria. *Nat Rev Microbiol* **17**, 651–664 (2019).
264. Virág, D. *et al.* Current Trends in the Analysis of Post-translational Modifications. *Chromatographia* **83**, 1–10 (2020).
265. Seo, J.-W. & Lee, K.-J. Post-translational Modifications and Their Biological Functions: Proteomic Analysis and Systematic Approaches. *BMB Reports* **37**, 35–44 (2004).
266. Johnson, L. N. The regulation of protein phosphorylation. *Biochemical Society Transactions* **37**, 627–641 (2009).
267. Khoury, G. A., Baliban, R. C. & Floudas, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep* **1**, 90 (2011).
268. Nishi, H., Shaytan, A. & Panchenko, A. R. Physicochemical mechanisms of protein regulation by phosphorylation. *Front. Genet.* **5**, (2014).
269. Narita, T., Weinert, B. T. & Choudhary, C. Functions and mechanisms of non-histone protein acetylation. *Nat Rev Mol Cell Biol* **20**, 156–174 (2019).
270. Bilbrough, T., Piemontese, E. & Seitz, O. Dissecting the role of protein phosphorylation: a chemical biology toolbox. *Chem. Soc. Rev.* **51**, 5691–5730 (2022).
271. Ardito, F., Giuliani, M., Perrone, D., Troiano, G. & Muzio, L. L. The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *International Journal of Molecular Medicine* **40**, 271–280 (2017).
272. Krüger, R., Kübler, D., Pallissé, R., Burkovski, A. & Lehmann, W. D. Protein and Proteome Phosphorylation Stoichiometry Analysis by Element Mass Spectrometry. *Anal. Chem.* **78**, 1987–1994 (2006).
273. Rubin, C. S. & Rosen, O. M. Protein phosphorylation. *Annu Rev Biochem* **44**, 831–887 (1975).
274. Cohen, P. The origins of protein phosphorylation. *Nat Cell Biol* **4**, E127–E130 (2002).
275. Collins, M. O., Yu, L. & Choudhary, J. S. Analysis of protein phosphorylation on a proteome-scale. *PROTEOMICS* **7**, 2751–2768 (2007).
276. Cheng, H.-C., Qi, R. Z., Paudel, H. & Zhu, H.-J. Regulation and Function of Protein Kinases and Phosphatases. *Enzyme Res* **2011**, 794089 (2011).

277. Wang, Z. & Cole, P. A. Catalytic Mechanisms and Regulation of Protein Kinases. in *Methods in Enzymology* vol. 548 1–21 (Elsevier, 2014).
278. Allen, J. F. Protein phosphorylation in regulation of photosynthesis. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1098**, 275–335 (1992).
279. Kitchen, J., Saunders, R. E. & Warwicker, J. Charge environments around phosphorylation sites in proteins. *BMC Structural Biology* **8**, 19 (2008).
280. Johnson, L. N. & Barford, D. The effects of phosphorylation on the structure and function of proteins. *Annu Rev Biophys Biomol Struct* **22**, 199–232 (1993).
281. Chen, Z. & Cole, P. A. Synthetic approaches to protein phosphorylation. *Curr Opin Chem Biol* **28**, 115–122 (2015).
282. Pawson, T. & Scott, J. D. Protein phosphorylation in signaling – 50 years and counting. *Trends in Biochemical Sciences* **30**, 286–290 (2005).
283. Barford, D. Protein phosphatases. *Current Opinion in Structural Biology* **5**, 728–734 (1995).
284. Cohen, P. THE STRUCTURE AND REGIJLATION OF PROTEIN PHOSPHAT ASES.
285. Cohen, P. & Cohen, P. T. W. Protein phosphatases come of age. *Journal of Biological Chemistry* **264**, 21435–21438 (1989).
286. Ingebritsen, T. S. & Cohen, P. Protein Phosphatases: Properties and Role in Cellular Regulation. *Science* **221**, 331–338 (1983).
