ABSTRACT

Title of Dissertation: STRUCTURAL AND BIOPHYSICAL EXPLORATIONS OF PROTEIN DEGRADATION TAGS

Steven Michael Bonn

Doctor of Philosophy, 2023

Dissertation directed by: Professor David Fushman

Department of Chemistry and Biochemistry

Ubiquitin is a 76-amino acid, well-folded and highly stable protein that is highly conserved across all eukaryotes. It is a post-translational modifier of other proteins through its

attachment via an isopeptide bond to a substrate lysine sidechain. Multiple ubiquitin units can be stacked to form a polyubiquitin chain, with chain topologies and cellular outcomes varying based on which of ubiquitin's lysine residues they are attached to. The most common and well-studied outcome of polyubiquitin attachment is degradation by the proteasome, a massive barrel-shaped protease complex responsible for general protein quality control as well as cell cycle progression.

Proteasomes have been discovered in archaea and bacteria, and are controlled by the small archaeal modifier protein (SAMP) and the disordered prokaryotic ubiquitin-like protein (Pup), respectively. Recently, a second bacterial proteasome operon was discovered with a new putative signaling protein, ubiquitin bacterial (UBact).

Here, the first investigation of the UBact proteasomal operon is presented. Using nuclear magnetic resonance (NMR) spectroscopy and a variety of biophysical techniques, UBact is demonstrated to be disordered in solution and interact with its putative proteasomal receptor. This sets the groundwork for further studies of the UBact system.

Additionally, NMR is used to explore the activity and directionality of various deubiquitinase enzymes responsible for breaking down polyubiquitin chains, and for exploring small molecule binding to ubiquitin chains themselves. This likewise provides a groundwork for further studies of the ubiquitin system, whose dysregulation is responsible for many diseases and is an area of intense therapeutic development.

STRUCTURAL AND BIOPHYSICAL EXPLORATIONS OF PROTEIN DEGRADATION TAGS

by

Steven Michael Bonn

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 Jinwoo Lee

Professor Lai-Xi Wang

Assistant Professor Myles Poulin

Associate Professor Amy Karlsson, Dean's Representative

Acknowledgements

As with any project several years in the making, there are so many people to thank.

First, Dr. Fushman: thank you for welcoming me into your lab, for your guidance over the years, your suggestions, and your willingness to let me branch out and explore something new. You have been so patient with me as I worked through challenges, and you have guided me back to the right path more times than anyone could count. I will be forever thankful for your help in my growth into the scientist I am today.

To my committee, Dr. Lee, Dr. Wang, Dr. Poulin, and Dr. Karlsson: thank you for taking the time away from your own busy labs to help guide me. Your suggestions, advice, and support have been invaluable.

This work was supported by NIH grant GM065334 and NSF grant MCB1818280, both to David Fushman. NMR experiments were performed on instruments supported in part by NSF grant DBI1040158. Funds for the purchase of the instrument used in sedimentation measurements were provided by NIH grant S10 RR15899 to Dorothy Beckett.

To the University of Maryland Department of Chemistry and Biochemistry, especially Wafaa Hussein and the business office staff: thank you for your help in finding answers to so many questions, and thank you for the additional support through the Dean's Fellowship, Sampugna/Keeney Summer Fellowship, and Sampugna/Keeney Travel Award. Wafaa, thank you for keeping the shelves stocked.

To the past and present members of the Fushman lab: thank you so much for putting up with my questions, helping me talk through analysis or troubleshoot experiments, and for making long days and nights in the lab more fun. Christina, thank you for showing me the ropes of ubiquitin and biomolecular NMR. And especially to Wes, who finally decided to just walk into the woods rather than listen to one more question, thank you for being a great friend over the years and helping me with literally everything.

To Dr. Daniel Dries, Dr. Greta Loring, Dr. John Matter, Matt Bush, and Rich Watson: you were my first scientific mentors and have been supporters from afar. Your influence is felt every day, and I hope you are proud of the scientist you helped mold.

To my friends: thank you for your open ears and encouraging words.

To my family: thank you for your unconditional love and encouragement as I remained a professional student. Thank you for always being available for a supportive conversation or getaway. And thank you for helping keep the day to day in perspective.

Finally, to Loren and Isla: thank you for the endless love and support. Loren, you have never wavered in your support of this goal, and you have always been there on the good days and the bad. Thank you for taking on so much responsibility, even while in school yourself and working a real job to help real people while I moved very small amounts of liquid back and forth. Isla, thank you for always running to the door when I get home for a big hug, and for always reminding me what matters.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAA	ATPases associated with diverse cellular activities
ATP	Adenosine triphosphate
AUC	Analytical ultracentrifugation
CD	Circular Dichroism
CP	Core Particle
CryoEM	Cryogenic electron microscopy
CSI	Chemical shift indexing
CSP	Chemical shift perturbation
DMF	Dimethyl formamide
DMSO	Dimethyl sulfoxide
DUB	Deubiquitinating enzyme
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E2-25K	Ubiquitin-conjugating enzyme of 25 kDa molecular weight
E3	Ubiquitin-ligating enzyme
FAS	Fatty acid synthase
HEPES	4-(2-hydroxyethyl)-1-piperazineethansulfonic acid
HMQC	Heteronuclear multiple quantum coherence
HSQC	Heteronuclear single quantum coherence
IB	Inclusion body
IsoT	Isopeptidase T; Usp5

JAMM	JAB1/MPN/Mov34 metalloenzyme
MBP	Maltose binding protein
MINDY	Motif interacting with Ub-containing novel DUB family
MJD	Machado-Joseph disease/Josephin domain
Mpa	Mycobacteria proteasome activator
MPD	2-methyl-2,4-pentanediol
MTSL	*S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate
OTU	Ovarian tumor domain
polyUb	Polymeric ubiquitin
PAN	Proteasome-activating nucleotidase
PBS	Phosphate-buffered saline
PGP	Proteasomal Gate Protein
PMSF	Phenylmethylsulfonyl fluoride
PTM	Post-translational modification/modifier
Pup	Prokaryotic ubiquitin-like protein
SAMP	Small archaeal modifier protein
Ub	Ubiquitin
UBact	Ubiquitin bacterial
UBL	Ubiquitin-like protein/domain
UCH	Ubiquitin C-terminal hydrolase
USP	Ubiquitin-specific protease
Yuh1	Yeast ubiquitin C-terminal hydrolase 1

Chapter 1: Introduction

1.1: The Proteasome: Nature's Protein Recycling System

The production of new proteins and the recycling of old proteins is a vital part of intracellular homeostasis. Cells need a method to carry out controlled protein degradation to remove proteins that are no longer needed, clear away damaged or misfolded proteins, and recycle the amino acids from those proteins for future protein synthesis. Such a system must be both highly specific, to prevent random destruction of all proteins, as well as fairly promiscuous, to be able to degrade a variety of proteins.

The essential macromolecular machine responsible for balancing this need for selectivity and promiscuity in eukaryotes is the 26S proteasome, a massive 2 MDa protein complex. The 26S proteasome (named for its sedimentation coefficient) consists of two parts, the 20S core particle (CP) and the 19S regulatory particle (RP). The CP is the site of proteolytic cleavage, while the RP serves to specifically recognize substrates and unfold them (Figure 1)[1].

1.1a The 20S Core Particle

The CP is a 28-mer barrel-shaped complex made from four stacked heteroheptameric rings. The outer rings are made up of α-subunits 1-7, and the inner rings are made up of β-subunits 1-7 and form the proteolytic active site of the entire proteasome complex. Despite having seven theoretically possible active sites in each of the β-rings, only three such sites are functional in each, giving the eukaryotic CP

six total active sites. Wilk and Orlowski showed in 1983 that the CP exhibits three distinct proteolytic activities, chymotrypsin-like (cleavage on the C-terminal side of an aromatic residue), trypsin-like (cleavage on the C-terminal side of a basic residue), and caspase-like (cleavage on the C-terminal side of a glutamate residue) cleavages[2]. Each of these activities could be specifically inhibited without necessarily inhibiting the other activities, suggesting that the differing activities are not due to one highly

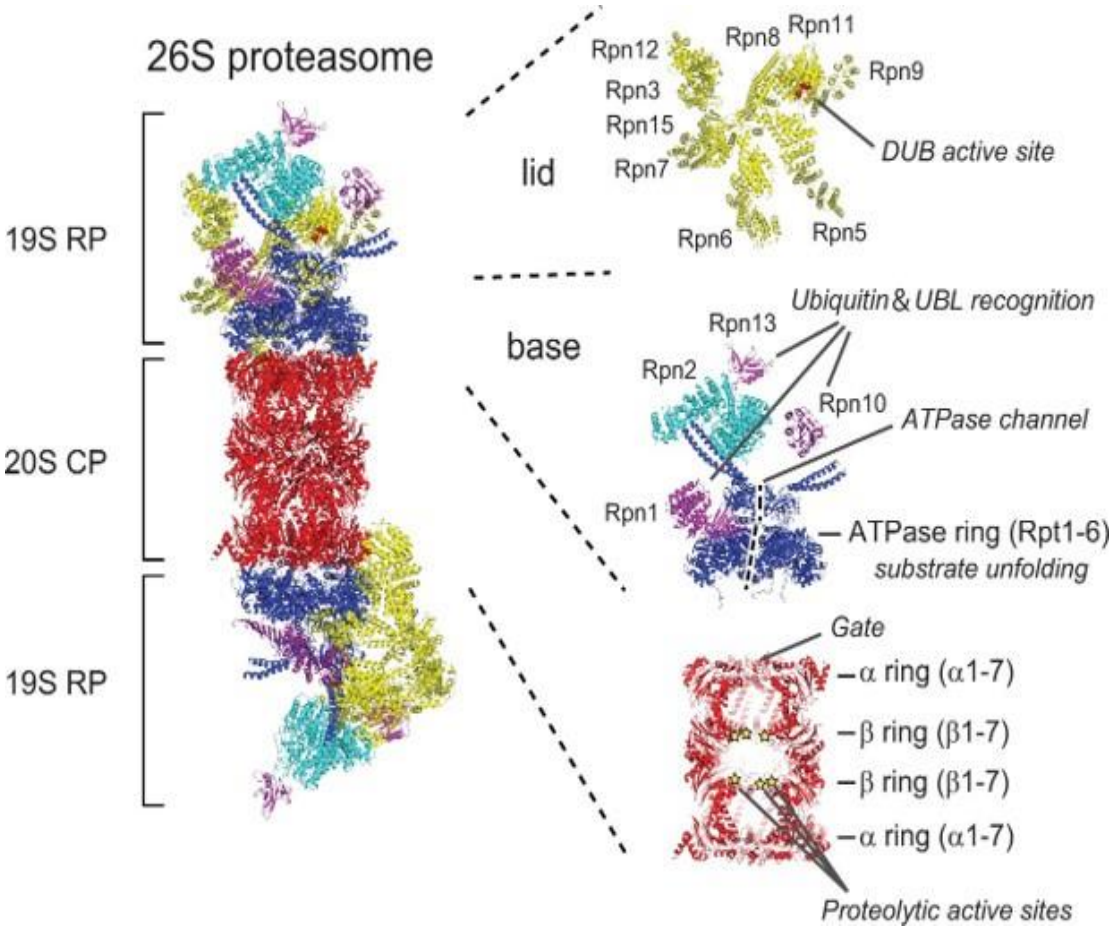

Figure 1-Structure of the 26S proteasome, subdivided into its component parts. The 19S RP docks to either end of the 20S CP, creating a symmetrical structural. Ub receptor proteins Rpn1, Rpn10, and Rpn13 facilitate recognition of ubiquitin-tagged substrate proteins, while Rpn11 is able to remove ubiquitin units. Removing Ub allows the Rpt proteins in the base of the RP to participate in ATP-dependent substrate unfolding and passing the unfolded substrate protein down a channel between the Rpt proteins and into the CP for degradation by the active sites in the β-subunits. Figure from Saeki *et al*[1]. Used with permission.

promiscuous active site[2]. It is worth noting that these are not strict specificities in peptide bond recognition but represent broad trends. Proteosomes have been shown to cleave after essentially every amino acid[3, 4]. The presence of these differential cleavage abilities satisfies the requirement of promiscuity, allowing for the degradation of many different proteins by the CP, and thus facilitating recycling of amino acids for reuse in new protein synthesis.

The outer α-subunits form a cap to the CP barrel. These form a gate structure that prevents nonspecific entry of substrates into the active site further down the barrel. This gate is formed from loops on the N-termini of several of the individual α-subunits[5], which stabilize a closed structure. The closed gate prevents folded proteins from randomly entering the proteolytic cavity and being degraded, but disordered or unfolded proteins are sometimes able to bypass the closed gate mechanism and enter the active site for degradation[6, 7], though the process is not very efficient. Mutagenic removal of these loops creates so-called "open gate" proteasomes[8], which can fairly efficiently degrade unfolded proteins. The gate formed by the α-subunits serves as just one part of the required specificity of the proteasome, with more specific recognition provided by the 19S RP.

1.1b The 19S Regulatory Particle

The 19S RP serves as the point of substrate recognition for tagged substrates (more on substrate tagging by ubiquitin in the next section). The RP itself is made of 19 subunits, divided into a lid and a base. The lid is made up of nine non-ATPase proteins, Rpn3, Rpn5-9, Rpn11, Rpn12, and Sem1. The lid primarily serves as the recognition site for proteins to be degraded. The base, made of the Rpt 1-6 AAA

ATPases[9], Rpn1, Rpn2, Rpn10, and Rpn13, serves as the immediate contact point to the CP, and facilitates substrate entry into the CP.

Currently, the lid proteins are not well characterized, though it is thought that it serves as a scaffolding system. Rpn11 is a metalloprotease responsible for removing ubiquitin (a deubiquitinase, or DUB) and polyUb moieties[10] that sits directly above a pore leading into the base of the RP[11]. On the base, Rpn1[12-14], Rpn10[15], and Rpn13[16] are receptors for ubiquitin and ubiquitin-like (UBL) proteins, and the Rpt proteins serve to assist in CP gate opening and pass substrates into the CP for final degradation[17].

1.2 Ubiquitin-Mediated Protein Degradation by the 26S Proteasome

1.2a Ubiquitin and Polyubiquitin

Ubiquitin (Ub) is a 76-amino acid, extraordinarily stable[18], highly structured β-grasp protein[19]. Ub is present in all eukaryotes, from yeast to humans, and is highly conserved across evolutionary space[20]. Most interactions involving Ub are facilitated by the highly conserved hydrophobic patch, L8/I44/V70[21], located across the β-sheets of the protein. Ub also features a highly flexible tail and ends with a C-terminal Gly-Gly motif.

Ubiquitin serves as a post-translational modifier (PTM) of other proteins. Ub is attached to a substrate lysine ε-amine by a three-enzyme cascade (more below) through its C-terminal glycine, forming an isopeptide bond in a process called

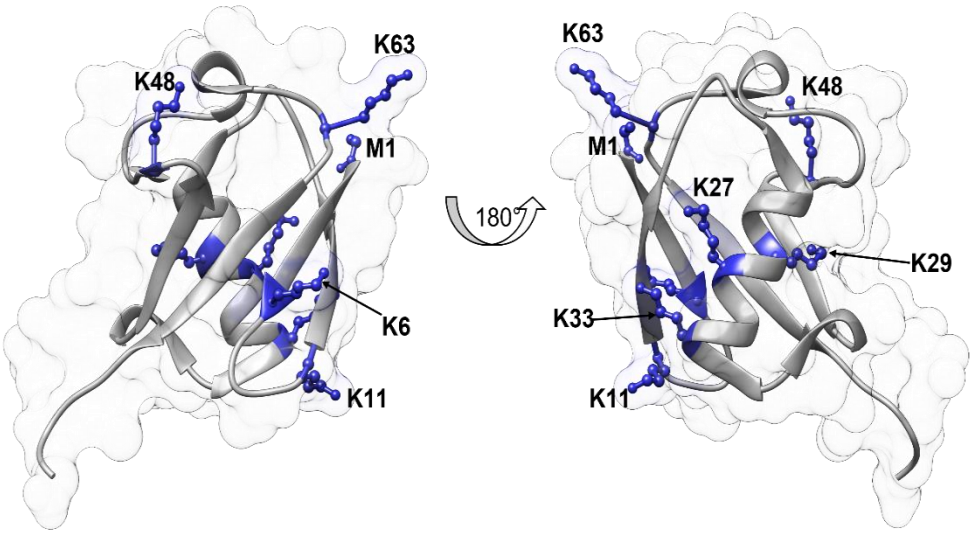

Figure 2-Structure of human ubiquitin (PDB 1D3Z), with semi-transparent surface shown. Residues available for conjugation are shown as blue ball and stick models. The highly flexible C-terminal tail extends out from the rest of the structure. All the conjugation-capable amine groups are solvent exposed, except that of K27, which is buried inside the core of the protein. Despite being buried, K27 is involved in chain formation.

ubiquitination. Ub itself has seven lysine residues (K6, K11, K27, K29, K33, K48, and K63) and its N-terminal amine and can therefore be ubiquitinated[22] (Figure 2)[23].

Ubiquitination of Ub forms polymeric Ub (polyUb) chains, with connections through the different lysine residues encoding for differential trafficking and signaling pathways. Different linkages and lengths produce chains of varying topology and dynamics, which helps explain the diverse interactions and outcomes involving ubiquitin and ubiquitin chains[24]. PolyUb chains of all linkages have been discovered, indicating physiological relevance for all of these chains[25]. Beyond simple homotypic polyUb linkages, different linkages can be utilized at the same time, resulting in mixed-linkage polyUb. Ubiquitin is also subject to a range of other

PTMs, such as acetylation[26], phosphorylation[27], and recently even carbamylation[28, 29], that can impact both receptor binding and protein dynamics.

1.2b Enzymatic Attachment of Ubiquitin to Substrates

As mentioned above, Ub is covalently attached to substrate lysine residues by a three-enzyme cascade. First, ubiquitin is activated by E1 (ubiquitin-activating enzyme), a large enzyme with two active sites that serves to prepare ubiquitin by forming a thioester bond with the C-terminal glycine carboxylic acid in an ATP-dependent manner. A single E1 enzyme is responsible for Ub activation. E1 then passes off the activated ubiquitin to an E2 (ubiquitin-conjugating enzyme), again through a thioester bond at the C-terminus of Ub. Finally, an E3 (ubiquitin-ligating enzyme) facilitates transfer of ubiquitin to the ε-amine of a substrate lysine sidechain. There are several dozen E2 enzymes (some E2s can function without an associated E3), and several hundred E3 enzymes. Different combinations of E2/E3 enzymes account for the huge diversity of tagged substrates, including the formation of the many polyUb chains. For example, K48-linked polyUb chains are formed first by activation of G76 of Ub with E1, then passing the activated Ub to E2-25K (an aptly named 25 kDa E2

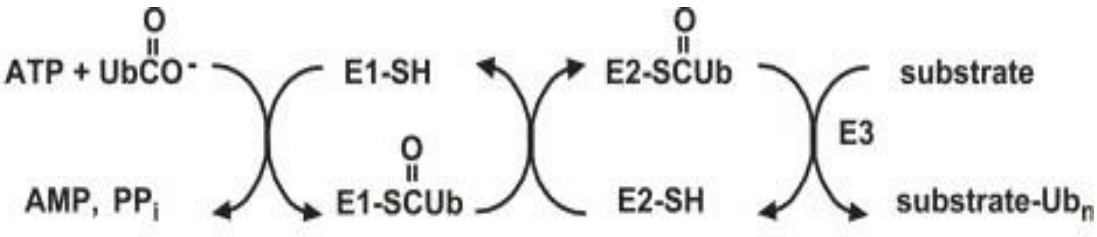

Figure 3-Diagram of ubiquitination by the E1/E2/E3 enzyme cascade. E1 first hydrolyzes ATP, adenylating Ub and releasing pyrophosphate. Adenylated Ub is then passed to the second active site and forms a C-terminal thioester with the active site cysteine on E1. Ub is then passed to E2 and forms another thioester, and a corresponding E3 coordinates selection and positioning of a substrate for isopeptide bond formation. Figure adapted from Pickart and Eddins[31]. Used with permission.

enzyme), which then facilitates the nucleophilic attack of a Ub K48 ε-amine on the E2-25K:Ub thioester, and the formation of the isopeptide bond[30](Figure 3)[31]. Following attachment of a single Ub to a substrate, more Ub moieties can be added to create a polyUb chain.