287. Gautschi, M. *et al.* The Yeast Nα-Acetyltransferase NatA Is Quantitatively Anchored to the Ribosome and Interacts with Nascent Polypeptides. *Mol Cell Biol* **23**, 7403–7414 (2003).
288. Polevoda, B. & Sherman, F. Composition and function of the eukaryotic N-terminal acetyltransferase subunits. *Biochemical and Biophysical Research Communications* **308**, 1–11 (2003).
289. Polevoda, B., Brown, S., Cardillo, T. S., Rigby, S. & Sherman, F. Yeast Nα-terminal acetyltransferases are associated with ribosomes. *Journal of Cellular Biochemistry* **103**, 492–508 (2008).
290. Arnesen, T. *et al.* Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 8157–8162 (2009).
291. Soppa, J. Protein Acetylation in Archaea, Bacteria, and Eukaryotes. *Archaea* **2010**, 1–9 (2010).
292. Bienvenut, W. V. *et al.* Comparative Large Scale Characterization of Plant versus Mammal Proteins Reveals Similar and Idiosyncratic N-α-Acetylation Features. *Molecular & Cellular Proteomics* **11**, M111.015131 (2012).
293. Liszczak, G. *et al.* Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. *Nat Struct Mol Biol* **20**, 1098–1105 (2013).
294. Forte, G. M. A., Pool, M. R. & Stirling, C. J. N-Terminal Acetylation Inhibits Protein Targeting to the Endoplasmic Reticulum. *PLOS Biology* **9**, e1001073 (2011).
295. Behnia, R., Panic, B., Whyte, J. R. C. & Munro, S. Targeting of the Arf-like GTPase Arl3p to the Golgi requires N-terminal acetylation and the membrane protein Sys1p. *Nat Cell Biol* **6**, 405–413 (2004).

296. Behnia, R., Barr, F. A., Flanagan, J. J., Barlowe, C. & Munro, S. The yeast orthologue of GRASP65 forms a complex with a coiled-coil protein that contributes to ER to Golgi traffic. *Journal of Cell Biology* **176**, 255–261 (2007).
297. Setty, S. R. G., Strochlic, T. I., Tong, A. H. Y., Boone, C. & Burd, C. G. Golgi targeting of ARF-like GTPase Arl3p requires its Na-acetylation and the integral membrane protein Sys1p. *Nature Cell Biology* **6**, 414–424 (2004).
298. Drazic, A., Myklebust, L. M., Ree, R. & Arnesen, T. The world of protein acetylation. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1864**, 1372–1401 (2016).
299. Quivy, V. & Van Lint, C. Regulation at multiple levels of NF-κB-mediated transactivation by protein acetylation. *Biochemical Pharmacology* **68**, 1221–1229 (2004).
300. Sharma, S. K. Protein acetylation in synaptic plasticity and memory. *Neuroscience & Biobehavioral Reviews* **34**, 1234–1240 (2010).
301. Shemorry, A., Hwang, C.-S. & Varshavsky, A. Control of Protein Quality and Stoichiometries by N-Terminal Acetylation and the N-End Rule Pathway. *Molecular Cell* **50**, 540–551 (2013).
302. Varshavsky, A. The N-end rule pathway and regulation by proteolysis. *Protein Science* **20**, 1298–1345 (2011).
303. Singh, B. N. *et al.* Nonhistone protein acetylation as cancer therapy targets. *Expert Review of Anticancer Therapy* **10**, 935–954 (2010).
304. Guan, K.-L. & Xiong, Y. Regulation of intermediary metabolism by protein acetylation. *Trends in Biochemical Sciences* **36**, 108–116 (2011).
305. Mischerikow, N. & Heck, A. J. R. Targeted large-scale analysis of protein acetylation. *PROTEOMICS* **11**, 571–589 (2011).
306. Anderson, K. A. & Hirschey, M. D. Mitochondrial protein acetylation regulates metabolism. *Essays in Biochemistry* **52**, 23–35 (2012).
307. Seto, E. & Yoshida, M. Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. *Cold Spring Harbor Perspectives in Biology* **6**, a018713–a018713 (2014).
308. Verdin, E. & Ott, M. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat Rev Mol Cell Biol* **16**, 258–264 (2015).
309. Christensen, D. G. *et al.* Mechanisms, Detection, and Relevance of Protein Acetylation in Prokaryotes. *mBio* **10**, e02708-18 (2019).
310. Diallo, I. *et al.* Current trends in protein acetylation analysis. *Expert Review of Proteomics* **16**, 139–159 (2019).
311. Marín-Hernández, Á., Rodríguez-Zavala, J. S., Jasso-Chávez, R., Saavedra, E. & Moreno-Sánchez, R. Protein acetylation effects on enzyme activity and metabolic pathway fluxes. *Journal of Cellular Biochemistry* **123**, 701–718 (2022).
312. Kamieniarz, K. & Schneider, R. Tools to Tackle Protein Acetylation. *Chemistry & Biology* **16**, 1027–1029 (2009).
313. Deng, S. & Marmorstein, R. Protein N-Terminal Acetylation: Structural Basis, Mechanism, Versatility, and Regulation. *Trends in Biochemical Sciences* **46**, 15–27 (2021).

314. Bu, B. *et al.* N-Terminal Acetylation Preserves α-Synuclein from Oligomerization by Blocking Intermolecular Hydrogen Bonds. *ACS Chem. Neurosci.* **8**, 2145–2151 (2017).
315. Ghanta, S., Grossmann, R. E. & Brenner, C. Mitochondrial protein acetylation as a cell-intrinsic, evolutionary driver of fat storage: Chemical and metabolic logic of acetyl-lysine modifications. *Critical Reviews in Biochemistry and Molecular Biology* **48**, 561–574 (2013).
316. Baldensperger, T. & Glomb, M. A. Pathways of Non-enzymatic Lysine Acylation. *Frontiers in Cell and Developmental Biology* **9**, (2021).
317. Baeza, J., Smallegan, M. J. & Denu, J. M. Site-Specific Reactivity of Nonenzymatic Lysine Acetylation. *ACS Chem. Biol.* **10**, 122–128 (2015).
318. Wolfe, A. J. Bacterial protein acetylation: new discoveries unanswered questions. *Curr Genet* **62**, 335–341 (2016).
319. Shi, L. & Tu, B. P. Protein acetylation as a means to regulate protein function in tune with metabolic state. *Biochemical Society Transactions* **42**, 1037–1042 (2014).
320. Farkaš, P. & Bystrický, S. Chemical conjugation of biomacromolecules: A mini-review. *Chemical Papers* **64**, (2010).
321. Harris, T. K. & Turner, G. J. Structural Basis of Perturbed pKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)* **53**, 85–98 (2002).
322. Isom, D. G., Castañeda, C. A., Cannon, B. R. & García-Moreno, B. Large shifts in pKa values of lysine residues buried inside a protein. *Proc Natl Acad Sci U S A* **108**, 5260–5265 (2011).
323. Simon, S. M. Role of organelle pH in tumor cell biology and drug resistance. *Drug Discovery Today* **4**, 32–38 (1999).
324. Matsuyama, S. & Reed, J. C. Mitochondria-dependent apoptosis and cellular pH regulation. *Cell Death Differ* **7**, 1155–1165 (2000).
325. Wu, M. M. *et al.* Organelle pH studies using targeted avidin and fluorescein–biotin. *Chemistry & Biology* **7**, 197–209 (2000).
326. Grabe, M. & Oster, G. Regulation of Organelle Acidity. *Journal of General Physiology* **117**, 329–344 (2001).
327. Demaurex, N. pH Homeostasis of Cellular Organelles. *Physiology* **17**, 1–5 (2002).
328. Weisz, O. A. Organelle Acidification and Disease. *Traffic* **4**, 57–64 (2003).
329. Chen, Y. & Arriaga, E. A. Individual Acidic Organelle pH Measurements by Capillary Electrophoresis. *Anal. Chem.* **78**, 820–826 (2006).