The most recently discovered polyUb chains are linear (M1-linked) ubiquitin chains, the only polyUb chains that do not contain an isopeptide bond between ubiquitin units; rather, they are attached via a regular peptide bond, with the ubiquitin units assembled head-to-tail. Though the *UBB* and *UBC* genes encode several head-to-tail ubiquitin units, these are co-translationally disassembled by the DUBs IsoT[32] and OTULIN[33] to give free ubiquitin. M1-linked polyubiquitin chains are assembled by an E3 complex called LUBAC (linear ubiquitin assembly complex), which both attaches the chain to substrates via an isopeptide bond and extends the chain via backbone peptide bonds[34]. M1-linked polyUb chains adopt an open conformation[35], and this flexibility imparts substrate specificity. M1-polyUb is involved in a variety of cellular processes, the most studied of which is interaction with NEMO (NF-κB Essential Modifier). M1-polyUb interaction with NEMO through its coiled-coil domain facilitates activation of the nearby IκB kinase complex, which in turn activates the NF-κB pathway[36]. Disruption of M1-polyUb chain synthesis and disassembly is associated with a number of pathologies[37], including disrupted embryonic development in mice[38], ectodermal dysplasia[39], increased susceptibility to bacterial infections[40], dermatitis[41, 42], and general inflammation[38].

K6-linked polyUb chains are assembled in solution by the E3 enzyme HUWE1[43], and appear to be involved in numerous signaling pathways. The chains

have been implicated in regulation of mitophagy, the processing of damaged mitochondria. The E3 enzyme Parkin auto-ubiquitinates itself and builds K6-polyUb chains, the formation of which slows down Parkin's recruitment to a damaged mitochondrion. Removal of the K6-polyUb moieties by Usp8 allowed for more efficient recruitment, and therefore autophagy[44]. There is evidence that K6-polyUb serves as a signal for protein degradation, as inhibition of the proteasome leads to accumulation of K6-polyuiquitinated substrates as detected by Western blot[43, 45]. Conversely, there are also scenarios in which ubiquitination with K6-linked chains stabilizes proteins[46] and might be used to out-compete attachment of other chains that might signal for substrate destruction[47]. These chains are also involved in DNA binding, insofar as their attachment to IFN regulatory factor 3 is required for DNA binding and subsequent gene induction[48]. The role of K6-polyUb signaling is a hot area of study, and new roles continue to be discovered.

K11-linked polyubiquitin chains are commonly involved in trafficking to the proteasome ending in substrate degradation[49, 50]. While K11-linked chains are sufficient for recruitment to the proteasome, branched K11/K48-linked chains are common, and enhance affinity for the proteasomal receptor Rpn1[51]. Additionally, K11-polyUb chains appear to be involved in innate immunity by assisting in autophagy regulation. Ligation of K11-polyUb to the MITA/STING protein prevents attachment of K48-linked chains, and therefore prevents its proteasomal degradation and therefore regulates type I IFN expression[52]. Type I IFN production is also abrogated upon K11-polyUb removal from Beclin-1[53]. Enhancement of K11-poyUb removal by upregulation of the DUB Cezanne is associated with increases levels of

hypoxia-induced factors 1α and 2α in several cancers, leading to studies about K11-poyUb applications in cancer research[54].

Lysine 27 is buried in the core of ubiquitin, and the CH_2 groups of the sidechain form significant hydrophobic contacts, and the ε-amine forms a stabilizing salt bridge with Q41 and D52, dramatically shaping the overall dynamics of the ubiquitin backbone[55]. Chains linked through K27 are involved in histone modification to facilitate DNA binding and signal for DNA damage[56] Formation and attachment of K27-linked polyUb chains has been implicated in resistance to viral infection. Attachment of K27-polyUb to NS4B of classical swine fever virus[57] and Nsp12 of porcine reproductive and respiratory syndrome virus[58] encoded for their destruction by the proteasome. K27-polyUb chains are also involved in IRF3 signaling, as conjugation with the chains promotes destruction of IRF3 in zebrafish[59] and humans[60]. K27-polyUb may also play a role in protection from neurodegenerative disease by helping (along with K29-linked polyUb) to encode for aggregation of leucine-rich repeat kinase 2[61] , DJ-1, and α-synuclein in Parkinson's Disease[62] and mutant huntingtin in Huntington's Disease[63].

K29-polyUb has been implicated in neurodegenerative diseases as described above. In addition, K29-polyUb decorates the E3 enzyme responsible for K27- and K29 ubiquitination of huntingtin protein[64]. K29-polyubiquitination is also involved in autophagy regulation by signaling for degradation of ULK1, which prevents the initiation of the phosphorylation pathway responsible for triggering autophagosome formation[65]. These chains are also attached to axin, which disrupts its interaction with LRP5/6, which in turn disrupts downstream phosphorylation events and triggers β-

catenin degradation, halting the Wnt signaling pathway[66] and giving rise to many cancer pathologies and neurodegenerative disease[67]. Interestingly, K29-polyUb is also utilized to inhibit the proteasome by conjugation to Rpn13, which precludes interaction with ubiquitinated proteins[68].

K33-linked chains are the least-well characterized polyubiquitin species. The chains are thought to generally be non-degrative, as the abundance of the chain does not increase upon proteasomal inhibition[69]. The chains are somehow involved in innate immune modulation. Removal of K33-polyUb chains from TBK1 in the interferon signaling pathway allows for degradation of TBK1 (following addition of K48-polyUb chains), and thus K33 chains serve to stabilize the protein[70]. Addition of K33-polyUb to the DNA-binding domain of STAT1 facilitated dissociation of STAT1 from DNA, thus blocking interferon-stimulated gene transcription[71], and addition to IRF3 disrupts interactions with nuclear transport receptors, thus suppressing subsequent gene transcription[72]. K33-polyUb linkages have been identified in autophagy pathways, alone and in concert with mixed-linkage chains, though their specific roles or importance has not been described[73, 74]. Additionally, attachment of K33-polyUb plays a role in degradation of leucine zipper-like transcriptional regulator 1 either by directly signaling for degradation or by enhancing attachment of K48-polyUb to the protein[75]. Many of these studies revealed K33-polyUb chains coexist with other linkage types either as mixed-linkage chains or as separate monotypic chains attached to the same substrate, though the relevance of this occurrence has not been described. Further studies are required to fully explore the

significance of the K33 linkage in conjunction with other linkages and as a standalone chain.

Polyubiquitin chains connected through lysine 48 are the most common type of linkage seen *in vivo*[25]. Labeling of substrates with K48-polyUb initiates transport to and degradation by the 26S proteasome[76], with a minimum of four ubiquitin units forming the basis of selectivity[77]. Proteasomal degradation is responsible for protein quality control in eukaryotic cells (reviewed extensively in Pohl and Dikic[78]) and in the mitrochondria[79]. Decoration with K48-polyUb as a degradation signal is utilized in immune signaling to both propagate signaling pathways and to shut them down[80]. K48-polyUb-mediated degradation of the Sic1 inhibitor protein is vital for advancement from G_1 to S phase in mitosis[81, 82]. Enhanced proteasome-mediated advancement of the cell cycle is implicated in a variety of cancers[83, 84]. Tau is monoubiquitinated as well as tagged with K48-polyUb[85], and post-translational modification of tau is associated with its potentially pathological liquid/liquid phase separation[86]. Additionally, aggregated proteins in neurons such as α-synuclein[87, 88] and huntingtin[89-91] are often ubiquitinated with K48-chains up to three units[92], below the threshold for efficient proteasomal degradation. Aggregation of these proteins additionally prevents their removal by the proteasome[93]. The wide variety of substrates of K48-polyubiquitination prove the importance of this linkage, and ongoing studies promise to reveal more substrates in diverse pathways.

The final type of linkage is through lysine 63. K63 chains are the second most common type of polyUb chain[22], and, together with K48-polyUb, are considered the "canonical" chains, with the bulk of research focusing on their physiological roles.

The most well-studied role of K63-polyUb signaling is in the context of DNA damage repair. Yeast with mutations in the Ubc13 E2, MMS2 E3, or ubiquitin (at position 63) are unable to properly repair UV-induced DNA damage[94]. These proteins are responsible for extending K63-linked chains, suggesting a role for the chains in DNA damage repair. Upon double stranded DNA breaks, due to radiation exposure, recombination processes[95, 96], and apoptosis[97], histone 2AX (H2AX) is phosphorylated at S139[98]. Phosphorylation allows RNF8, a ubiquitin ligase, to attach ubiquitin to H2AX. RNF8 recruits Ubc13 to extend K63-chains[99]. Rap80 interacts with these chains and facilitates the BRCA1-A complex formation on the polyUb chains[99], which in turn facilitates non-homologous end joining and homologous recombination[100]. Free-floating K63-polyUb chains might be involved in NF-κB signaling by binding to the TAK1 kinase and facilitating autophosphorylation and activation of TAK1[101].

1.2c Deubiquitinase Enzymes

In addition to Rpn11, the proteasomal deubiquitinase enzyme responsible for removing ubiquitin chains from substrates immediately prior to transport into and degradation by the proteasome, there are numerous deubiquitinase (DUB) enzymes. Some DUBs are linkage specific, interacting and cleaving only specific types of polyUb linkages, while others are nonspecific and will promiscuously cleave a variety of linkages. There are six families of DUB enzymes: ubiquitin C-terminal hydrolase (UCH), ubiquitin-specific protease (USP), ovarian tumor domain (OTU), MINDY (motif interacting with Ub-containing novel DUB family) and the Josephin-domain (MJD) families that encode papain-like cysteine proteases, while the

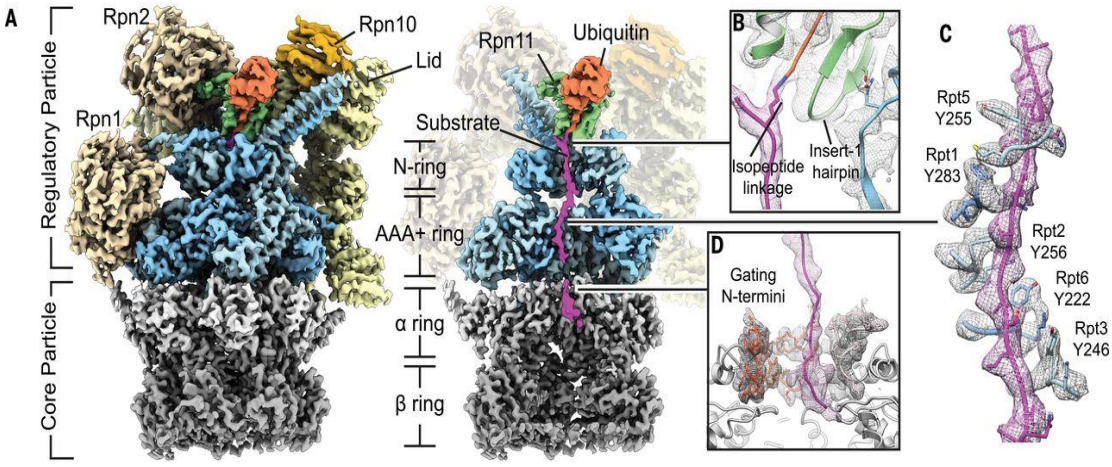

Figure 4-CryoEM reconstruction of substrate-engaged proteasome. A) Exterior and interior views, showing substrate (purple) extending from an isopeptide bond with ubiquitin down through the central pore into the top of the CP. B) The isopeptide bond is located inside the catalytic groove of Rpn11. C) The unfolded substrate (purple mesh) is stabilized by many tyrosine residues on the inside of the Rpt pore (gray mesh), permitting translocation through the pore and into the CP. D) Substrate entering the open gate CP. Figure from de la Peña, A. H.; Goodall, E. A.; Gates, S. N.; Lander, G. C.; Martin, A. Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-drive translocation. Science **2018**, 362 (6418).[107] Reprinted with permission from AAAS.

JAB1/MPN/Mov34 metalloenzyme (JAMM) domain encodes zinc-dependent metalloprotease. UCH DUBs are mostly involved in removing short peptide sequences from the C-terminus of ubiquitin, and are thought to be involved in processing of proubiquitin into physiologically active ubiquitin units[102]. USP family DUBs typically process larger leaving groups and contain a central catalytic domain with N-or C-terminal extensions containing regulatory elements that impart specificity for substrates. The USP family is the largest family of DUBs. OTU domain DUBs are so named based on their similarity to tumor genes in *Drosophila*. Humans encode 16 OTU domain proteins, though not all demonstrate DUB activity[103]. The MJD family contains four DUBs, which are not well studied currently. The most well-studied example is ataxin-3, which preferentially cleaves

K63-linked polyUb chains and interacts with a variety of cellular pathways[104]. MINDY family DUBs have extreme specificity for K48-linked chains[105]. The metalloprotease JAMM family of DUBs[106] include Rpn11, the proteasomal DUB that removes ubiquitin chains from substrates to be degraded[10].

1.2d Recognition of Ubiquitinated Substrates by the Proteasome

K48-linked polyUb is the most common polyUb species in the eukaryotic cell. Work from Cecile Pickart's lab involving carefully controlled polyUb chain synthesis revealed K48-linked tetraUb to be the minimum sufficient signal encoding substrate transport to and degradation by the proteasome[77]. Shorter chains did not efficiently bind to the proteasome, but longer chains appear to bind more tightly based on the number of tetraUb units in the chain. The polyUb units interact with the various receptors and are removed by Rpn11. At this point, the Rpt proteins hydrolyze ATP to provide energy for substrate unfolding and translocation down a central channel and into the CP[107] (Figure 4). The discovery and elucidation of the ubiquitin-proteasome system, and its characterization as essential for protein homeostasis, earned Aaron Ciechanover, Avram Hershko, and Irwin Rose the 2004 Nobel Prize in Chemistry.

1.3 A Simplified Proteasome System is Present in Archaea and some Bacteria

Once thought to be a hallmark of the cellular complexity of eukaryotic cells, the discovery of a large barrel-shaped protease complex in *Frankia*[108] challenged the paradigm that proteasomes could only exist in eukaryotes. Genomic analysis revealed a simplified proteasomal operon present in Archaea and some bacteria, notably

Actinomycetes and Nitrospirotae. Rather than seven different α- and β-subunits, non-eukaryotic proteasomes consisted of just a single form of the relevant subunit, repeated seven times to form a similar 28-mer CP[109]. Indeed, rather than the 19-mer RP of eukaryotes, these simpler proteasomes were capped by just an analogue of the Rpt proteins, a homohexameric protein termed *Mpa* (Mycobacterial proteasome activator) or *PAN* (proteasome-activating nucleotidase), which is likewise responsible for substrate recognition, unfolding, and transport into the CP[110]. Structural studies revealed that Mpa/PAN have similar structures to the Rpt proteins, forming N-terminal coiled-coils, a double stack of oligonucleotide-binding domains (as opposed to the single domain in Rpt proteins), and a C-terminal ATPase domain[111, 112].

Despite the presence of a proteasome in Archaea and some bacteria, as well as the documented hijacking of ubiquitin by bacteria[113], there is to date no evidence of ubiquitin in either kingdom's genome. Thus, it falls to some other system to build the recognition signal for interaction with the proteasome. In Archaea, this signal consists of the small Archaeal modifier protein (SAMP), which is also a β-grasp protein[114]. SAMP is activated by a single enzyme analogue to eukaryotic E1 (UbaA, ubiquitin-like activating protein of Archaea), and is ligated to substrate lysine residues by an isopeptide bond through the C-terminal glycine carboxylate, as in ubiquitin[115]. UbaA appears to be responsible for the entire activation/ligation process, rather than the multi-enzyme cascade involved in the ubiquitin system. SAMP was even capable of forming chains, though the function or physiological relevance of these chains remains unknown[115].

Actinomycetes and Nitrospirotae do not contain either ubiquitin or SAMP, however. Most studies performed on the bacterial proteasome have been performed in *Mycobacteria*, either in *M. tuberculosis* (the pathogenic bacteria that infects human lungs) or *M. smegmatis* (non-pathogenic model organism for *Mycobacteria*, with a fast growth rate relative to other species), as proteasome knockout mutants were found to dramatically decrease infectivity in a mouse model[116]. Genomic and mass spectrometric study of a known bacterial proteasome substrate revealed the identity of a small protein that was covalently attached via an isopeptide bond through its C-terminal residue to the substrate lysine sidechain[117]. This protein was named Pup (prokaryotic ubiquitin-like protein) due to its seemingly similar role. Knockouts of Pup, Mpa, or either proteasomal subunit yielded the same phenotype of reduced infectivity and lowered survival of *M. tuberculosis*. Rather than following the pattern of β-grasp fold proteins like ubiquitin or SAMP, Pup was found to be intrinsically disordered when free in solution and upon attachment to a substrate[118, 119]. Pup varies wildly in length from species to species, from as short as about 50 amino acids to almost 100 amino acids. Interestingly, Pup does not share the common Gly-Gly motif at its C-terminus, and instead ends with Gln or Glu, depending on the exact species. Crystallographic and NMR studies show that Pup interacts with Mpa at a region from the middle of the protein towards the C-terminus, forming an α-helix along the coiled-coil domain of Mpa, with the disordered N-terminus extending into the central pore of the Mpa hexamer[111]. The immediate C-terminus remains disordered and is solvent accessible. Despite the apparent existence of three identical binding sites (three coiled coil domains), it appears that only one Pup molecule is capable of

interaction at a time, as nonspecific electron density in the central Mpa pore, most likely corresponding to the disordered extreme N-terminus of Pup, would preclude a similar mode of binding at the other two equivalent sites[111].

Unlike in Archaea, no homolog of E1 could be found in *Mycobacteria*. Perhaps this is not surprising, since Pup does not share a similar structure to ubiquitin or SAMP, but a search for a similar enzyme responsible for ligation to a substrate did not yield immediate results. Pulldown and knockout studies identified PafA (proteasome accessory factor A) as a vital piece of the pupylation system, with PafA knockout cells showing no pupylated species[117]. Further characterization of PafA lead to the elucidation of its ligation mechanism[120, 121]. If Pup contains a C-terminal Gln, it is first deamidated by Dop (deamidase of Pup)[122]. Once deamidated, the resulting Glu sidechain carboxylic acid is phosphorylated by PafA, which then facilitates nucleophilic attack on the phosphoester by a substrate lysine residue and formation of the isopeptide bond *through the sidechain* of the C-terminus of Pup[120]. In addition to serving to activate Pup by deamidation, Dop also acted as a depupylase, able to remove Pup from substrates[123, 124]. Reconstitution of the Pup proteasome pathway in *E. coli*, which does not contain a natural proteasome system, showed that the presence of Pup, Dop, PafA, Mpa, and the proteasome genes were sufficient to initiate proteasomal degradation of several native *E. coli* proteins[125]. This suggests that there are intrinsic recognition signals in putative substrates, though no common sequence or structure motif has yet been identified, nor does there appear to be a common pathway or interaction beyond pupylation. It does not appear that Actinomycetes utilize their proteasome for general protein recycling.

Isopeptide linkage of Pup to substrates through its C-terminus immediately suggests a mechanism of proteasomal recognition and degradation. Following conjugation of Pup, the disordered Pup acts as a "fly fishing rod" and interacts with Mpa[126], whereupon it adopts the helical conformation, which pulls the substrate closer to the proteasomal complex. Mpa utilizes ATP hydrolysis to feed the disordered segment of Pup down through the pore and into the proteasome for degradation, as well as unfolding the covalently attached substrate and translocating that into the proteasome as well. This degrades both Pup and the substrate[127]. Degradation of Pup as part of the recycling process is in contrast with ubiquitin, which is generally removed and preserved by Usp14 and Rpn11 prior to substrate unfolding, translocation, and destruction in the CP.

1.4 The Putative UBact Proteasomal Operon

In 2017, a genomic search for novel proteasome genes revealed a potential new proteasome in the phylum Nitrospirota[128]. Ciechanover's group identified the presence of the Pup proteasomal operon in these bacteria, but they also discovered the presence of a second set of apparent PafA/Dop genes, as well as proteasome genes for α- and β-subunits. Previous studies showed that the Actinobacteria described above

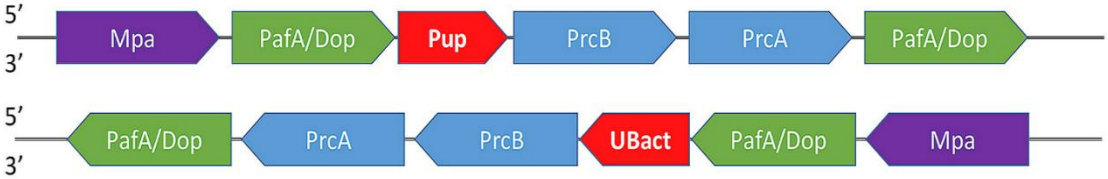

Figure 5-Arrangement of Pup (top) and putative UBact (bottom) operons in *Nitrospira* bacteria. The similar arrangement, albeit on opposite strands of the bacterial chromosome, allowed for the identification of the open reading frame the authors termed "UBact". PafA/Dop are difficult to differentiate purely from sequence, as the enzymes are extremely similar both by primary sequence and structure. Adapted from Lehmann *et al*[128]. Used with permission.

only contained one proteasome system. Exploration of this second putative proteasomal operon revealed an open reading frame that could potentially be conjugated to a substrate, based on its length, sequence, and genomic positioning in a similar arrangement to the Pup operon (Figure 5). Notably, this open reading frame, which the group called UBact (Ubiquitin bacterial) to differentiate it from Pup, does not share much sequence similarity with Pup (33% similar), but the other proteins in the operon are remarkably similar to Pup operon proteins. The proteasome genes themselves are similar to Pup operon proteasome genes, Archaeal proteasome subunits, and show general similarity to eukaryotic α- and β-subunits.

A search for other UBact genes yielded the presence of more proteasomal operons across other Gram-negative phyla as well, as well as candidate phyla and at least one Archaeal phylum[128]. Notably, several of these phyla also do not contain a corresponding Pup proteasomal operon. The appearance of the putative UBact proteasomal operon raises questions about the need for such a system, especially in organisms that already have a proteasome, as well as the evolutionary relationships between all proteasome systems.

1.5 Research Motivations

The occurrence of proteasomes across all domains of life highlights the importance of a controlled protein recycling system. Understanding the proteasomes and their associated signaling and trafficking pathways is vital toward designing functional therapeutics for a wide variety of conditions. As noted above, proteasomes are vital for full infectivity of *M. tuberculosis*[129]. Disrupting the Pup proteasome system has become an avenue for tuberculosis treatment[130-132]. Dysregulated

proteasomal degradation has been implicated in a variety of cancers, as the eukaryotic proteasome is required for proper cell division[133]. Disruption of the ubiquitin system, including altered proteasomal degradation, increased or decreased E2/E3 activity, and dysregulated DUB activity have all been implicated in disease states[134]. The ubiquitination system is also involved in the immune response, with cells utilizing the ubiquitin tagging system to target foreign proteins for proteasomal destruction, and interference with this system has been seen in many viral diseases[135, 136].

The wide usage of proteasomes across disparate phyla also raises the question of evolution. The similarity of the proteasomes themselves suggests a common ancestor that has changed little across evolutionary time, while the varied mechanisms for identifying and labeling substrates suggest divergent evolution for full utilization of the system. An exploration of the different proteasome systems, and their different signaling and recognition systems, can help scientists identify the evolutionary relationships and interplay of species across evolutionary time. To these ends, this work describes the first attempts at characterization of the UBact proteasome system. Additionally, NMR is used to study DUB kinetics and determine the directionality of ubiquitin unit removal from a polyUb chain, in order to further understand how these enzymes regulate post-translational modification with ubiquitin. Finally, NMR is used to characterize the interactions of small molecules with ubiquitin, as the first steps toward development of therapeutics that interact directly with ubiquitin itself.

Chapter 2: NMR Signal Assignment and Dynamics of *Nitrospira nitrosa* UBact

Chapter 2 was adapted from Bonn SM, Fushman D. Backbone NMR resonance assignment of the intrinsically disordered UBact protein from *Nitrospira nitrosa.* *Biomol NMR Assign* 2022; **16**: 129-134.[137]

2.1 Objective

Characterize UBact and complete an NMR signal assignment as a first step toward further interaction studies.

2.2 Expression and Purification of UBact

Escherichia coli-optimized *N. nitrosa* UBact (Uniprot accession number A0A256WXF2) DNA sequence was synthesized and cloned into the KpnI/XhoI restriction sites in the pET41a vector by Genscript (Piscataway, NJ, USA). This construct consists of an N-terminal GST tag, thrombin cleavage site, S-tag, enterokinase cleavage site, followed by the UBact gene with the amino acid sequence MIQSLMPERRERPGDPMPKSPSPLEEGGGPRRPETGSPDKDSLLKRMRRVDP KQAERYRQRTGE[128].

50 ng of this plasmid was transformed into chemically competent BL21 (DE3) *E. coli* cells by heat shock for 45 seconds at 42°C. Cells were then immediately cooled on ice for 2 min and rescued with SOC media for 1 hour at 37°C. Cells were grown on LB-agar plates supplemented with 50 µg/mL kanamycin for 16 hours at 37°C to select for successful transformation. Single colonies were picked to inoculate

5-20 mL cultures of LB supplemented with 50 µg/mL kanamycin, and were grown overnight at 37°C. Overnight growth was used to inoculate 1L LB supplemented with 50 µg/mL kanamycin and 1 mM $MgCl_2$, and the cultures were grown to OD_{600}=0.8, at which point protein expression was induced with the addition of 1 mM isopropyl-β-d-1-thiogalactopyranoside (IPTG) at 18 °C for at least 16 hours for maximal yield.

Cells were harvested by centrifugation at 4680 x *g* for at least 15 minutes. Cells were resuspended in phosphate buffered saline (PBS; 10 mM sodium phosphate dibasic, 1.8 mM potassium phosphate monobasic, 2.7 mM potassium chloride, 137 mM sodium chloride, pH 7.4) supplemented with 5 mM $MgCl_2$, DNase, 0.4 mg/mL lysozyme, 1% Triton X-100, and 1 mM phenylmethylsulfonyl fluoride (PMSF) as a general protease inhibitor. Cells were lysed by flash freezing in liquid nitrogen and rapid thawing at 42°C, followed by sonication for 3 minutes at output 5, duty 60% on ice. Lysate was clarified by ultracentrifugation at 48,900 x *g* for 30 minutes. Lysate supernatant was filtered through a 0.22 µm membrane, then applied to PBS-equilibrated glutathione agarose beads in a gravity column at 4°C. Beads were washed extensively with PBS, until flowthrough ran clear (usually 15-20 bead volumes, performed in 3 wash steps). Bound GST-UBact fusion protein was eluted from the beads with ~30 mL 50 mM tris pH 8.0, 10 mM glutathione. Eluted GST-UBact was concentrated, and buffer exchanged into 50 mM tris pH 8.0 in a 10 kDa molecular weight cutoff centrifugal filter unit to a volume of 0.5-1 mL.

Concentrated GST-UBact was incubated with 16-32 U bovine enterokinase (New England Biolabs, Ipswitch, MA, USA) for 4-12 hours at 18°C to cleave UBact from the fusion protein. Following incubation, reaction solution was diluted to ~4 mL

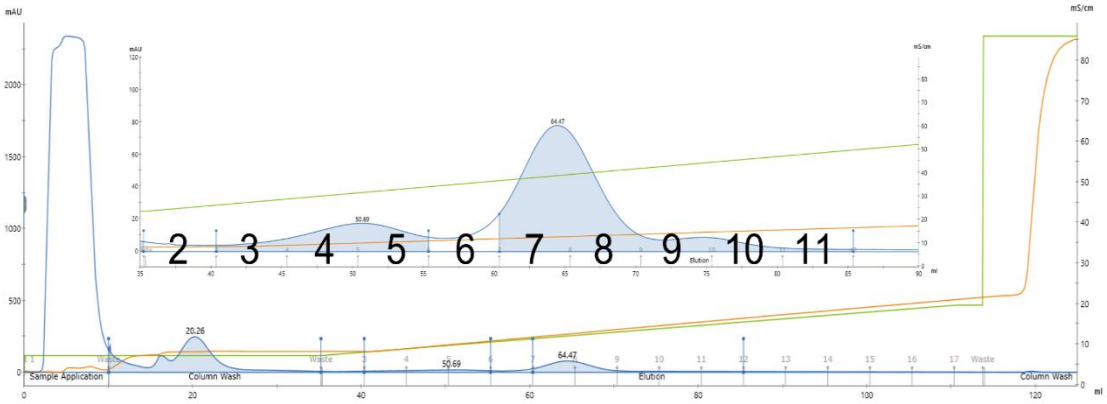

Figure 6-Chromatogram of UBact M17W cation purification. Absorbance at 280 nm (mAbs) is shown in blue. Conductivity (mS/cm) is shown in orange. Gradient buffer percentage is shown in green. Flowthrough/wash contains the N-terminal GST fragment removed by enterokinase. The small peak in fractions 4-5 is a truncated UBact form as well as a very small amount of GST fragment that sticks to the resin. Fractions 7-8 contain full length UBact. Note that this chromatogram shows the M17W mutant, allowing for easy tracking of purification success. WT UBact does not display any absorbance at 280 nm, despite a tyrosine at position 58.

with 20 mM sodium phosphate pH 8.0 and loaded onto an SPHP cation exchange column (GE Healthcare), equilibrated in the same buffer. The column was washed until absorbance dropped, and bound protein was eluted with a gradient to 200 mM NaCl. Highly pure UBact eluted between 100-150 mM NaCl (Figure 6). Purified UBact was concentrated, and buffer exchanged into 20 mM sodium phosphate pH 6.8 (for backbone assignment) or 20 mM sodium phosphate pH 6.8, 130 mM NaCl (for binding studies) in a 3 kDa molecular weight cutoff centrifugal filter unit. UBact is stable at 23°C for up to 5 days, at 4°C for up to a month, and for up to three freeze/thaw cycles from -80°C. UBact was typically stored at 1-10 mM concentration, depending on intended usage. UBact did not appear to be sensitive to added dimethyl sulfoxide (DMSO) up to 50%, dimethyl formamide up to 50%, or methanol up to 10%, though no longer-term stability testing was done.

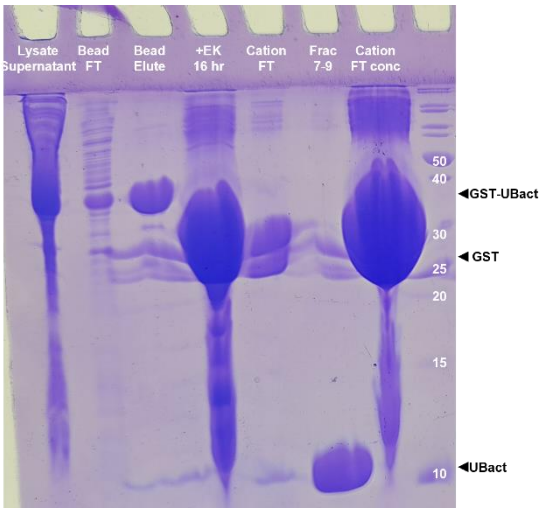

Figure 7- SDS-PAGE of UBact expression and purification. Glutathione bead purification yields highly pure GST-UBact, and 12–16-hour incubation with enterokinase is generally sufficient for complete cleavage. A small amount of UBact sometimes flows through the cation column. Cation purification yields highly pure UBact. UBact runs slightly higher than its calculated molecular weight of 7.6 kDa. This is common for IDPs, as enrichment in charged residues can disrupt SDS binding, which lowers migration through the gel.

Purity was analyzed by SDS-PAGE (Figure 7)[138] and ESI mass spectrometry.

1L of cell culture typically yielded ~3-5 mg purified UBact.

Point mutations M17W, G63C, and E64Q were introduced using site directed mutagenesis. UBact mutants were purified identically as above, except that 1 mM tris(2-carboxyethyl)phosphine (TCEP) was added to cation exchange buffers and 10 mM TCEP was added to storage buffers when working with cysteine-containing

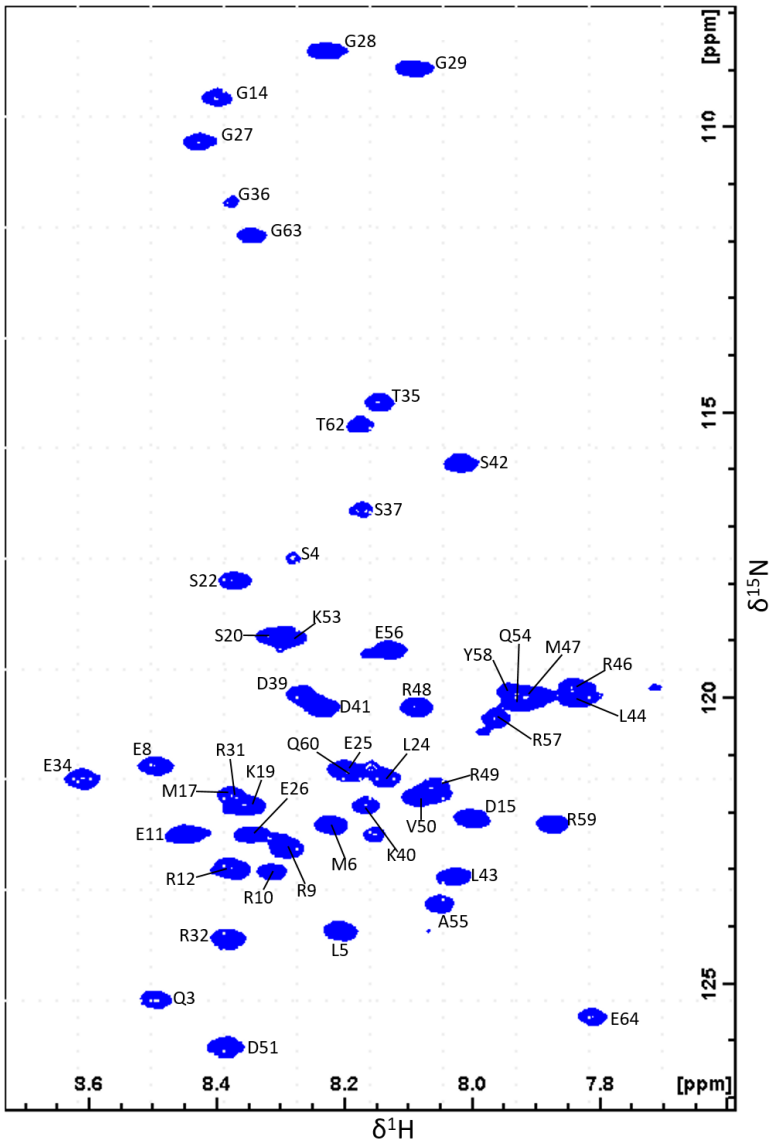

Figure 8- Assigned ^{1}H-^{15}N HSQC spectrum of UBact. Signals show a narrow dispersion in the proton dimension, a hallmark of a disordered protein. I2 does not show a resonance, possibly due to being in fast exchange with water. The L45 resonance is buried between L44 and R46 and could not be resolved. R61 could not be definitively identified. Adapted from Bonn and Fushman[137]. Reproduced with permission from Springer Nature.

UBact. M17W was incorporated to aid in concentration determination, as the lone tyrosine at position 58 did not exhibit any absorbance at any concentration tested. Lowry[139] and Bradford[140] dye-binding assays did not display a linear response at any concentrations tested. Position 17 was chosen as it is between two proline residues, and therefore might be conformationally locked such that mutation of the sidechain does not affect surrounding residues much.

Isotope-enriched UBact species were produced identically to above, except that the large LB culture was replaced by 1L M9 minimal media supplemented with either 1 g/L $^{15}NH_4Cl$ and 4 g/L glucose for ^{15}N-enriched UBact species or 1 g/L $^{15}NH_4Cl$ and 2 g/L ^{13}C-glucose for $^{15}N/^{13}C$ enriched UBact species. Yields typically decreased 50% or greater per isotope added. 1L of $^{15}N/^{13}C$ enriched growth yielded approximately 1 mg purified $^{15}N/^{13}C$ UBact.

2.3 Backbone Signal Assignment of UBact

The HSQC spectrum of a ^{15}N UBact sample (Figure 8) shows narrow dispersion in the proton dimension, which is often indicative of a disordered protein, or a protein consisting of non-interacting α-helices. The narrow spread of signals in the proton dimension prevented easy signal assignment using a TOCSY/NOESY walk, as too many signals overlapped to unambiguously assign connectivity.

The backbone signal assignment of UBact was completed using a series of 2D and 3D heteronuclear NMR experiments to determine both sequential linkages between residues and to identify the specific type of residue represented by each signal. ^{15}N-edited TOCSY and NOESY, HNCO, HN(CA)CO, HNCA, HN(CO)CA, and HNCACB experiments were conducted on a 500 µM ^{15}N/^{13}C UBact sample.

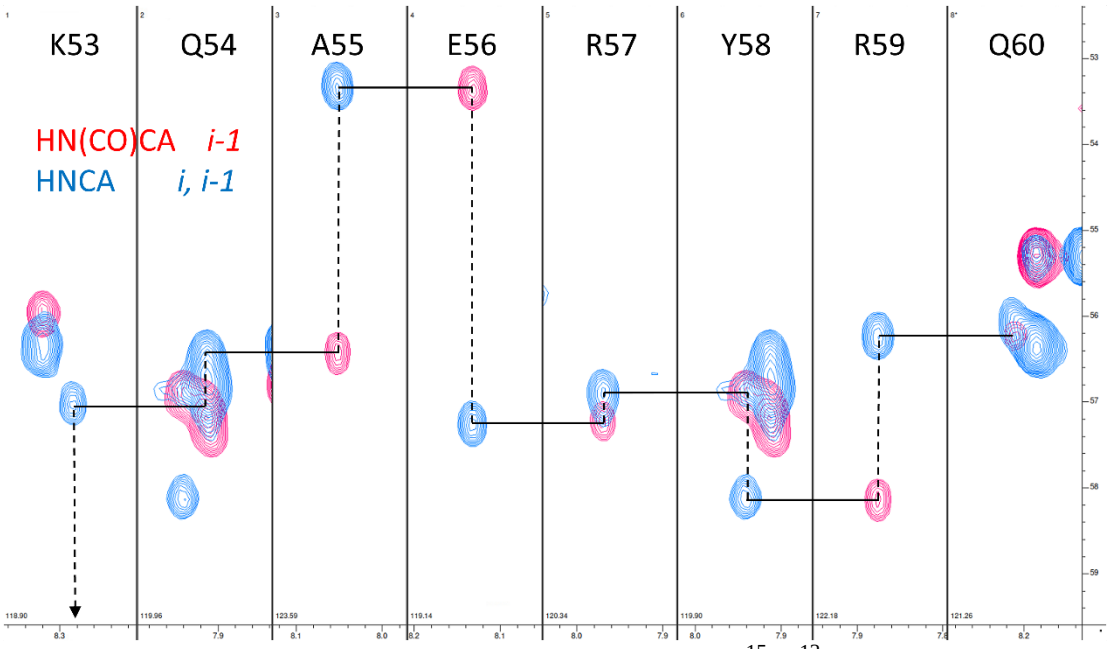

Figure 9-Example of the *i* and *i-1* connectivities derived from the HNCA/HN(CO)CA experiment pair. HN(CO)CA is shown in red and shows only signals corresponding to the Cα from the preceding (i-1) residue. The HNCA shows signals from both the self (i) and preceding (i-1) residue. In the case of the resonances shown above, the HNCA i-1 signals were significantly weaker in intensity, allowing for contours to be manipulated to remove ambiguities and signal overlaps. Strips show the same frequency scale in the carbon dimension, at different frequencies in proton and nitrogen dimensions.