330. Casey, J. R., Grinstein, S. & Orlowski, J. Sensors and regulators of intracellular pH. *Nat Rev Mol Cell Biol* **11**, 50–61 (2010).
331. Santo-Domingo, J. & Demaurex, N. The renaissance of mitochondrial pH. *Journal of General Physiology* **139**, 415–423 (2012).
332. Cao, L. *et al.* In vivo observation of the pH alternation in mitochondria for various external stimuli. *Chemical Communications* **51**, 17324–17327 (2015).
333. Trougakos, I. P., Sesti, F., Tsakiri, E. & Gorgoulis, V. G. Non-enzymatic post-translational protein modifications and proteostasis network deregulation in carcinogenesis. *Journal of Proteomics* **92**, 274–298 (2013).

334. Khan, S. & Bhat, A. A. Chapter 10 - Nonenzymatic Posttranslational Protein Modifications: Mechanism and Associated Disease Pathologies. in *Protein Modificomics* (eds. Dar, T. A. & Singh, L. R.) 229–280 (Academic Press, 2019). doi:10.1016/B978-0-12-811913-6.00010-2.
335. Kazansky, Y., Lai, M.-Y., Singh, R. K. & Fushman, D. Impact of different ionization states of phosphorylated Serine-65 on ubiquitin structure and interactions. *Sci Rep* **8**, 2651 (2018).
336. Song, L. & Luo, Z.-Q. Post-translational regulation of ubiquitin signaling. *Journal of Cell Biology* **218**, 1776–1786 (2019).
337. Sharma, B. S. *et al.* Post-translational Modifications (PTMs), from a Cancer Perspective: An Overview. *Oncogen* **2**, (2019).
338. Lacoursiere, R. E. & Shaw, G. S. Acetylated Ubiquitin Modulates the Catalytic Activity of the E1 Enzyme Uba1. *Biochemistry* **60**, 1276–1285 (2021).
339. Zhou, H. *et al.* Toward a Comprehensive Characterization of a Human Cancer Cell Phosphoproteome. *J. Proteome Res.* **12**, 260–271 (2013).
340. Mertins, P. *et al.* Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534**, 55–62 (2016).
341. Kazlauskaite, A. *et al.* Parkin is activated by PINK1-dependent phosphorylation of ubiquitin at Ser65. *Biochemical Journal* **460**, 127–141 (2014).
342. Walser, F. *et al.* Ubiquitin Phosphorylation at Thr12 Modulates the DNA Damage Response. *Molecular Cell* **80**, 423-436.e9 (2020).
343. Swaney, D. L., Rodríguez-Mias, R. A. & Villén, J. Phosphorylation of ubiquitin at Ser65 affects its polymerization, targets, and proteome-wide turnover. *EMBO Rep* **16**, 1131–1144 (2015).
344. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Research* **26**, 399–422 (2016).
345. Ohtake, F. *et al.* Ubiquitin acetylation inhibits polyubiquitin chain elongation. *EMBO Rep* **16**, 192–201 (2015).
346. Kienle, S. M. *et al.* Electrostatic and steric effects underlie acetylation-induced changes in ubiquitin structure and function. *Nat Commun* **13**, 5435 (2022).
347. Lee, S. *et al.* Nonenzymatic acetylation of ubiquitin Lys side chains is modulated by their neighboring residues. *FEBS J* **285**, 1277–1289 (2018).
348. Kaur, S. & Saper, C. B. Neural Circuitry Underlying Waking Up to Hypercapnia. *Front Neurosci* **13**, 401 (2019).
349. Tsakiris, T. S., Konstantopoulos, A. I. & Bourdas, D. I. The Role of CO2 on Respiration and Metabolism During Hypercapnic and Normocapnic Recovery From Exercise. *Research Quarterly for Exercise and Sport* **92**, 537–548 (2021).