Each experiment pair provides resonance for the *i* (self) residue, as well as the *i-1* (preceding) residue (Figure 9). The HNCO/HN(CA)CO experiments were most useful in helping to establish connectivity between residues. The HNCA/HN(CO)CA experiments were less helpful, as there is significant sequence redundancy, and many C_α signals overlapped, or connections were unclear. The addition of the HNCACB

clarified almost all of the connection ambiguities, as most amino acids show a characteristic combination of C_α and C_β signals. The HNCACB experiment is also able to discriminate between *i* and *i-1* signals through differential intensities, as the transfer to the *i-1* residue is less efficient, leading to significantly reduced signal intensity. Combining the characteristic C_α and C_β resonances with the TOCSY-derived number of connected protons yielded unambiguous identification of the type of residue, and all experiments together yielded multiple levels of confirmation of connectivity.

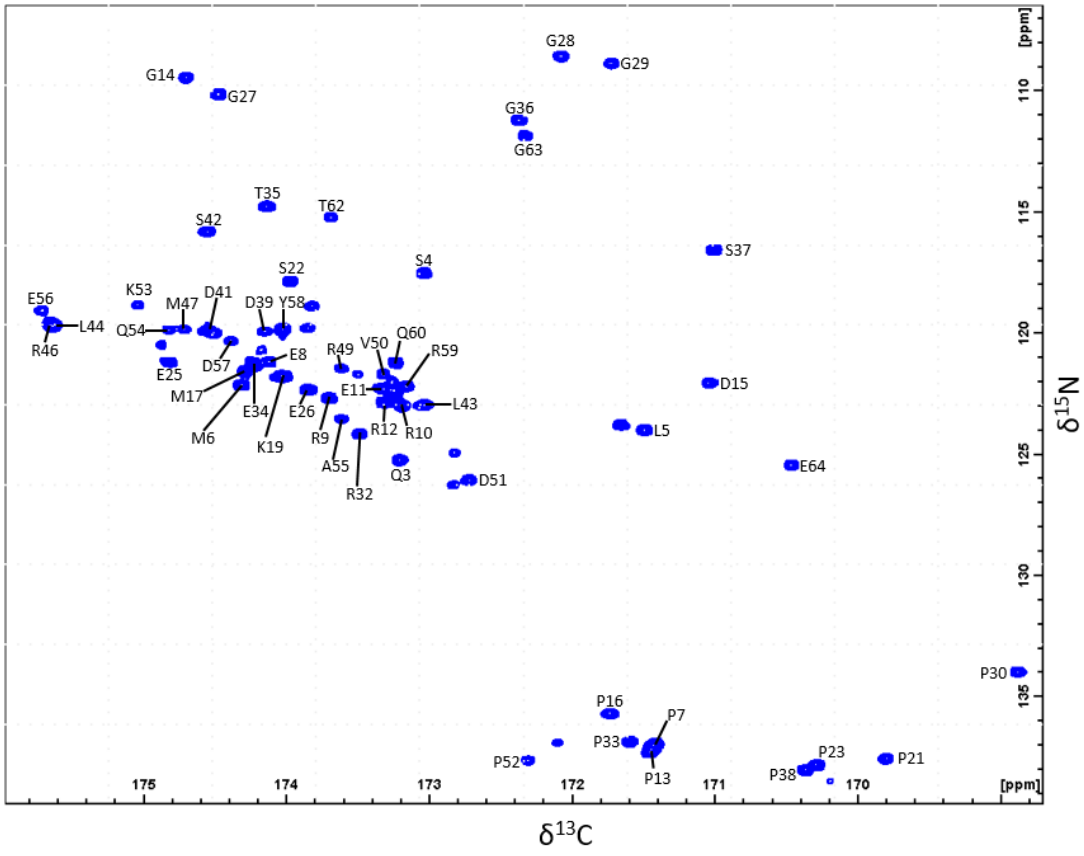

Figure 10-Assigned CON spectrum of UBact. A major advantage of the CON experiment is the appearance of proline residues in the bottom right of the spectrum. Unannotated peaks showed no cross peaks in any triple resonance experiments and could be degradation products due to the long experiment time required. Adapted from Bonn and Fushman[137]. Reproduced with permission from Springer Nature.

In addition to assigning the HSQC spectrum signals, which correspond only to backbone amides, the triple resonance experiments allowed for resonance assignment of the C_α and C' backbone resonances, as well as many of the C_β resonances. In the end, 397 resonances were assigned, including 94% of backbone resonances, and 45% of residues had at least partial sidechain proton assignments. Assignments for K45 and R61 were not possible, as they overlapped too severely with other signals.

Since the HSQC spectrum is complicated by crowded signals, a carbon direct detected CON spectrum was acquired. This spectrum dramatically increases

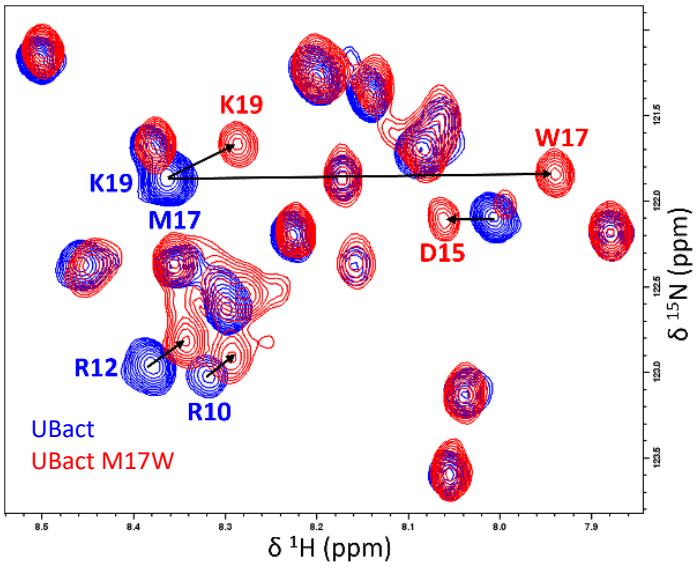

Figure 11- Incorporation of the W17 mutation does not dramatically alter the UBact spectrum. There is a large shift in the signal corresponding to residue 17, which is to be expected upon mutation. Other shifts are small and in residues immediately near to the mutation site.

resolution of the signals, at the cost of significant experimental time (12-24 hours for a CON versus 2 hours for a HSQC or 20 minutes for a SOFAST HMQC). Since the CON avoids starting with proton magnetization, it has the bonus of providing signals for proline residues, which can only be detected in proton-start experiments indirectly. The fully assigned CON spectrum of UBact is shown in Figure 10. In

theory, a full suite of carbon direct-detected experiments for signal assignment as well as relaxation and dynamics measurements exists. The additional experimental time required was determined to be unneccessary for further studies of UBact, especially in light of UBact's short stability window on the scale of multi-day NMR experiments, the low yield and high monetary costs associated with addition of multiple isotopes.

Incorporation of the M17W mutation only led to the disappearance of the signal corresponding to M17, and the appearance of a new signal corresponding to the backbone amide of W17 (Figure 11) and two sidechain indole resonances. The two indole signals most likely result from conformational exchange from proline isomerization.

As part of the characterization of UBact, ^{15}N R_1 longitudinal and R_2 transverse relaxation rates were determined at 600 and 800 MHz (Figure 12 and Appendix Table 1 and 2). R_1 describes the rate for the *Z*-magnetization to return to thermodynamic equilibrium, or the rate at which the high energy state loses energy to the environment. R_2 describes the rate at which magnetization vectors in the perpendicular plane following excitation dephase, or the rate at which spins tansfer energy to each other. These relaxation processes are governed by the local magnetic envrinoments, and thus can provide information about how a particular residue interacts with surrounding residues. The relaxation rates are generally lower in the N-terminal half of the protein, and slightly elevated in the C-terminal half. Both N- and

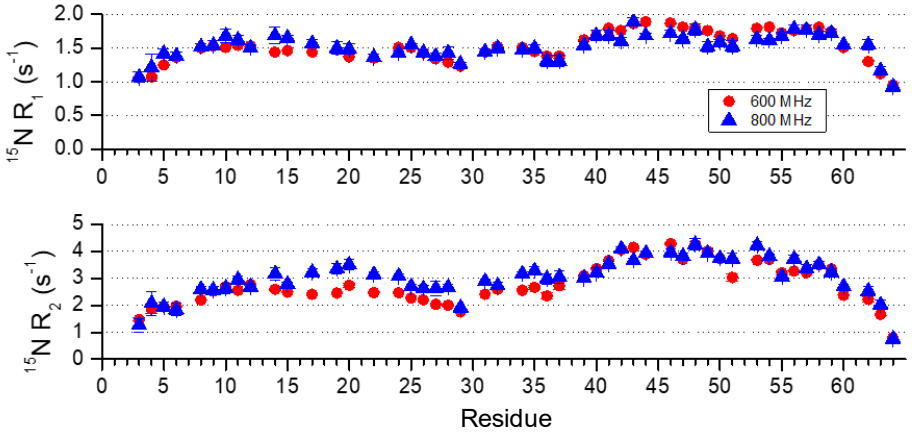

Figure 12-^{15}N R_1 (top) and R_2 (bottom) relaxation rates. The C-terminal half of the protein exhibits slightly elevated relaxation rates relative to the N-terminal half. Each of the N- and C-termini themselves display exceedingly slow relaxation, which is a hallmark of high flexibility. Errors were calculated from duplicate measurements at each relaxation delay.

C-termini show decreased relaxation rates relative to the rest of the protein, which is to be expected if the termini are flexible. The relaxation rates follow roughly the same trend between experiments and at differing field strengths. In isolation, relaxation rates do not mean much, but thy can be helpful in predicted rotational correlation times. However, most analyses rely on the assumption of a well-folded protein with a single overall correlation time, an assumption that is not satisfied in this case.

2.4 Exploring the Potential for Secondary Structure in UBact

2.4a Circular Dichroism

The narrow spread of amide signals in the HSQC spectrum are often a hallmark of disordered proteins. Since UBact is predicted to be disordered, a circular dichroism (CD) spectrum of UBact was acquired to try to identify any evidence of secondary structure (Figure 13). The CD spectrum shows no dips at 222 nm or 208 nm, which are evidence of the presence of α-helical secondary structure. The spectrum also does not show the positive elipticity around 200 nm, a hallmark of β-sheet secondary structure. Rather, the spectrum shows a strong negative band around 200 nm, which is typical of protein random coil.

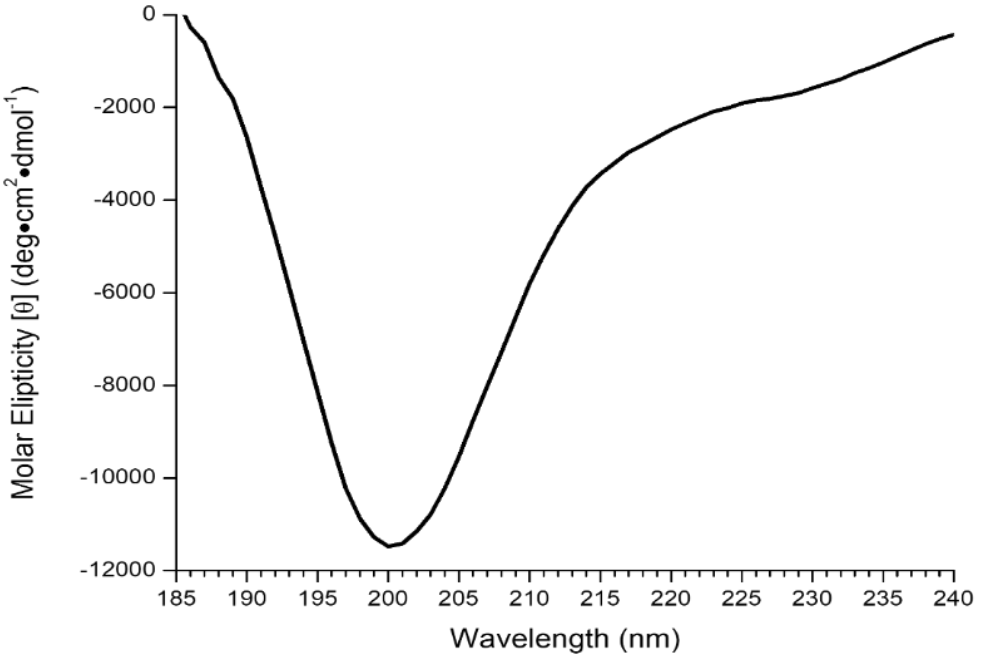

Figure 13- Circular dichroism spectrum of UBact. The strong negative band around 200 nm is indicative of protein random coil. The absence of characteristic bands for α-helix and β-sheet lend credence to the observed lack of secondary structure in UBact. Figure adapted from Bonn and Fushman[137]. Reproduced with permission from Springer Nature.

2.4b Chemical Shift Indexing

Chemical shift indexing (CSI) was used to confirm if any secondary structure could be derived from the secondary chemical shifts of UBact resonances, even if it

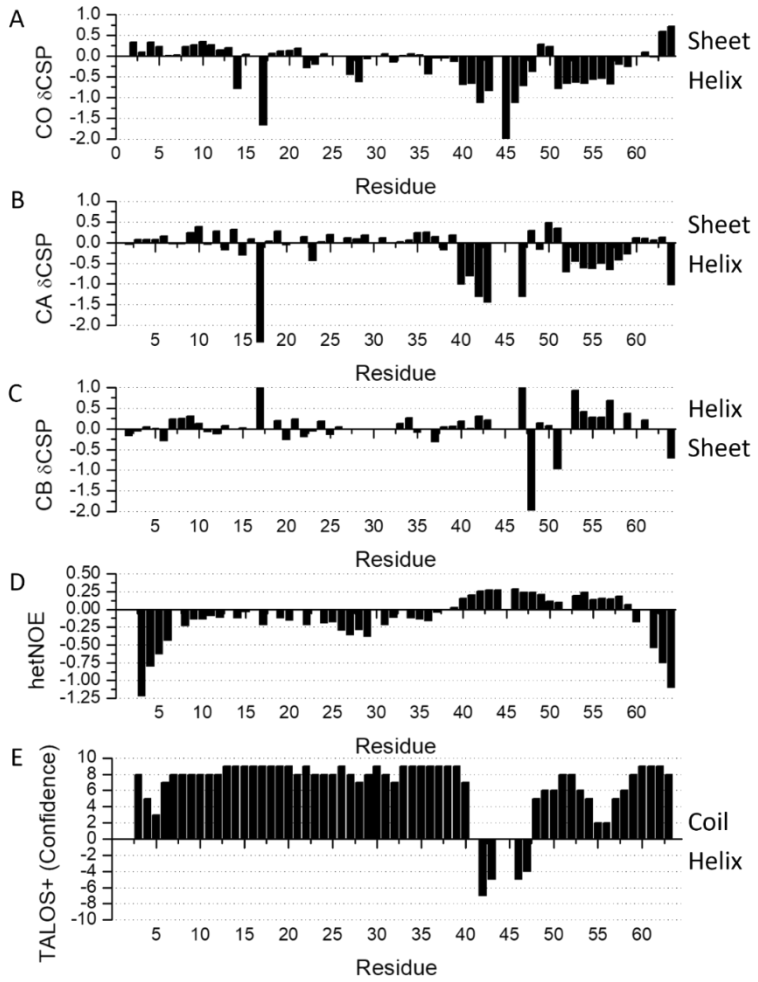

Figure 14-UBact secondary structure prediction from NMR data. A) CO secondary chemical shifts. Positive values are typical of β-sheet, while negative values are indicative of α-helix. B) CA secondary chemical shifts. C) CB secondary chemical shifts. The opposite trend is seen with CB shifts, with positive values corresponding to helical character and negative values corresponding to β-sheet. The large deviations at position 17 are probably due to its positioning between two proline residues. D) ^{1}H-^{15}N heteronuclear NOEs. The strong negative values seen throughout indicate a high degree of flexibility. The positive NOEs are generally below 0.25, which is still well below the ~0.5-0.8 that would be expected from a well-ordered protein. E) TALOS+ secondary structure from dihedral angle prediction. TALOS+ shows a high confidence of random coil throughout the protein, while showing only a short stretch of low-confidence helical character. Figure adapted from Bonn and Fushman[137]. Reproduced with permission from Springer Nature.

were too transient to be seen by CD. In order to avoid any bias in the CSI, random coil chemical shift values based on Flemming Poulsen's nearest neighbor corrections were utilized[141, 142]. CSI supports that UBact is disordered. The region from residue 40-58 shows a small deviation from pure random coil values (Figure 14 A-C) which suggest there could be some transient helix formation in this region.

2.4c Heteronuclear NOE

In order to determine just how flexible UBact was, and whether there was any real posibility of secondary structure formation in the region suggested by CSI, a steady state heteronuclear $^{1}H–^{15}N$ NOE experiment was run. The results (Figure 14 D) show many strongly negative NOEs, indicating large fluctuations in bond orientation, and thus long regions of highly flexible protein. A perfectly ordered, non-flexible protein would have hetNOEs on the order of 0.8, while most stably folded regions in protein have NOEs above 0.6-0.7. The region from 40-58 shows low intensity positive NOEs, suggesting a region that is less dynamic than the N- and C-termini, but not truly ordered, as would be expected for a region of secondary structure, especially a helix.

2.4d Dihedral Angle Prediction

As a prediction of the presence of secondary structure in this region, UBact backbone dihedral angles were predicted from chemical shifts using TALOS+[143]. The output shows extended regions of high confidence of random coil, with only a very short region of potential helix, albeit at low confidence (Figure 14 E).

2.4e Chemical Exchange Saturation Transfer and Relaxation Dispersion

A check to determine if there was in fact any secondary structure present, or if UBact sampled any conformations that were not visible on an NMR timescale, a chemical exchange saturation transfer (CEST) experiment was run. This experiment applies a saturating pulse at varying resonances, and other signals are watched to see if any off-resonance signals are affected, indicating a transfer of that saturation to the visible state. This is a powerful technique for determining if a protein samples NMR-invisble states, even at an extremely low population. In this case, no off-resonance attenuations were seen, providing still more evidence that UBact does not form secondary structure, even transient, in solution. As a final check to determine if any alternate states could be discovered, a CPMG relaxation dispersion experiment was conducted at 600 MHz and 800 MHz, which demonstrated no field dependence of ^{15}N R_2, indicative of no meaningful conformational exchange.

Taken all together, the CD evidence, chemical shift-derived secondary structure predictions (including predicted dihedral angles), heteronuclear NOE data, lack of saturation transfer, and lack of evidence for relaxtion dispersion all suggest that UBact is a highly flexible protein. Although there might be some predicted population of helical secondary structure, none was actually observed with the suite of experiments performed.

2.5 Conclusion

UBact from *N. nitrosa* was efficiently expressed and purified at milligram scale for initial NMR characterization. Utilizing standard experiments, the backbone and partial sidechain NMR signal assignment was completed. Using a variety of

methods, UBact was determined to be highly flexible in solution, and contain no detectable regions of secondary structure. This work serves as the basis for further studies of UBact and its interactions with potential binding partners, as well as the first investigation of any constituent of the UBact proteasome system. Resonance assignments were deposited in the Biological Magnetic Resonance Bank, ID 51116.

2.6 Contributions

Steven Bonn designed the project, performed all sample preparation, experimental setup, and data analysis. David Fushman oversaw the project and assisted with NMR experimental setup and data analysis.

Chapter 3: UBact Binds to the Proteasomal Gate Protein

3.1 Objective

Explore UBact's interaction with the putative proteasomal receptor and determine if this receptor behaves in a similar fashion to Pup's Mpa and eukaryotic Rpt proteins.

3.2 Expression and Purification of the Proteasomal Gate Protein

Escherichia coli-optimized *N. nitrosa* Proteasomal Gate Protein (PGP) 1-209 was synthesized and cloned into the pET28a expression vector between the NcoI/BamHI restriction sites by Genscript (Piscataway, NJ, USA). The final construct consisted of an N-terminal His_6-tag, followed by a thrombin cleavage site (LVPR|GS), followed by the gene sequence. Residue 209 was chosen as the final residue due to its predicted location in a flexible loop between elements of predicted secondary structure based on homology modeling (Figure 15). This plasmid was transformed into chemically competent *E. coli* BL21 (DE3) cells as previously described. Positive transformation was selected for by growth on LB-agar plates supplemented with 50 µg/mL kanamycin at 37°C overnight. Single colonies were picked to inoculate 5-20 mL LB cultures supplemented with 50 µg/mL kanamycin and were grown overnight at 37°C. Overnight growth was used to inoculate 1L LB supplemented with 50 µg/mL kanamycin and 1 mM $MgCl_2$, and cultures were grown to OD_{600}=0.6-1.0 at 37°C, approximately 3 hours, at which point protein expression was induced with addition of 500 µM IPTG at 18°C for at least 12 hours. Cells expressing PGP appeared to grow slightly faster than cells expressing other proteins

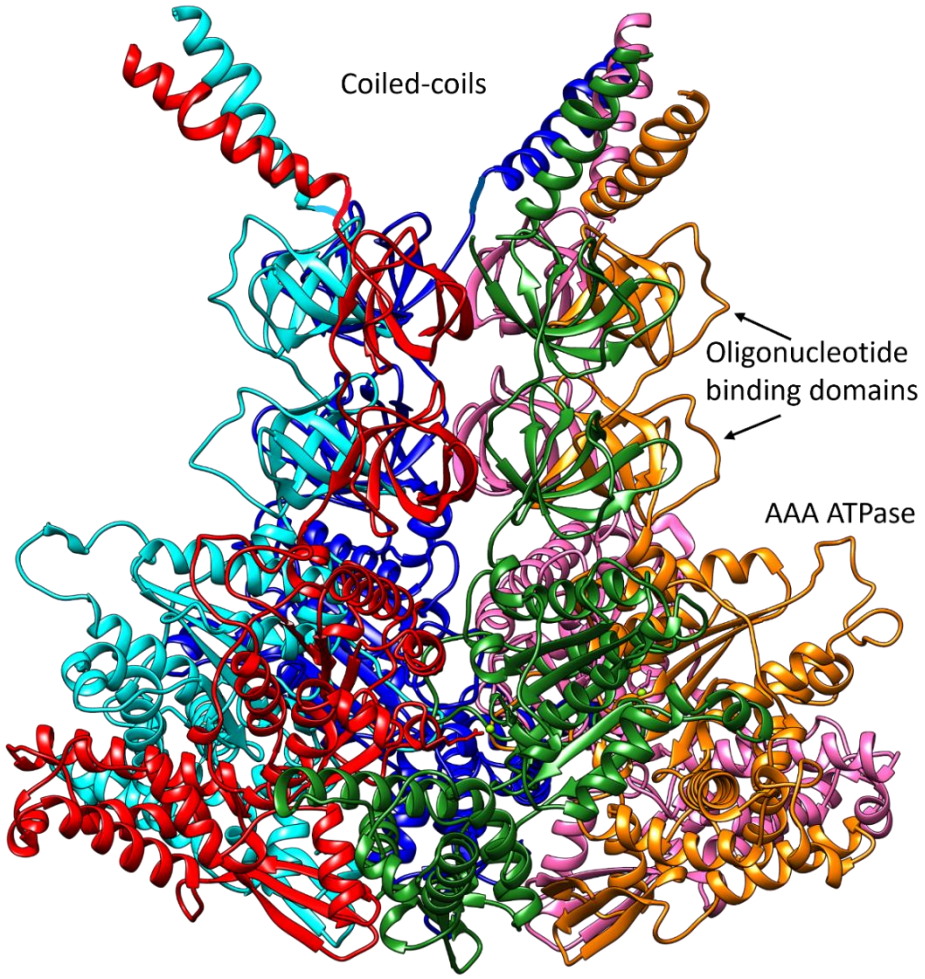

Figure 15- Homology model of the Proteasomal Gate Protein from *N. nitrosa*. The model is based off existing crystal structure of Mpa from *M. tuberculosis*, where the coiled coil and single oligonucleotide binding domains were crystalized, and the oligonucleotide binding domains and ATPase domains were crystalized. The predicted structures were aligned based on the overlapping oligonucleotide binding domains to prepare a model of the full-length protein. The predicted structure forms a homohexamer. Residue 209 is located in an unstructured loop between the second oligonucleotide binding domain and the ATPase domain. The first 50 amino acids are predicted to be disordered and are not modeled here.

in the lab, though growth rate was never formally quantified. ^{15}N PGP was expressed identically to the above, except that the large culture of LB was replaced with 1L M9 minimal media supplemented with 1 g/L $^{15}NH_4Cl$.

Cells were harvested by centrifugation at 4680 x *g* for at least 15 minutes. Cells were resuspended in PBS supplemented with 5 mM $MgCl_2$, DNase, 0.4 mg/mL lysozyme, 0.5% Triton X-100, and 1 mM PMSF, and were lysed by flash freezing in liquid nitrogen followed by rapid thawing at 42°C, followed by sonication for 3 minutes at output 5, duty 70% on ice. Lysate was clarified by ultracentrifugation at 48,900 x *g* for 35 minutes at 4°C. Lysate supernatant was filtered through a 0.45 µm membrane, then was loaded onto a 10 mL HisTrap column (GE Healthcare) preequilibrated with PBS at room temperature. Initially, when all sample was loaded, the column was washed with 20 mM imidazole to remove non-specifically bound material, and a large amount of material flowed through the column. When absorbance dropped, a gradient was run from 20-500 mM imidazole to elute bound protein, and there was a very shallow but very broad elution ranging from 250-500 mM imidazole. SDS-PAGE gel analysis showed that this broad peak was a highly pure protein of approximately 25 kDa, the expected weight of the PGP monomer. A significant amount of 25 kDa protein was present in the flowthrough and wash steps as well, suggesting that there was too much protein to bind to only 10 mL of resin. Upon concentration and buffer exchange into 20 mM sodium phosphate pH 6.8, the protein precipitated into a white powder if NaCl concentrations dropped below ~75 mM. The protein could be immediately redissolved if more NaCl was added to bring the final concentration above the critical concentration. PGP was essentially infinitely soluble and could be concentrated well past the point at which it could not be transferred in a pipette of any volume, yet it did not precipitate or aggregate. Highly concentrated PGP is light brown in color.

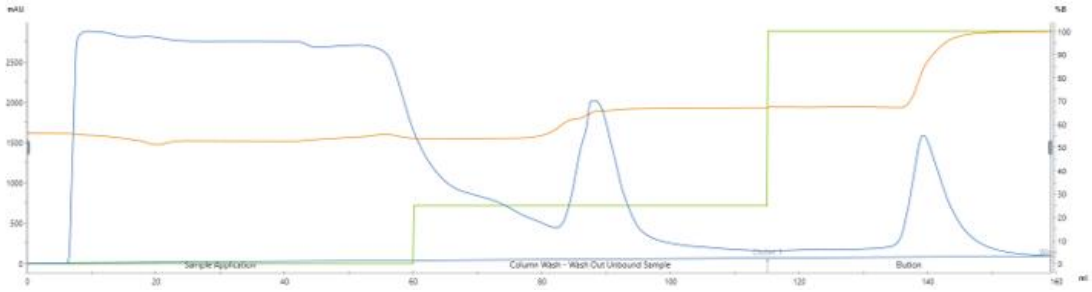

Figure 16- Nickel affinity purification of PGP. Absorbance at 280 nm (mAbs) is shown in blue. Conductivity (mS/cm) is shown in orange. Gradient buffer percentage is shown in green. The protein bound extremely tightly to the column, well above the normal affinity for a single His6 tagged protein. Such high affinity for the column allowed for fast and effective purification.

In order to shorten the amount of time required for purification, and to dramatically reduce the elution volume, the nickel purification method was altered to wash the column with 100 mM imidazole to prevent any non-specific interactions, flowing until absorbance dropped, followed by a single step elution at 500 mM imidazole to elute all bound protein at once (Figures 16 and 17). This protein was buffer exchanged and concentrated into 20 mM sodium phosphate, pH 6.8, 130 mM NaCl for storage, and was stored at -80°C.

A single liter of bacterial culture could yield up to 300 mg PGP, depending on the amount of time spent inducing under the influence of IPTG. 500 µM IPTG expressed roughly 25% more protein than either 200 µM or 1 mM IPTG, and expression at 18°C yielded dramatically more protein than expression at warmer temperatures. Without the use of a refrigerated incubator, room temperature induction was still feasible, yielding 100+ mg per liter of bacterial culture. PGP was soluble in all near-neutral buffers tested (20 mM sodium phosphate, pH 6.8, PBS pH 7.4, with and without high concentrations of imidazole, and 50 mM tris pH 8.0) as long as ionic strength was kept above roughly 100 mM. Sodium chloride, sodium fluoride,

and sodium sulfate were all appropriate for maintaining ionic strength. PGP was soluble up to and past ~60 mM. PGP solutions above 10 mM became too viscous to handle with any accuracy and could only be diluted by layering additional buffer on top of the nearly solid layer of protein and incubating for several hours to allow for natural diffusion of the protein through the solution. Even at these extremely high concentrations, PGP was stable at 37°C for up to four days before visible precipitation occurred, at room temperature for two weeks, at 4°C for over two weeks, and at -20°C indefinitely, for exactly one freeze/thaw cycle. Freezing the protein at -20°C a second time led to immediate precipitation upon thawing. Flash freezing in liquid nitrogen and indefinite storage at -80°C allowed for up to three freeze/thaw cycles before protein aggregated upon thawing. PGP was highly sensitive to and would irreversibly precipitate upon addition of dimethyl sulfoxide (DMSO), dimethyl

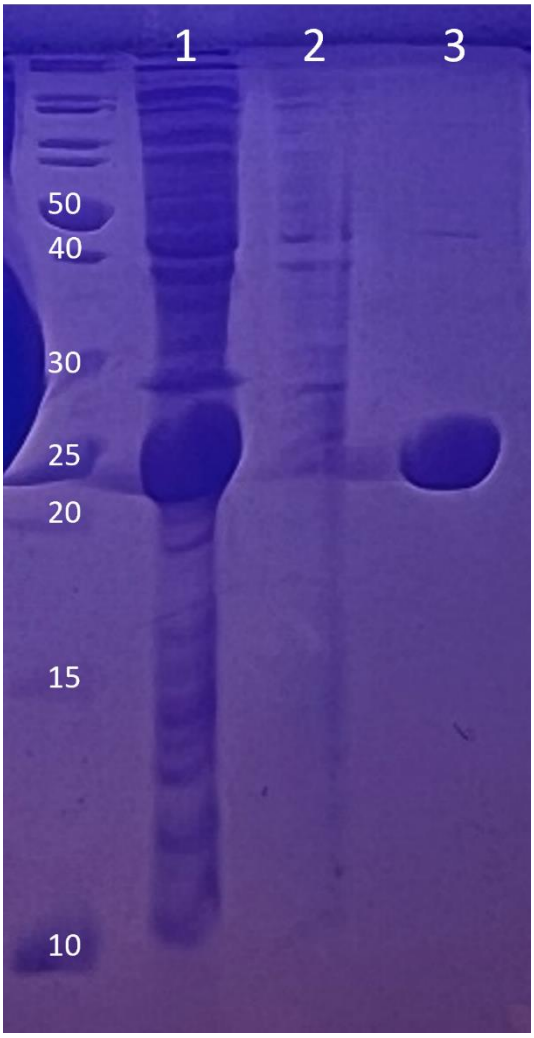

Figure 17- SDS-PAGE analysis of PGP nickel purification. Lane 1 is the clarified lysate supernatant. Lane 2 is the column flowthrough and wash. Lane 3 is the 500 mM imidazole elution. Essentially no PGP flows through the column, nor is it washed out with 100 mM imidazole. The very strong affinity for the nickel column allows for a very high degree of purity. The thick bands at 25 kDa represent a dissociated monomer of PGP.

formamide (DMF), ethanol, isopropanol, or methanol at even extremely low concentrations.

3.3 Determination of Oligomeric State of PGP

The Rpt proteins in the eukaryotic proteasome exist as a heterohexameric complex[9], and Mpa/PAN from the Pup proteasome system has been crystalized[111] and visualized using cryoEM[144] as a homohexamer. Structure prediction with SWISS-MODEL[145-149] suggested that PGP would also exist as a homohexamer, just as one would anticipate from its similarity to other proteasome-associated ATPases (Figure 15).

Proteins bearing a His_6-tag typically elute from an immobilized nickel column upon addition of 125-250 mM imidazole. The unusual elution profile of PGP, eluting above 300 mM imidazole, suggested that there were additional interactions than just a single His_6-tag. Size exclusion chromatography revealed that a concentrated solution of PGP eluted immediately at the void volume of the column (40 mL), whereas a 25 kDa monomer would be expected to elute closer to 60-70 mL. Native gel electrophoresis showed the presence of a species that barely migrated into resolving portion of the gel. A ^{1}H-^{15}N SOFAST HMQC and ^{1}H-^{15}N TROSY of ^{15}N-labeled PGP show only about 40 backbone signals rather than the nearly 200 one would expect to see based on the sequence (Figure 18). The lack of signals could be due to a larger oligomer of PGP forming in solution, as even a dimer would push the size limits of NMR. Together, these results strongly suggest that PGP forms a larger oligomeric species in solution.

To confirm that PGP is in fact folded in solution, a CD spectrum was obtained. Rather than showing a strong negative band around 195-200 nm (indicative of a disordered protein), the spectrum shows the characteristic double dip at ~208 nm and ~222 nm, indicative of helical character (Figure 19). Spectral decomposition with Dichroweb[150-152] suggests roughly 10% helix, 36% β-strand, 18% turn, and 36% other structure or unordered. This is essentially identical to the secondary structure

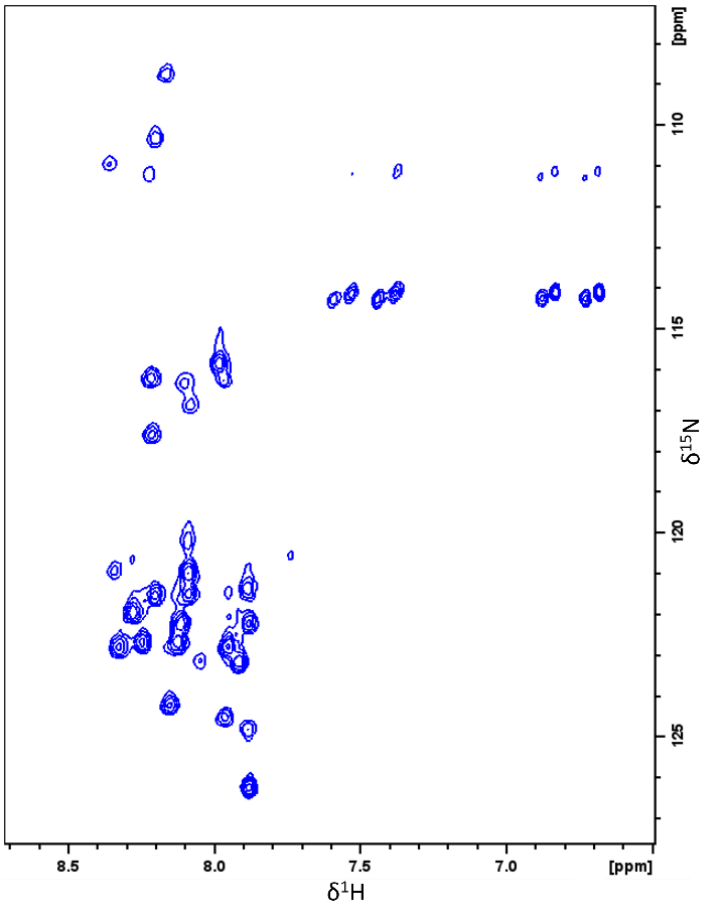

Figure 18- ^{1}H-^{15}N SOFAST HMQC spectrum of ^{15}N-labeled PGP. There are very few signals present relative to the number of expected signals for a 25 kDa monomeric protein. The narrow dispersion of the backbone signals in the proton dimension suggests that the signals that do appear are from unstructured regions or residues. Notably, there are four glycine residues in the predicted unstructured region, and there are four signals around 100 ppm in the ^{15}N dimension, where glycine signals often show up. Additionally, there are two residues that contain sidechain amides (Q28 and N32), and there are only two strong sidechain amide signals. This suggests that the resonances visible correspond primarily to an unstructured region at the N-terminus of the protein.

breakdown when the predicted structure (Figure 13, truncated at position H209) is used to predict a CD spectrum and that spectrum is used to predict secondary structure[153].

In order to determine what size oligomer PGP formed in solution, analytical ultracentrifugation (AUC) was employed. Sedimentation velocity (SV) experiments suggested a large population with a molecular weight of 150 kDa, a hexamer of PGP, with a small population of lower weight oligomers (Figure 20A). Sedimentation equilibrium experiments support this conclusion (Figure 20B). Global and single experiment fits suggest a molecular weight (142 ± 1.5 kDa) approaching that of PGP_6 (151 kDa) for all concentrations tested. This is consistent with both the predicted structure of a homohexamer, as well as comparisons to other proteasome-associated ATPases which exist as hexameric structures. The construct under study is predicted to exclude the C-terminal ATPase domain, indicating that the N-terminal portion is

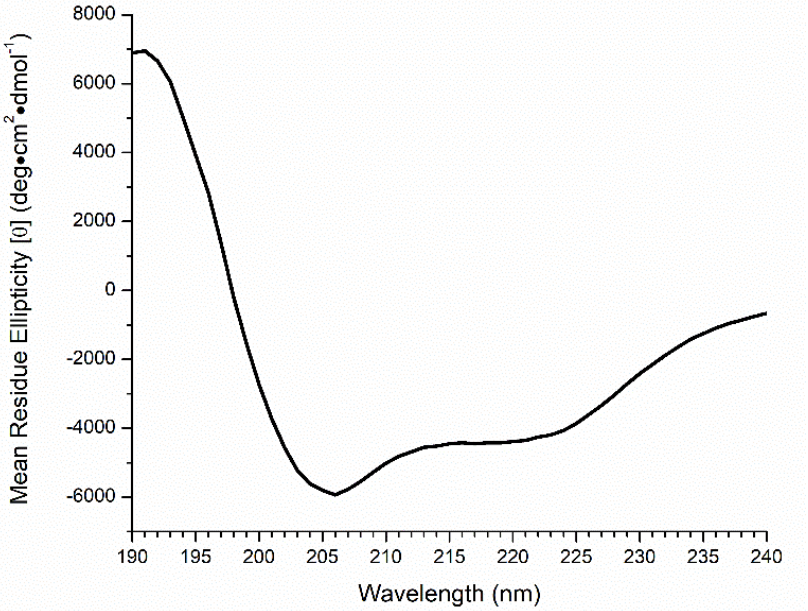

Figure 19- CD spectrum of PGP. The double dips at 208 and 222 nm are indicative of helical character. Spectral decomposition suggests that the protein is folded in solution, and the secondary structure breakdown is consistent when the homology model is used to predict a CD spectrum.

sufficient for oligomerization. There is a population of interconverting lower oligomers, though they are estimated to represent less than 10% of the population at the relatively low concentrations tested.