350. Cattano, C., Claudet, J., Domenici, P. & Milazzo, M. Living in a high CO_2 world: a global meta-analysis shows multiple trait-mediated fish responses to ocean acidification. *Ecol Monogr* **88**, 320–335 (2018).
351. Clements, J. C., Sundin, J., Clark, T. D. & Jutfelt, F. Meta-analysis reveals an extreme "decline effect" in the impacts of ocean acidification on fish behavior. *PLoS Biol* **20**, e3001511 (2022).
352. Heuer, R. M. & Grosell, M. Physiological impacts of elevated carbon dioxide and ocean acidification on fish. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **307**, R1061–R1084 (2014).

353. Thauer, R. K. Citric-acid cycle, 50 years on. *European Journal of Biochemistry* **176**, 497–508 (1988).
354. Curley, G., Laffey, J. G. & Kavanagh, B. P. Bench-to-bedside review: Carbon dioxide. *Critical Care* **14**, 220 (2010).
355. Amaral Silva, D. *et al.* Simulated, biorelevant, clinically relevant or physiologically relevant dissolution media: The hidden role of bicarbonate buffer. *European Journal of Pharmaceutics and Biopharmaceutics* **142**, 8–19 (2019).
356. D. Cassel & M. Rotman. REGULATION OF THE INTRACELLULAR pH IN THE PRESENCE AND ABSENCE OF BICARBONATE BUFFERS. *Journal of Basic and Clinical Physiology and Pharmacology* **1**, 393–398 (1990).
357. Lorimer, G. H. Carbon dioxide and carbamate formation: the makings of a biochemical control system. *Trends in Biochemical Sciences* **8**, 65–68 (1983).
358. Meldon, J. H. Blood-gas equilibria, kinetics and transport. *Chemical Engineering Science* **42**, 199–211 (1987).
359. Geers, C. & Gros, G. Carbon Dioxide Transport and Carbonic Anhydrase in Blood and Muscle. *Physiological Reviews* **80**, 681–715 (2000).
360. Arnone, A. Mechanism of Action of Hemoglobin. *Annu. Rev. Med.* **25**, 123–130 (1974).
361. Schneider, G., Lindqvist, Y. & Brändén, C. I. RUBISCO: structure and mechanism. *Annu Rev Biophys Biomol Struct* **21**, 119–143 (1992).
362. Spreitzer, R. J. & Salvucci, M. E. RUBISCO: Structure, Regulatory Interactions, and Possibilities for a Better Enzyme. *Annual Review of Plant Biology* **53**, 449–475 (2002).
363. Salvucci, M. E. & Crafts-Brandner, S. J. Inhibition of photosynthesis by heat stress: the activation state of Rubisco as a limiting factor in photosynthesis. *Physiologia Plantarum* **120**, 179–186 (2004).
364. Cleland, W. W., Andrews, T. J., Gutteridge, S., Hartman, F. C. & Lorimer, G. H. Mechanism of Rubisco: The Carbamate as General Base. *Chem. Rev.* **98**, 549–562 (1998).
365. Muradian, K. & Fraifeld, V. Hypercapnia-inducible factor: a hypothesis. *A&L* **2**, 27–31 (2021).
366. Shigemura, M., Welch, L. C. & Sznajder, J. I. Hypercapnia Regulates Gene Expression and Tissue Function. *Front. Physiol.* **11**, 598122 (2020).
367. Helenius, I. T. *et al.* Elevated CO2 suppresses specific Drosophila innate immune responses and resistance to bacterial infection. *Proc Natl Acad Sci U S A* **106**, 18710–18715 (2009).
368. Helenius, I. T. *et al.* Identification of *Drosophila* Zfh2 as a Mediator of Hypercapnic Immune Regulation by a Genome-Wide RNA Interference Screen. *The Journal of Immunology* **196**, 655–667 (2016).
369. Sharabi, K. *et al.* Elevated CO2 levels affect development, motility, and fertility and extend life span in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **106**, 4024–4029 (2009).
370. Zuela, N., Friedman, N., Zaslaver, A. & Gruenbaum, Y. Measuring the effects of high CO_2 levels in Caenorhabditis elegans. *Methods* **68**, 487–491 (2014).
371. Li, G. *et al.* Effect of carbon dioxide on neonatal mouse lung: a genomic approach. *J Appl Physiol (1985)* **101**, 1556–1564 (2006).

372. Bharat, A. *et al.* High CO_2 Levels Impair Lung Wound Healing. *Am J Respir Cell Mol Biol* **63**, 244–254 (2020).
373. Casalino-Matsuda, S. M., Nair, A., Beitel, G. J., Gates, K. L. & Sporn, P. H. S. Hypercapnia Inhibits Autophagy and Bacterial Killing in Human Macrophages by Increasing Expression of Bcl-2 and Bcl-xL. *The Journal of Immunology* **194**, 5388–5396 (2015).
374. Casalino-Matsuda, S. M. *et al.* Hypercapnia Alters Expression of Immune Response, Nucleosome Assembly and Lipid Metabolism Genes in Differentiated Human Bronchial Epithelial Cells. *Sci Rep* **8**, 13508 (2018).
375. Meigh, L. CO2 carbamylation of proteins as a mechanism in physiology. *Biochemical Society Transactions* **43**, 460–464 (2015).
376. Cummins, E. P., Strowitzki, M. J. & Taylor, C. T. Mechanisms and Consequences of Oxygen and Carbon Dioxide Sensing in Mammals. *Physiological Reviews* **100**, 463–488 (2020).
377. Cummins, E. P., Selfridge, A. C., Sporn, P. H., Sznajder, J. I. & Taylor, C. T. Carbon dioxide-sensing in organisms and its implications for human disease. *Cell. Mol. Life Sci.* **71**, 831–845 (2014).
378. Linthwaite, V. L. *et al.* The identification of carbon dioxide mediated protein post-translational modifications. *Nat Commun* **9**, 3092 (2018).
379. Kougentakis, C. M. *et al.* Anomalous Properties of Lys Residues Buried in the Hydrophobic Interior of a Protein Revealed with ^{15}N-Detect NMR Spectroscopy. *J. Phys. Chem. Lett.* **9**, 383–387 (2018).
380. Zhu, L., Kemple, M. D., Yuan, P. & Prendergast, F. G. N-Terminus and Lysine Side Chain pKa Values of Melittin in Aqueous Solutions and Micellar Dispersions Measured by 15N NMR. *Biochemistry* **34**, 13196–13202 (1995).
381. Webb, H. *et al.* Remeasuring HEWL pKa values by NMR spectroscopy: Methods, analysis, accuracy, and implications for theoretical pKa calculations. *Proteins: Structure, Function, and Bioinformatics* **79**, 685–702 (2011).
382. Blaum, B. S. *et al.* Lysine and Arginine Side Chains in Glycosaminoglycan−Protein Complexes Investigated by NMR, Cross-Linking, and Mass Spectrometry: A Case Study of the Factor H−Heparin Interaction. *J. Am. Chem. Soc.* **132**, 6374–6381 (2010).
383. Chiu, L.-Y., Hung, K.-W., Tjong, S.-C., Chiang, Y.-W. & Sue, S.-C. NMR characterization of the electrostatic interaction of the basic residues in HDGF and FGF2 during heparin binding. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1844**, 1851–1859 (2014).
384. André, I., Linse, S. & Mulder, F. A. A. Residue-Specific pKa Determination of Lysine and Arginine Side Chains by Indirect 15N and 13C NMR Spectroscopy: Application to apo Calmodulin. *J. Am. Chem. Soc.* **129**, 15805–15813 (2007).
385. Esadze, A., Li, D.-W., Wang, T., Brüschweiler, R. & Iwahara, J. Dynamics of Lysine Side-Chain Amino Groups in a Protein Studied by Heteronuclear ^{1}H– ^{15}N NMR Spectroscopy. *Journal of the American Chemical Society* **133**, 909–919 (2011).