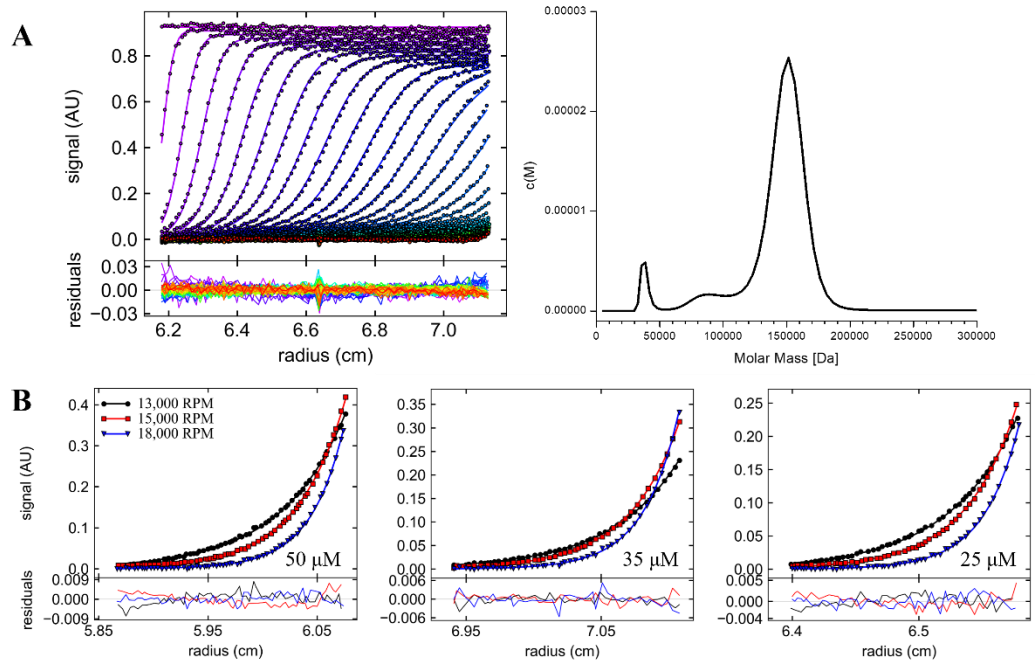

Figure 20- Analytical ultracentrifugation analysis of PGP in order to determine the oligomeric state in solution. A) Sedimentation velocity at 42,000 rpm reveals a molecular mass distribution heavily favoring ~150 kDa, the predicted mass of a PGP hexamer. Representative curves are shown with every fifth point for clarity. B) Sedimentation equilibrium data at three rotor speeds and concentrations again support existence as a hexamer as the primary state, with a predicted molecular weight of 142 ± 1.5 kDa based on $\bar{\upsilon}$=0.7375 ml/g and a solvent density of 1.00613 g/mL. Representative curves are shown with every second data point for clarity. Black circles represent data at 13,000 rpm, red squares at 15,000 rpm, and blue triangles at 18,000 rpm. Data acquired by Dorothy Beckett.

3.4 UBact Binding to PGP_6 Can Be Measured by NMR

3.4a UBact Signals Attenuate Upon Titration of PGP_6

In order to qualitatively determine if there UBact binds to PGP_6 in solution, a concentrated solution of PGP_6 was titrated into a ^{15}N UBact solution, and SOFAST HMQC spectra were acquired at each titration point. Rather than observing chemical

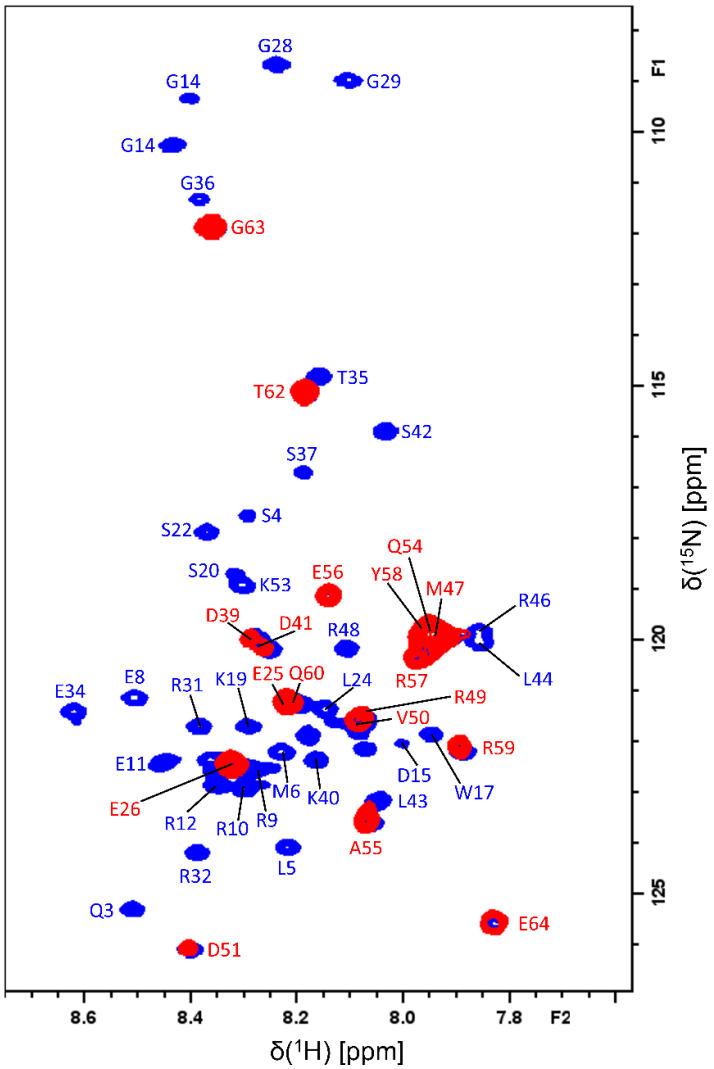

Figure 21-Overlay of SOFAST HMQC spectra of UBact M17W alone (blue) and upon saturation with PGP_6 (red). Resonances are annotated to show which signals remain upon binding with PGP_6. Attenuated signals are primarily located in the N-terminal portion of the protein, while the C-terminal residues remain relatively strong. There is a sharp cutoff in attenuations at residue 39, with subsequent residues generally remaining fairly strong and preceding residues attenuating.

shift perturbations (CSPs), which are a typical measurement of binding seen in NMR experiments, signals from the N-terminus of UBact already attenuated at low concentrations of PGP_6. C-terminal signals attenuated, though not nearly as early in the titration (Figures 21 and 22). At apparent saturation, only signals from the C-terminal half of UBact remain visible, though these signals are somewhat attenuated. This attenuation is indicative of binding between UBact and PGP_6, and the residue-specificity indicates that the binding is selective (versus non-specific, Residues preceding D39 are generally attenuated (with the exception of E26) while the following resonances are still visible. E25 appears to remain visible, but this is most likely due to overlap with Q60. The attenuated region most likely represents the interface between UBact and PGP_6.

A CEST experiment was also conducted to try and identify any states that were invisible by usual methods. Using a high concentration of UBact and low

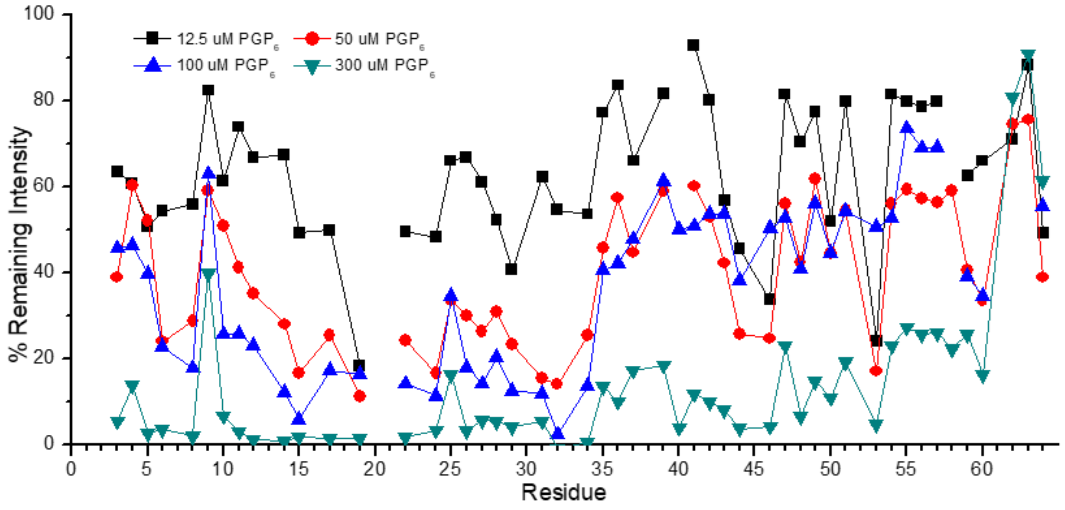

Figure 22- UBact NMR signal intensity upon titration with PGP_6. All signals attenuate, but the N-terminal half of the protein attenuates much earlier in the titration. Even at equimolar ratios of UBact and PGP_6, the C-terminal half of the protein maintains ~10-20% of initial intensity, while the N-terminal portion shows essentially no remaining intensity.

concentration of PGP_6, saturation of the invisible bound state and a long mixing time should allow for that saturation to transfer to the visible state and display some amount of signal attenuation. CEST experiments have been used to identify invisible population states with as low as 2% abundance[154] to identify exchange rates and chemical shifts between states. No visible resonances were influenced by off-resonance saturation, and thus no apparent shifts could be identified to assist in binding quantification.

3.4b Transferred Relaxation Can Be Used To Quantify UBact:PGP_6 Binding

NMR signal attenuation is often due to signal broadening as a result of faster T_2 relaxation times, as signal width is inversely proportional to T_2[155]. In general, short T_2 is associated with larger molecular weight and therefore slower tumbling (and thus the practical upper limit of molecular weight for NMR detection of roughly 50 kDa). Since PGP_6 is a 150 kDa complex, any UBact binding to PGP_6 would effectively increase the apparent molecular weight, so this signal attenuation is to be expected, though inconvenient.

In 2007, Gottfried Otting presented a method by which binding could be measured by transferred relaxation[156]. Since UBact is much smaller than PGP_6, this method is, on its face, appropriate. Essentially, in a fast-exchanging system, the observed UBact relaxation rates at varying titration points with PGP_6 should reflect the presence of even small populations of bound UBact:PGP_6. Unfortunately, the experiment set is costly both in terms of material and experimental time, as relaxation experiments require generally high concentrations or many scans to generate

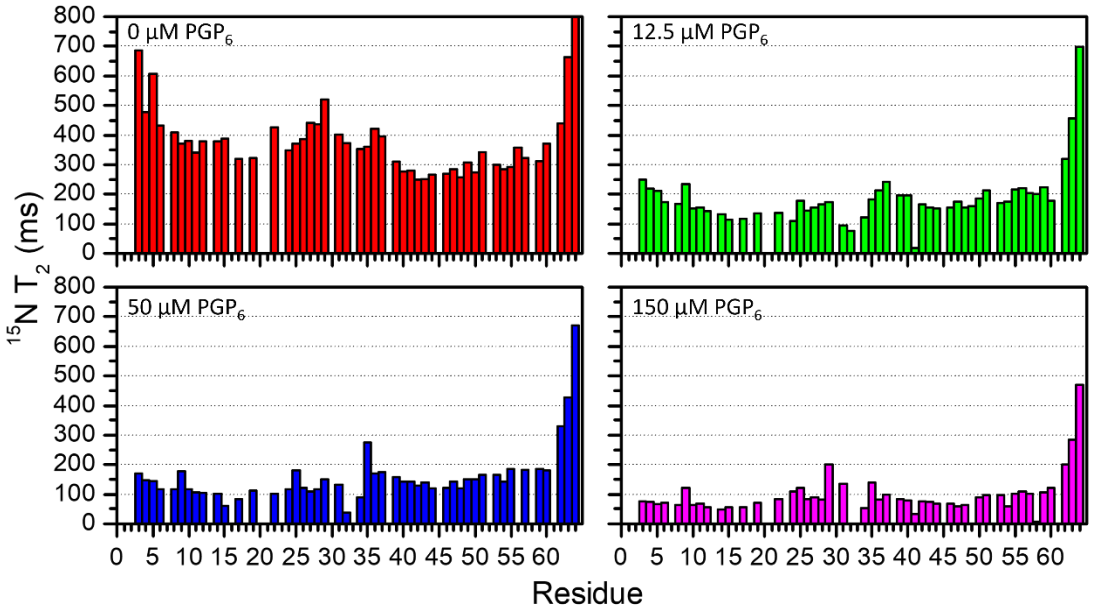

Figure 23- ^{15}N T_2 values upon titration with increasing concentrations of PGP_6. Upon addition of PGP_6, T_2 values decrease corresponding to an increase in apparent molecular weight of the complex.

acceptable signal to noise for accurate rate determination from peak intensity changes.

Separate samples of 300 μM UBact M17W were prepared with increasing concentrations of PGP_6. ^{15}N T_2 relaxation was measured at each titration point. Upon titration with PGP_6, ^{15}N T_2 relaxation times decreased, indicative of an increase in apparent molecular size (Figure 23). N-terminal residues that attenuate showed more dramatic shortening of relaxation times, as would be expected by their more dramatic attenuations, while the C-terminal residues show lessened relaxation enhancement.

Some relatively simple algebra is required to configure change in relaxation for fitting to a binding curve.

$$p_{free} + p_{bound} = 1 \quad \therefore \; p_{free} = 1 - p_{bound} \tag{1}$$

$$R_{obs} = R_{free} \cdot p_{free} + R_{bound} \cdot p_{bound} \tag{2}$$

$$R_{obs} = R_{free}(1 - p_{bound}) + R_{bound} \cdot p_{bound} \tag{3}$$

$$R_{obs} = R_{free} + R_{bound} \cdot p_{bound} - R_{free} \cdot p_{bound} \quad (4)$$

$$R_{obs} = R_{free} + p_{bound}\left(R_{bound} - R_{free}\right) \quad (5)$$

$$R_{obs} - R_{free} = p_{bound}\left(R_{bound} - R_{free}\right) \quad (6)$$

$$\left(R_{bound} - R_{free}\right) \equiv \Delta R_{max} \quad (7)$$

$$\Delta R_{obs} \equiv R_{obs} - R_{free} \quad (8)$$

$$\frac{R_{obs}}{R_{free}} - 1 = p_{bound} \cdot \frac{\Delta R_{max}}{R_{free}} = \frac{\Delta R_{obs}}{R_{free}} \quad (9)$$

$$\frac{R_{obs}}{R_{free}} = \frac{T_{free}}{T_{obs}} \quad (10)$$

$$\frac{T_{free}}{T_{obs}} - 1 = \ p_{bound} \cdot \Delta R_{max} \cdot T_{free} \quad (11)$$

where p_{free} and p_{bound} are the population in either free or bound state, respectively, R_{obs} is the relaxation rate observed at each titration point, R_{free} is the relaxation rate of free UBact, R_{bound} is the relaxation rate of UBact upon saturated binding to PGP_6, ΔR_{max} is the change in relaxation upon complete binding, T_{free} is the relaxation time (inverse of the rate) of free UBact, and T_{obs} is the observed relaxation time at each titration point. These can be fit to a quadratic equation to solve for the dissociation constant using the in-house Matlab program kd_fit[157].

Fitting individual residues to the quadratic equation referenced shows two general regions of interaction: residues 6-36 show dramatic decrease in T_2, while the residues 3-5 and 36-64 show a dramatically lessened response. It is reasonable to believe that the 6-35 region is the primary point of contact between UBact and PGP_6, with the T_2 of residues 36-64 more influenced by being anchored to PGP_6 and not directly interacting with the hexamer. The fact that the relaxation times do not

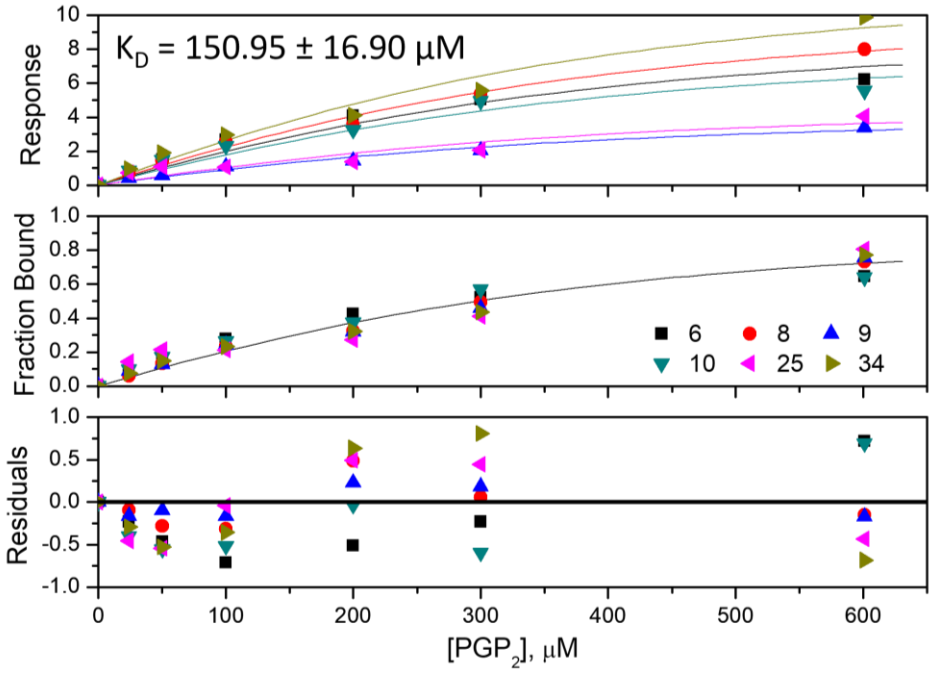

Figure 24- UBact binding to PGP_6 quantified by transverse relaxation. Top-The response, $(T_{free}/T_{obs})-1$, at each point fit to a global analysis using representative residues. Middle- the normalized fraction bound. Bottom-Residuals of the response curve fit. Error was determined by fitting residues individually and calculating the standard deviation of those fits.

dramatically decrease suggests that these residues are probably still mostly free in solution. A global fit of representative residues is shown in Figure 24. Error was calculated by fitting residues individually, then determining the standard deviation of those fits.

The residuals of both the individual and global fit suggest that UBact binds with a 3:1 stoichiometry, with three identical binding sites on PGP_6. With a K_D of 150.95 ± 16.90 µM, this interaction is observed to be three times weaker than that of Pup for the coiled-coil of Mpa, and substantially weaker than that of Pup for full length Mpa[126]. Based on the structural homology and the known interaction of Pup and Mpa from *Mycobacteria*, it would be reasonable to assume that UBact interacts with the coiled-coil domain of PGP_6, as in the Pup:Mpa interaction[111]. Since there are

predicted to be three coiled-coil regions, it makes sense that UBact could bind at each coiled-coil. In this case, it appears that three UBact molecules can bind simultaneously, and that the region of interaction is much closer to the N-terminus in the UBact system. These differences could arise from differences between the two systems, or the results could be influenced by the use of an N-terminal truncation of PGP_6. Experiments with the full length PGP_6 protein would clarify this difference. Interaction with full-length PGP_6 might allow for UBact to bind into the central pore of the protein, and thus preclude interaction of three UBact molecules simultaneously. Presenting that additional binding surface could also serve to increase the affinity of UBact for PGP_6 to be more in line with literature values for Pup interacting with Mpa if they really are similar systems.

To better understand how UBact interacts with PGP_6, a spin label could be attached to PGP_6 to provide rough distance information from UBact to the spin label. Site-specific attachment of *S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1H-pyrrol-3-yl)methyl methanesulfonothioate (MTSL), a nitroxide-containing compound, results in resonances up to ~25 Å away experience paramagnetic relaxation enhancement (PRE) and a decrease in signal intensity due to faster relaxation[158]. Thus, attachment of MTSL at various sites on PGP_6 would yield attenuations in the UBact signals if they were within the 25 Å radius of the nitroxide functional group. Placing MTSL at multiple sites in PGP_6 could help to create a rough map of which region of PGP_6 is involved in interactions with UBact. MTSL attachment at a region directly involved in interacting with UBact would yield dramatic signal attenuations in UBact, while MTSL at a position far away from UBact would have little influence on UBact

signals. Attempts to attach MTSL to PGP_6 via a disulfide linkage resulted in the immediate precipitation of the protein at all concentrations tested. This was most likely due to sensitivity to the methanol solvent of MTSL, rather than any induced insolubility from MTSL attachment. This sensitivity to non-aqueous solvents could pose a barrier to further studies using various extrinsic labels with marginal solubility or stability in water, such as fluorophores or other paramagnetic labels.

3.5 UBact Binding to PGP_6 Confirmed by Orthogonal Methods

Since transferred relaxation is not a widely utilized method, orthogonal techniques were attempted to confirm the NMR-derived dissociation constant. In addition, this analysis gave residue-specific dissociation constants, rather than a single unified number to characterize the interaction. Utilizing methods that characterize the interaction as a whole will provide a single K_D and lend support to the use of transferred relaxation as a viable method for measuring binding. Surface plasmon resonance (SPR) has previously been used to quantify the binding between Pup and Mpa[159], showing a sub micromolar K_D. Isothermal titration calorimetry (ITC) has also been previously used to characterize the binding between Pup and full-length Mpa[126], yielding a K_D of ~4 μM. This was subsequently confirmed by fluorescence anisotropy[126].

3.5a Surface Plasmon Resonance

SPR is a fantastic technique for determining the kinetics of biomolecular interactions. SPR measures the on and off rates of interactions by tracking the increase in response upon the binding of an analyte to its immobilized binding partner. In this case, immobilized UBact can be used as bait for PGP_6. Since UBact is

very small relative to PGP_6, the response should be much larger upon binding of PGP_6. Attempts to immobilize UBact on a CM5 chip by amine coupling for SPR resulted in no discernable immobilization at any concentrations attempted, as determined by a lack of any persistent response following the binding reaction.

3.5b Fluorescence Quenching

Fluorescence quenching is a fairly straightforward and sensitive technique to determine the interaction between two molecules. Fluorescence is highly dependent on the local environment of the fluorophore, and changes to that environment, such as direct contact in binding, reduction in mobility, or change in solvent exposure, can alter the fluorescence profile. Changes in fluorescence could be seen as a change in emission intensity or a shift in the wavelength of maximal emission.

The incorporation of tryptophan at position 17 provides an internal fluorophore, and the NMR signal attenuations suggest that the region containing rseidue17 is involved in binding to PGP_6; thus, it serves to reason that there should be some sort of fluorescence change upon interaction with PGP_6. There are no tryptophan residues present in PGP_6, thus the only fluorescent emission should come from UBact M17W.

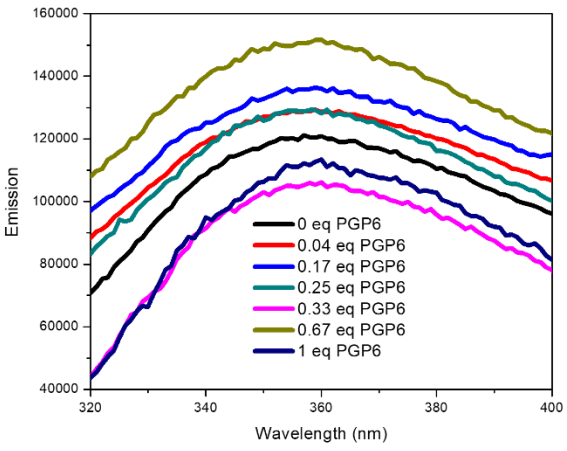

Figure 25- Fluorescence quenching of tryptophan in UBact M17W upon binding to PGP_6. The response does not show any pattern consistent with quenching of tryptophan fluorescence in a [PGP_6]-dependent manner.

Excitation at 295 nm should provide for specific excitation of

tryptophan without bleeding over into exciting any tyrosine residues, of which there are 30 per PGP hexamer. Indeed, there is modest emission from the tryptophan on UBact upon excitation at 295 nm, with maximal emission at 360 nm. However, there is still emission from tyrosine residues with a local maximum around 340 nm. Altering the excitation wavelength or narrowing excitation bandwidth immediately abolished all emission. Subtraction of PGP_6-only spectra at each concentration and correction for the inner filter effect did not yield any dose-dependent change in tryptophan emission from UBact. Fluorescence emission jumped up and down at each point tested, and there was no trend to any change in the emission, nor was there a red- or blue-shift in the emission maximum (Figure 25).

3.5c Isothermal Titration Calorimetry

ITC is a powerful tool for determining the thermodynamic parameters of intermolecular interactions. ITC automates a titration and measures how much heat must be added or removed from a system in order to maintain a constant temperature. This directly measures the heat of the interaction and any conformational changes that take place upon binding according to:

$$\Delta G = -RTlnK_A = \Delta H - T\Delta S \quad (1)$$

Where ΔG is the Gibbs free energy of a reaction, R is the ideal gas constant, T is the temperature in Kelvin, K_A is the association constant of a binding reaction, ΔH is the enthalpy, and ΔS is the change in entropy. ITC directly measures the heat of interaction, and by integrating the injection peaks and plotting the molar ratio of interacting species against ΔH (in kcal/mol), K_A and stoichiometry can be determined. In addition, solving for enthalpy reveals the degree to which hydrogen

bonding and van der Waals forces are involved in the interaction, while solving for entropy reveals the influence of hydrophobic interactions (due to excluded water). A test to scan near the range of the SPR-derived K_D value of ~100-200 nM for the Pup:Mpa interaction[159] yielded no heat of binding at all, suggesting that there is no meaningful interaction at such low concentrations, consistent with the NMR data. In order to test a 25-100 μM K_D on the VPITC machine, one would need a tremendous amount of material (upwards of 50 mg UBact per run), assuming 1 UBact:1 PGP_6 stoichiometry. Thus, it was decided to take advantage of the National Cancer Institute's Biophysical Resource's ITC200, which requires dramatically less material at the cost of sensitivity. Several experiments yielded very small heats of binding, and no meaningful estimate of stoichiometry. When testing for a 25-100 μM K_D, no tests ever approached saturation, nor was the data generated ever sufficient to estimate binding affinity or stoichiometry. Interestingly, upon titration of UBact into a solution of PGP_6, there was an initial heat peak, consistently followed by a second smaller peak (Figure 26). The first peak is most likely the heat of binding as UBact and PGP_6 interact, and the second peak could be a conformational shift in UBact as it adopts a binding-competent structure, possibly a helix[111]. The extremely high viscosity of concentrated PGP_6 prevented its use as the material in the syringe, so a reverse titration could not be attempted.

3.5d Analytical Ultracentrifugation

Analytical ultracentrifugation (AUC) is a useful tool for determining the hydrodynamic properties of a species in solution. AUC is particularly powerful when determining the interaction between two species of very large size difference, as is the

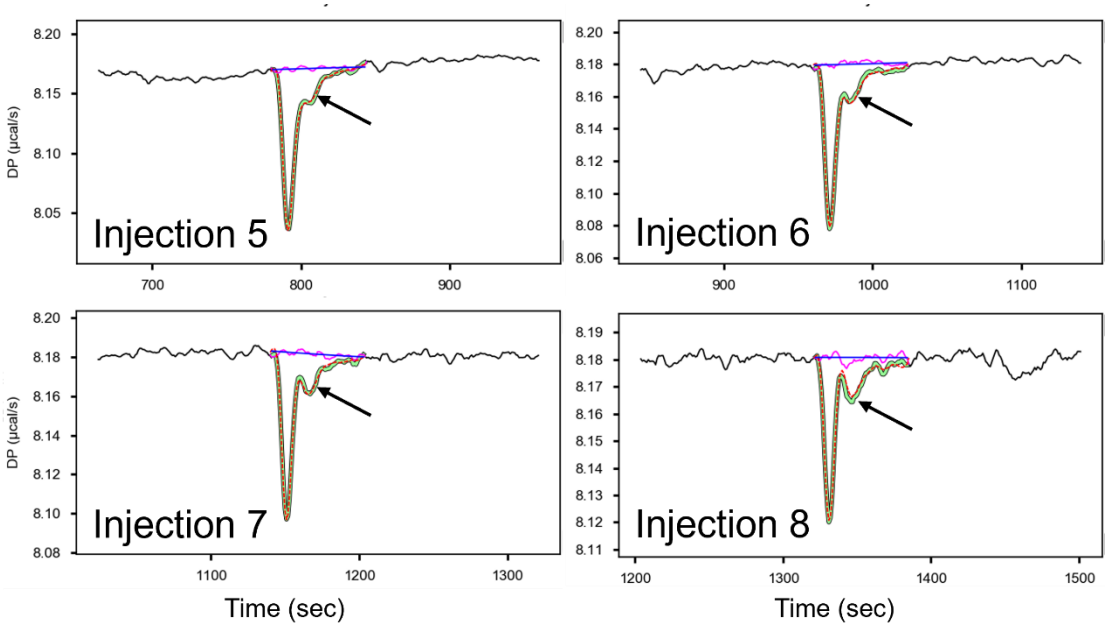

Figure 26- Two heat peaks appear upon titration of UBact into a solution of PGP_6. The first peak is most likely the interaction between UBact and PGP_6. The second peak, denoted by the arrow, could be UBact undergoing a conformational shift. This peak persisted across all trials. Representative injections are shown.

case between UBact and PGP_6. The use of differential detection for each species allows for the tracking of an apparent change in molecular weight of each species, which can be used to determine the strength of a binding interaction. Sedimentation equilibrium (SE) is a method in which a sample is allowed to reach diffusional equilibrium, at which point the dispersion of molecules in the sample is determined. This is done at several different rotor speeds, allowing for accurate determination of hydrodynamic properties such as molecular weight. The primary downside of SE experiments is the extremely long time it can take for species to reach equilibrium, and the relatively low number of data points acquired as compared to sedimentation velocity (SV) experiments. The long times to equilibrium can be shortened by using a reduced cell length, which has the bonus of allowing for more samples to be run simultaneously. Using the longer 12 mm cell lengths can lead to equilibration times

up to two days per rotor speed tested, while using the 3 mm cells can allow for equilibrium to develop in as little as several hours (eight- and nine-hour tests are standard). The advantage is that because SE is an equilibrium method, many components of analysis simply cancel each other out, leading to a straightforward analysis. SE is also more forgiving of hydrodynamic non-ideality, in which diffusional transport is influenced by concentration. SE also allows for molecular weight determination independent of molecular shape, which is important as most SV models are tailored to globular proteins. AUC has been used to determine the strength of interactions between molecules across nine orders of magnitude, from low nanomolar dissociation all the way to molar K_D. The development of fluorescence detection systems is extending this capability to sub-nanomolar range.

In order to utilize SE AUC to characterize the binding between UBact and PGP_6, a specific label needs to be attached to UBact, since that will theoretically undergo the larger molecular weight shift (7.9 kDa to ~150 kDa upon complexation with PGP_6). A cysteine was incorporated at position 63 in UBact, and fluorescein-5-maleimide was conjugated at this position. Since the binding site of UBact was already mapped by NMR, position 63 was chosen to keep the fluorophore away from PGP_6 and prevent it from influencing the interaction in any meaningful way. Fluorescein absorbs at 494 nm. In this case, emission was not measured, as the instrument utilized does not have fluorescence detection capabilities. However, the large absorbance at 494 nm allowed for tracking of UBact specifically, as PGP_6 does not absorb at all near that wavelength. In addition, measurements were taken at 280 nm, though these would be increasingly influenced by increasing concentrations of

PGP_6. Excess tag was removed from the solution by extensive buffer exchange in 3 kDa MWCO filters. Multiple attempts at tagging showed that this is a low efficiency reaction, with only 25-30% of cysteine being tagged, as determined by absorbance, and utilizing the correction factors provided by the manufacturer. The fluorescein-tagged UBact construct did not fly on ESI-MS. Dissolving the fluorophore in DMSO to a relatively high 10 mM concentration gave the highest yield, significantly higher than using a DMF-dissolved tag.

Upon titration of PGP_6 into solutions of UBact-fluorescein (UBact-F), absorbance at 494 nm increased at increasing radius, suggesting an increase in the molecular weight of the species absorbing at 494 nm. At higher concentrations of PGP_6 and higher rotor speeds, essentially all the absorbance was piled up at the outermost radius of the cell, and the data was unusable. At lower molar ratios and relatively slower rotor speeds, fitting of 494 nm data together to a simple A+B→AB interaction model, using the dimeric PGP as the competent binding unit, gave a binding constant of 154 ± 4.54 μM (Figure 27). This number is consistent with the global fit of the NMR relaxation data. Since SE-AUC is a well-known and robust method for quantifying binding across a wide range of affinities, its agreement with the transferred relaxation method lends credence to the applicability of the NMR methodology.

3.6 Conclusions

NMR titration studies show that UBact and PGP_6 specifically interact along the N-terminus of UBact, while the C-terminus remains essentially freely rearranging in solution. SE AUC confirmed the NMR-derived K_D numbers as robust and show the

usefulness of using transferred relaxation to expand the range of molecular sizes that can be studied by NMR. The K_D is different from the observed K_D of Pup interacting with the full length Mpa or the N-terminal coiled-coil domain fragment of Mpa. The PGP_6 construct utilized here contains the same putative coiled-coil and two putative β-barrel domains responsible for oligomerization, but not the large C-terminal ATPase domain. Full length Mpa binds Pup approximately five times more tightly than the coiled-coil domain, suggesting that there are interactions between Pup and the more C-terminal domains that are vital for full functionality. In the case of UBact, there could be more interactions between UBact and the PGP_6 ATPase domain, which could be probed in future experiments utilizing the full-length construct.

The regions involved in the interaction with the proteasomal receptor proteins are slightly different between UBact and Pup. Pup interacts primarily with residues through its middle and C-terminus of the protein, while UBact utilizes its N-terminal half. Again, this apparent difference could be due to the use of different proteasomal receptor constructs, or it could reflect a fundamental difference in the Pup and UBact systems. The longer region of N-terminal residues that appear "free" in the Pup system could be residues that sit inside the central pore of the Mpa hexamer, facilitating transport through the pore and into the 20S CP of the proteasome, while the extreme C-terminus remains covalently attached to a substrate and drags that substrate to its ultimate degradation. Further experiments with full-length PGP_6 could clarify this discrepancy in apparent affinity and binding site. If a full-length PGP_6 construct shows a similar binding site and affinity as that of Pup:Mpa, then perhaps this truncated construct demonstrates an aborted binding event, where an initial

lower-affinity interaction facilitates ATP-dependent translocation into the central pore of PGP_6 and then further down into the 20S CP.

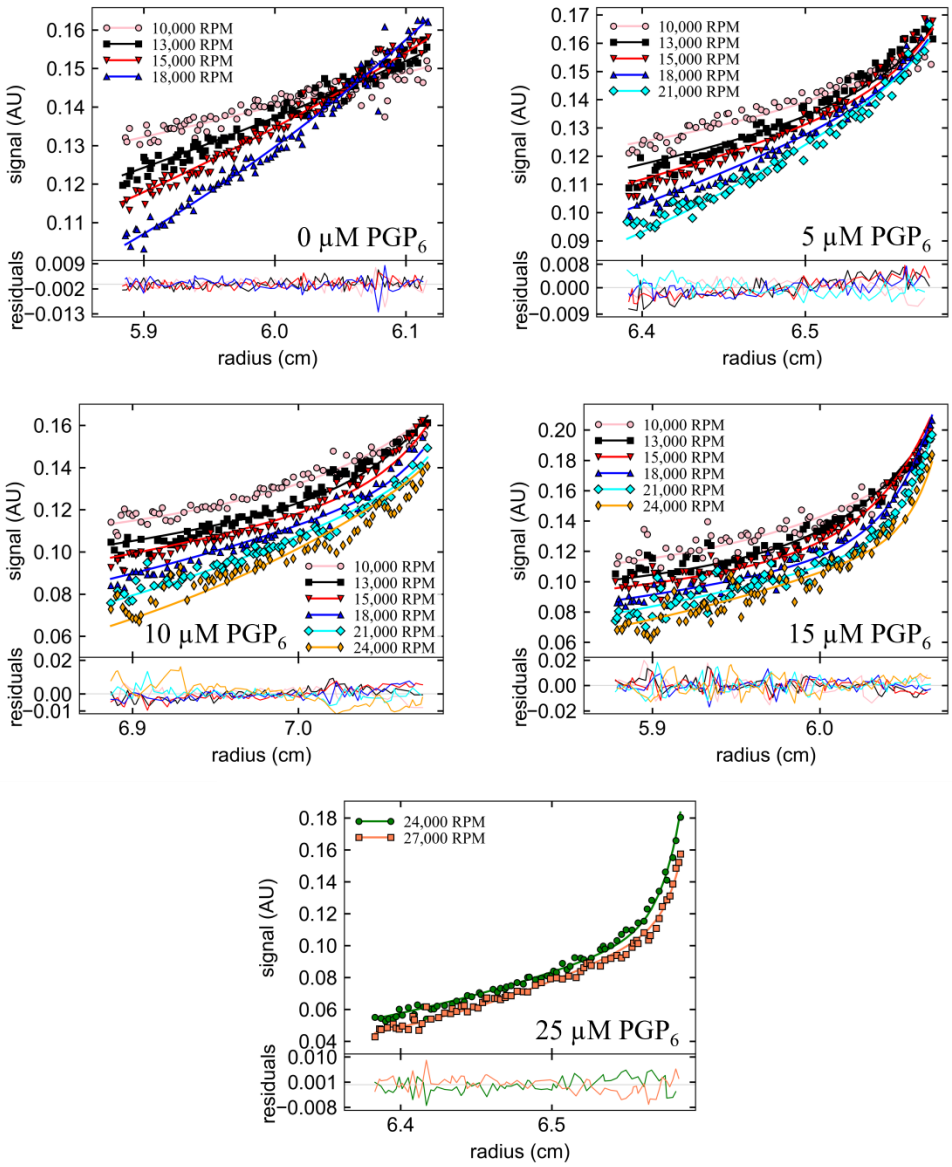

Figure 27-Sedimentation equilibrium analytical ultracentrifugation titration of UBact and PGP_6. Fluorescein-labeled UBact is held constant at 5 µM, PGP_6 is varied from 0 to 5 molar excess. Absorbance is measured at 495 nm, so PGP_6 does not influence the absorbance at all. At increasing PGP_6 concentration and increased rotor speed, more absorbance appears at a wider radius, qualitatively suggesting an apparent increase in molecular weight. Lines represent a global fit to a 3 UBact to 1 PGP_6 model.

3.7 Contributions

Dorothy Beckett (UMD, NIH) performed sedimentation velocity and equilibrium experiments on PGP. Dorothy Beckett is currently the Director of the BBCB Division of NIGMS, NIH. This work is not endorsed by either the NIH or NIGMS. Connor Donahue and Guanghui Zong (UMD) provided advice for SPR experimental design and operated the Biacore instrument. Sergey Tarasov (NCI) set up and performed ITC experiments on the ITC200 and assisted with data analysis. Ian Ferencz (UMD) assisted in setting up SE-AUC titration experiments. Steven Bonn designed the overall project, performed all sample preparation, performed all other experiments, and handled all data analysis. David Fushman oversaw the project and assisted with data analysis.

Chapter 4: Attempted Expression, Purification, and Characterization of UBact Ligase and Deamidase Enzymes

4.1 Objective

Express, purify, and determine the activity of the putative enzymes involved in UBact ligation to and deamidation/removal from substrates.

4.2 Identification of the Ligase and Deamidase Enzymes

In the Pup proteasomal operon, the ligase (PafA) and deamidase (Dop) are sequentially and structurally very similar to each other. The putative genes encoding the corresponding UBact ligase and deamidase enzymes are very similar as well and cannot immediately be differentiated. The main sequence and structural difference between the two genes is the presence of an additional loop in the N-terminus of Dop (and probably the UBact deamidase) that is thought to potentially block substrate binding[123, 160]. In order to identify the putative ligase enzyme from the UBact operon, both putative enzyme genes were aligned (Figure 28) and were run through SWISS-MODEL[146, 148, 149] and the resulting structures were compared (Figure 29). The primary sequences are 45% identical, and share many chemically conserved regions as well The predicted structures are essentially the same, and they overlay extremely well with previously solved crystal structures of PafA and Dop[160]. Notably, structure prediction and sequence alignment show the presence of a loop inserted near the N-terminus of the putative UBact deamidase enzyme (shown in red in Figure 29). This is enough of a prediction to differentiate the putative ligase and deamidase enzymes.