386. Lee, C. K. & Manning, J. M. Kinetics of the Carbamylation of the Amino Groups of Sickle Cell Hemoglobin by Cyanate. *Journal of Biological Chemistry* **248**, 5861–5865 (1973).

387. Stark, G. R. Reactions of Cyanate with Functional Groups of Proteins. III. Reactions with Amino and Carboxyl Groups *. *Biochemistry* **4**, 1030–1036 (1965).
388. Piotrowski, J. *et al.* Inhibition of the 26 S Proteasome by Polyubiquitin Chains Synthesized to Have Defined Lengths. *Journal of Biological Chemistry* **272**, 23712–23721 (1997).
389. Haas, A. L. & Rose, I. A. The mechanism of ubiquitin activating enzyme. A kinetic and equilibrium analysis. *Journal of Biological Chemistry* **257**, 10329–10337 (1982).
390. Wee, K. E. *et al.* Steady-state kinetic analysis of human ubiquitin-activating enzyme (E1) using a fluorescently labeled ubiquitin substrate. *J Protein Chem* **19**, 489–498 (2000).
391. Singh, R. K., Kazansky, Y., Wathieu, D. & Fushman, D. Hydrophobic Patch of Ubiquitin is Important for its Optimal Activation by Ubiquitin Activating Enzyme E1. *Anal. Chem.* **89**, 7852–7860 (2017).
392. Amero, C. *et al.* Fast Two-Dimensional NMR Spectroscopy of High Molecular Weight Protein Assemblies. *J. Am. Chem. Soc.* **131**, 3448–3449 (2009).
393. Schanda, P., Kupče, Ē. & Brutscher, B. SOFAST-HMQC Experiments for Recording Two-dimensional Deteronuclear Correlation Spectra of Proteins within a Few Seconds. *J Biomol NMR* **33**, 199–211 (2005).
394. Schanda, P., Van Melckebeke, H. & Brutscher, B. Speeding Up Three-Dimensional Protein NMR Experiments to a Few Minutes. *J. Am. Chem. Soc.* **128**, 9042–9043 (2006).
395. Dixon, E. K., Castañeda, C. A., Kashyap, T. R., Wang, Y. & Fushman, D. Nonenzymatic assembly of branched polyubiquitin chains for structural and biochemical studies. *Bioorganic & Medicinal Chemistry* **21**, 3421–3429 (2013).
396. Lobato-Gil, S. *et al.* Proteome-wide identification of NEDD8 modification sites reveals distinct proteomes for canonical and atypical NEDDylation. *Cell Reports* **34**, 108635 (2021).
397. Vijayasimha, K., Tran, M. V., Leestemaker-Palmer, A. L. & Dolan, B. P. Direct Conjugation of NEDD8 to the N-Terminus of a Model Protein Can Induce Degradation. *Cells* **10**, 854 (2021).
398. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **160**, 65–73 (2003).
399. Prestegard, J. H., Mayer, K. L., Valafar, H. & Benison, G. C. Determination of Protein Backbone Structures from Residual Dipolar Couplings. *Methods Enzymol* **394**, 175–209 (2005).
400. Lipsitz, R. S. & Tjandra, N. Residual Dipolar Couplings in NMR Structure Analysis. *Annual Review of Biophysics and Biomolecular Structure* **33**, 387–413 (2004).
401. Fushman, D., Varadan, R., Assfalg, M. & Walker, O. Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar coupling measurements. *Progress in Nuclear Magnetic Resonance Spectroscopy* **44**, 189–214 (2004).

402. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **56**, 227–241 (2013).
403. Linge, J. P., Habeck, M., Rieping, W. & Nilges, M. ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **19**, 315–316 (2003).
404. de Vries, S. J., van Dijk, M. & Bonvin, A. M. J. J. The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* **5**, 883–897 (2010).
405. Chojnacki, M. *et al.* Polyubiquitin-Photoactivatable Crosslinking Reagents for Mapping Ubiquitin Interactome Identify Rpn1 as a Proteasome Ubiquitin-Associating Subunit. *Cell Chemical Biology* **24**, 443-457.e6 (2017).
406. Liang, J. *et al.* Chemical Synthesis of Diubiquitin-Based Photoaffinity Probes for Selectively Profiling Ubiquitin-Binding Proteins. *Angew Chem Int Ed Engl* **56**, 2744–2748 (2017).
407. Taupitz, K. F., Dörner, W. & Mootz, H. D. Covalent Capturing of Transient SUMO-SIM Interactions Using Unnatural Amino Acid Mutagenesis and Photocrosslinking. *Chemistry - A European Journal* **23**, 5978–5982 (2017).
408. Das, S. & Oliver, D. B. Mapping of the SecA·SecY and SecA·SecG Interfaces by Site-directed in Vivo Photocross-linking. *Journal of Biological Chemistry* **286**, 12371–12380 (2011).
409. Mori, H. & Ito, K. Different modes of SecY–SecA interactions revealed by site-directed in vivo photo-cross-linking. *Proceedings of the National Academy of Sciences* **103**, 16159–16164 (2006).
410. Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc Natl Acad Sci U S A* **99**, 11020–11024 (2002).
411. Chin, J. W. *et al.* Addition of p-Azido-l-phenylalanine to the Genetic Code of Escherichia coli. *J. Am. Chem. Soc.* **124**, 9026–9027 (2002).
412. Varadan, R., Assfalg, M., Raasi, S., Pickart, C. & Fushman, D. Structural Determinants for Selective Recognition of a Lys48-Linked Polyubiquitin Chain by a UBA Domain. *Molecular Cell* **18**, 687–698 (2005).
413. Beal, R., Deveraux, Q., Xia, G., Rechsteiner, M. & Pickart, C. Surface hydrophobic residues of multiubiquitin chains essential for proteolytic targeting. *Proc Natl Acad Sci U S A* **93**, 861–866 (1996).
414. Mueller, T. D., Kamionka, M. & Feigon, J. Specificity of the interaction between ubiquitin-associated domains and ubiquitin. *J Biol Chem* **279**, 11926–11936 (2004).
415. Ryu, K.-S. *et al.* Binding surface mapping of intra- and interdomain interactions among hHR23B, ubiquitin, and polyubiquitin binding site 2 of S5a. *J Biol Chem* **278**, 36621–36627 (2003).
416. Chojnacki, M. *et al.* Polyubiquitin-Photoactivatable Crosslinking Reagents for Mapping Ubiquitin Interactome Identify Rpn1 as a Proteasome Ubiquitin-Associating Subunit. *Cell Chemical Biology* **24**, 443–457 (2017).
417. Shi, Y. *et al.* Rpn1 provides adjacent receptor sites for substrate binding and deubiquitination by the proteasome. *Science* **351**, 10.1126/science.aad9421 aad9421 (2016).

418. Castaneda, C. *et al.* Nonenzymatic assembly of natural polyubiquitin chains of any linkage composition and isotopic labeling scheme. *J Am Chem Soc* **133**, 17855–17868 (2011).
419. Varadan, R., Walker, O., Pickart, C. & Fushman, D. Structural properties of polyubiquitin chains in solution. *Journal of Molecular Biology* **324**, 637–647 (2002).
420. Varadan, R. *et al.* Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. *The Journal of biological chemistry* **279**, 7055–7063 (2004).
421. Delaglio, F. *et al.* NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293 (1995).
422. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
423. Maciejewski, M. W. *et al.* NMRbox: A Resource for Biomolecular NMR Computation. *Biophysical Journal* **112**, 1529–1534 (2017).
424. Esadze, A., Zandarashvili, L. & Iwahara, J. Effective strategy to assign 1H-15N heteronuclear correlation NMR signals from lysine side-chain NH3 + groups of proteins at low temperature. *J Biomol NMR* **60**, 23–27 (2014).