4.3 His-Ligase Initial Tests

4.3a Expression

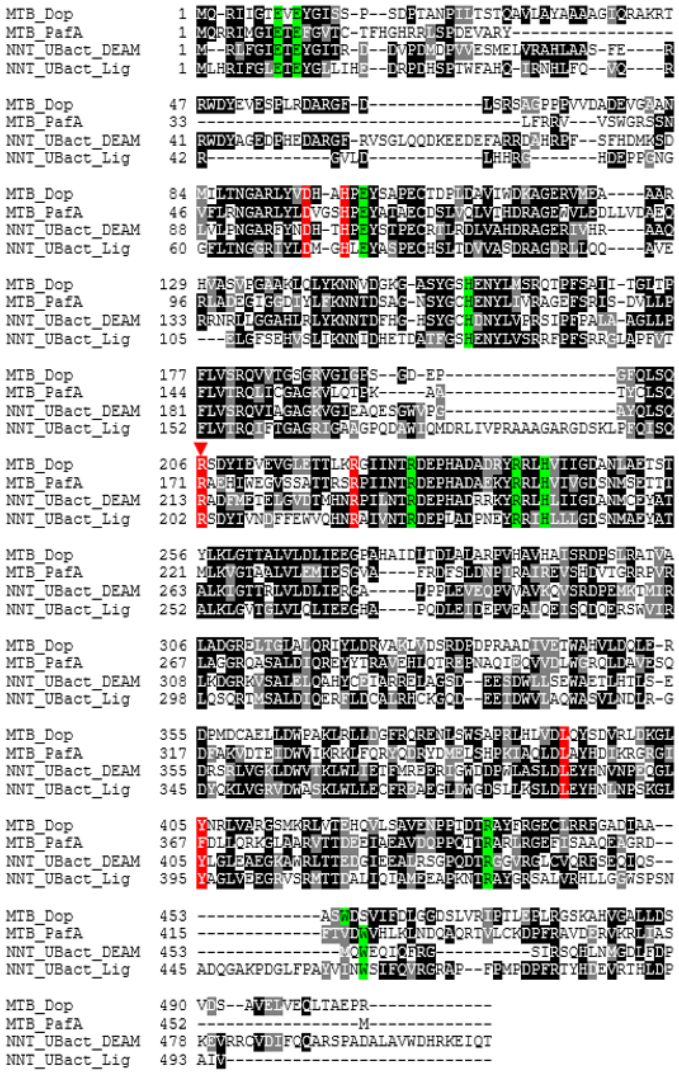

Figure 28- Multiple sequence alignment of Dop (deamidase of Pup) PafA (Pup ligase) from *Mycobacterium tuberculosis* with *Nitrospira nitrosa* putative UBact deamidase (DEAM) and ligase (Lig) genes. Identical residues are highlighted in black and conserved sidechain chemistry is highlighted in gray. Residues involved in ATP/ADP binding are colored green. Residues involved in Pup binding (and therefore possibly UBact binding) are shown in red. The arginine residue responsible for directly coordinating the C-terminal glutamate in Pup is highlighted with a red triangle. These residues are determined from the crystal structures of PafA with and without Pup and ATP and Dop with and without ADP. Dop and the putative UBact deamidase both have an N-terminal insertion that imparts deamidase activity in Dop (colored red in Figure 25).

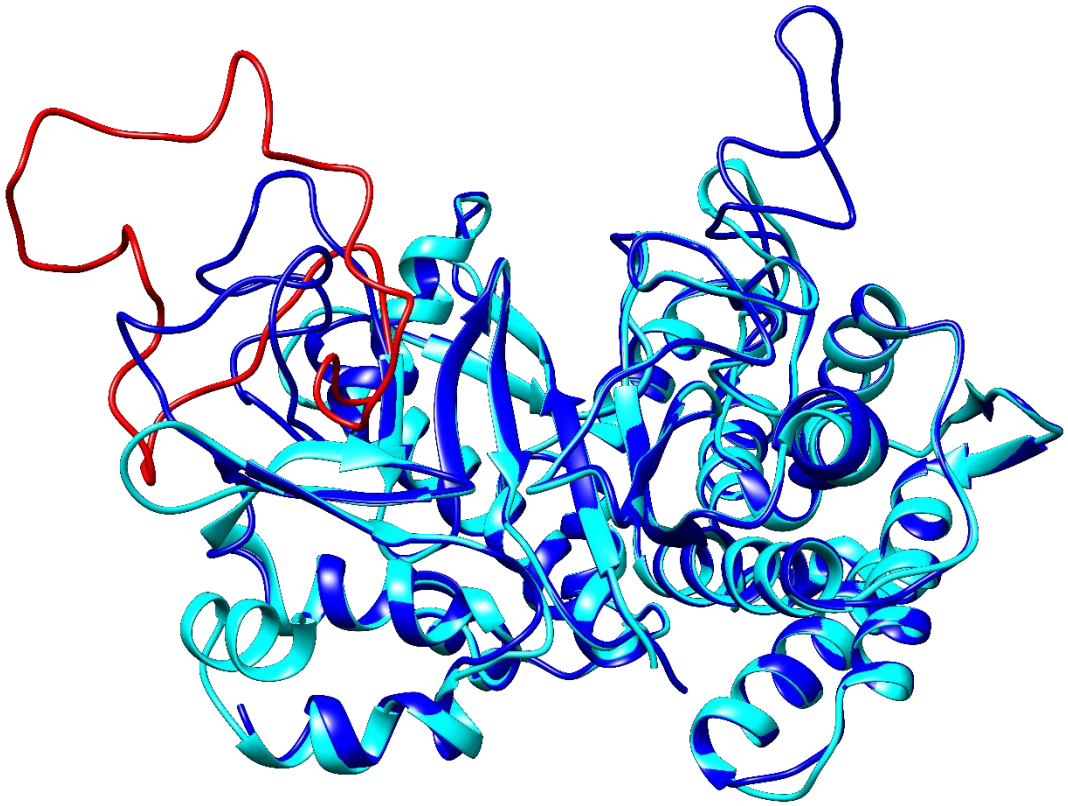

Figure 29- Overlay of the predicted structures of the putative ligase and deamidase enzymes from the *N. nitrosa* UBact operon. Ligase is shown in blue, deamidase is shown in cyan, and the putative deamidase loop is shown in red. The structures overlay almost perfectly, except at the deamidase loop. Homology models based off of crystalized PafA and Dop from *M. tuberculosis*. If the UBact operon follows a similar interaction pattern to the Pup operon, UBact would bind in the central cavity between the two domains of the protein.

E. coli optimized UBact ligase gene (Uniprot accession number A0A0S4L4F5) was synthesized and cloned into pET15b between NcoI and BamHI sites by Genscript (Piscataway, NJ, USA). The final construct contained an N-terminal His_6 tag, a thrombin cleavage site, and then the ligase gene. *M. tuberculosis* and *M. smegmatis* PafA (Pup ligase) and Dop (deamidase of Pup) have both been expressed in a soluble form in both *E. coli* and *M. smegmatis* with the His_6 tag under a variety of different conditions[123]. This plasmid was transformed into BL21 DE3

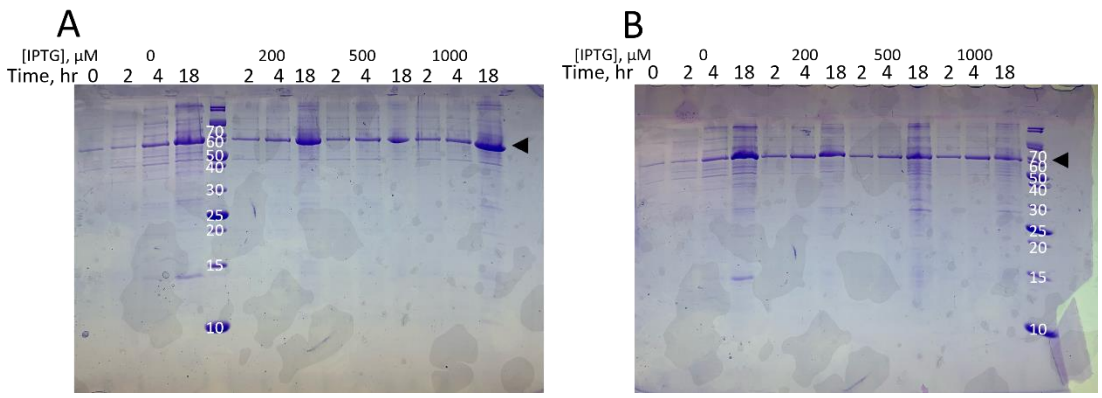

Figure 30-Test expression of pET15b His-Ligase from BL21 (DE3) cells. A) Expression at 18°C, B) Expression at 37°C. The protein overexpresses particularly well under all tested conditions and displays leaky expression in the absence of IPTG. It would later be discovered that none of this expression was soluble, and that even in the absence of inducing agent, all expressed protein was directed to inclusion bodies.

cells as previously described, and successful transformation was selected by plating on LB-agar plates supplemented with 100 µg/mL ampicillin. Single colonies were picked and grown in 5 mL overnight culture. An expression test was performed to determine optimal expression temperature and IPTG concentration. (Figure 30). The test showed that there was generally leaky expression, but that addition of 1 mM IPTG at 18°C overnight gave the strongest band by gel.

5 mL small cultures were used to inoculate 1L LB large cultures supplemented with 100 µg/mL ampicillin. Cultures were grown to OD_{600} = 0.6, then were induced were induced with 1 mM IPTG at 18°C to grow overnight.

4.3b Nickel Affinity Purification

Cells were harvested by centrifugation at 4680 x *g* for 20 minutes at 4°C. Cells were resuspended 50 mM tris pH 8.0, 500 mM NaCl, 5% (v/v) glycerol, 20 mM $MgCl_2$, DNase, 0.4 mg/mL lysozyme, 3 mM PMSF, and 1% Triton X-100, and were lysed by flash freezing in liquid nitrogen and thawing at 30°C, followed by light

sonication at output 4, duty 30% for three cycles of 1 minute on, 3 minutes off on ice. Lysate solution was clarified by ultracentrifugation at 48,900 x *g* for 30 minutes at 4°C. Supernatant was filtered through a 0.45µm membrane, and the filtered supernatant was loaded onto a HisTrap nickel affinity column (GE Healthcare) pre-equilibrated with 50 mM tris pH 8.0, 500 mM NaCl, 5% (v/v) glycerol. The column was washed until absorbance dropped, then a gradient was run from 0-500 mM imidazole. There was an elution almost as soon as imidazole hit the column, but no further evidence of protein elution at higher imidazole concentrations.

SDS-PAGE analysis of the lysis and nickel purification steps revealed that the strong band corresponding to His-Ligase was present in the crude lysate samples but not in the supernatant. When the pellet resulting from ultracentrifugation was dissolved in 7 M urea, a strong band at ~60 kDa was seen, most likely corresponding to the His-Ligase construct. The pellet was white, which is indicative of the presence of inclusion bodies (IBs) of insoluble protein. Lysate pellets without inclusions are typically dark gray or brown.

4.3c Inclusion Body Purification

To attempt to purify His-Ligase from inclusion bodies, fresh cultures were grown and harvested as above. Cells were lysed by more harsh sonication for five rounds of 3 minutes on, 5 minutes off at output 5, duty 50% on ice. Lysate was ultracentrifuged as above to pellet cell debris and inclusion bodies.

Lysate pellet was resuspended in 50 mM tris pH 8.0, 1 M urea, 2% Triton X-100 and incubated with shaking for one hour. Supernatant was set aside. Following shaking, pellet wash solution was ultracentrifuged at 48,900 x *g* again. The resulting

large pale white pellet was resuspended in 50 mM tris pH 8.0, 7 M urea, 500 mM NaCl, 3 mM $MgCl_2$, 5 mM TCEP by pipetting up and down. The solution was incubated with sharking at 4°C overnight to fully dissolve inclusion body material. Following incubation, the urea solution was ultracentrifuged at 95,800 x *g* for one hour at 4°C to pellet any remaining insoluble material. A small pellet formed, about the consistency of a thick jelly.

Our lab has had success with on-column refolding previously[161]. On-column refolding is a relatively fast and robust method for removing urea and gradually allowing a protein to refold. Immobilization on-column also prevents aggregation that can occur during refolding processes. On-column refolding has the added bonus of not requiring extensive concentration steps upon completion, as the bound protein can be eluted with a small volume of solvent[162].

The urea solution supernatant was then loaded onto a cold pre-equilibrated nickel column. When all solution was loaded, the column was washed with 50 mM tris pH 8.0, 7 M urea, 500 mM NaCl, 3 mM $MgCl_2$, 5 mM TCEP until absorbance dropped. At that point, a gradient was run to remove the 7 M urea and add 5% (v/v) glycerol at 0.3 mL/min over 300 mL. Following this gradient, a second gradient was run from 0-500 mM imidazole to elute the bound protein from the column. There was no elution from the column at any point over the entire run.

SDS-PAGE (Figure 31) gel analysis reveals that the His-Ligase construct flowed through the column and did not stick, despite the presence of the N-terminal His_6 tag.

Nickel column flowthrough was subjected to dialysis against 200X volume 50 mM tris pH 8.0, 130 mM NaCl, 3 mM $MgCl_2$ overnight at 4°C. The resulting dialysate was extremely cloudy the next morning and was centrifuged briefly to pellet precipitate. The supernatant contained no protein, while the pelleted precipitate contained a strong gel band corresponding to His-Ligase at ~60 kDa (Figure 32). In order to test whether slow removal of urea was more amenable to refolding, the 7 M urea solution was dialyzed against a 4 M urea solution, leading to formation of precipitate following overnight incubation. No further incremental steps

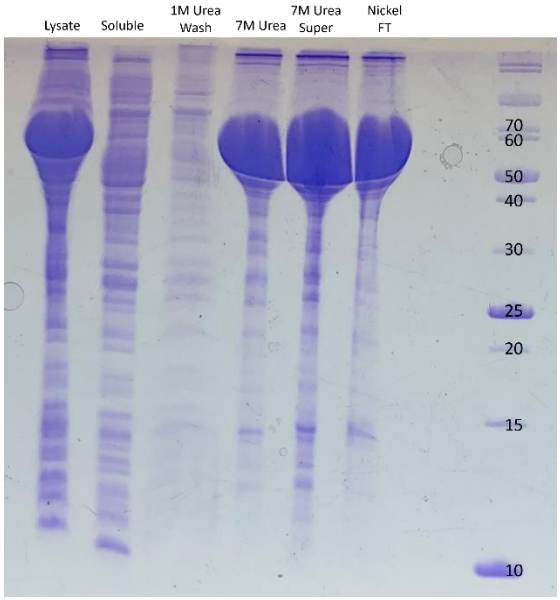

Figure 31- SDS-PAGE analysis of an on-column refolding for His-Ligase. The thick band around 60 kDa corresponds to the full weight of the construct. While this band is highly enriched in the raw lysate, none is present in the soluble fraction or upon washing with 1 M urea. The band is solubilized by addition of 7 M urea, but does not stick to the nickel column, even with an N-terminal His_6 tag.

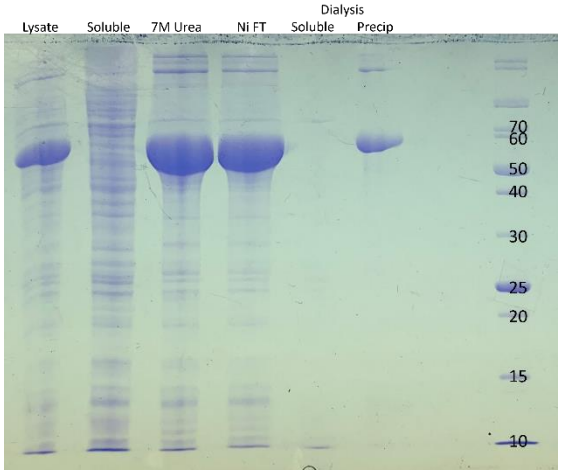

Figure 32-SDS-PAGE analysis of His-Ligase refolding by dialysis. There is no amount of the 60 kDa construct present in the soluble fraction of the dialysis.

were attempted, as the soluble fraction showed no evidence of protein by gel analysis.

4.3d Dilution Refolding Attempt

Since on-column refolding did not work at all, refolding by dilution was attempted. Refolding by dilution involves the rapid dilution of the denatured protein to avoid aggregation while simultaneously decreasing the amount of denaturant present in the local environment. Chaperone molecules such as glycerol, sucrose, arginine, and other amino acids, as well as redox reagents can be added to the refolding buffer to facilitate refolding. Refolding can be a relatively slow kinetic process, and thus it is vital to keep the protein solutions dilute to allow for proper folding. The primary disadvantages of this method are that the protein ends up extremely diluted, and the resulting concentration steps can take a tremendous amount of time, and that refolding buffers must be attempted purely through trial and error. There is no method for predicting which buffers or additives might facilitate efficient and correct folding of a given protein, and a buffer system that works for one protein might not work as well or at all for another protein[163, 164].

Several buffers were prepared for use in dilution refolding: 20 mM sodium phosphate pH 6.8, 200 mM NaCl, 3 mM $MgCl_2$; 20 mM sodium phosphate pH 6.8, 130 mM NaCl; 50 mM tris pH 8.0, 5 mM $MgCl_2$; 20 mM sodium phosphate pH 6.8, 130 mM NaCl, 5 mM $MgCl_2$; 20 mM sodium phosphate pH 8.0;, 20 mM potassium phosphate pH 6.8; 50 mM 4-(2-hydroxyethyl)-1-piperazineethansulfonic acid (HEPES) pH 7.6, 500 mM NaCl, 20 mM imidazole, 1 mM TCEP, 5% (v/v) glycerol.

His-Ligase urea solution was diluted 250X (200 µL into 50 mL) in each of the cold refolding buffers, pure water, and a 70% solution of perchloric acid (as a positive

precipitation control). No precipitate formed immediately upon addition to any of the refolding solutions, but precipitate did form when protein was added to water and perchloric acid. All solutions were incubated at 4°C for 30 minutes, at which point solutions turned cloudy. Solutions were centrifuged to pellet precipitate. Solutions at pH 6.8 contained a larger volume of precipitate than solutions at basic pH, suggesting that the His-Ligase construct is less amenable to slightly acidic conditions. Pelleted precipitate contained all traces of protein. No protein was detected in the supernatants by gel or by absorbance at 280 nm.

4.3e Size Exclusion-Mediated Refolding

Refolding by running a diluted, denatured protein solution over a size exclusion column (SEC) is thought to facilitate refolding. The slow flow rate and size separation allow for buffer exchanging and removal of the denaturant, while also theoretically separating misfolded, aggregated, or unfolded protein molecules from the properly folded molecules based on size[165, 166].

A small amount of the His-Ligase urea solution was diluted 10X with 50 mM tris pH 8.0, 130 mM NaCl, 10 mM $MgCl_2$. This slightly diluted solution was loaded onto a size exclusion column and run at 0.5 mL/min. There was a large elution peak around 40 mL, suggesting that a large protein had eluted. This was followed by a broad, low intensity elution over the following 40 mL (Figure 33). SDS-PAGE analysis revealed the presence of a band corresponding to the His-Ligase construct in the large elution peak, with essentially nothing in the broad elution peak (Figure 33). Fractions from the large elution peak were pooled and concentrated, but after concentration there was no absorbance at 280 nm, indicating that there was no protein

present. Nevertheless, a small amount of the concentrated “enzyme” solution was added to an NMR sample of ^{15}N UBact M17W with 1 mM adenosine triphosphate (ATP) and 1 mM lysine as a model substrate[121] at both pH 6.8 and pH 8.0. No shifts in UBact peaks were seen, nor was any change in the ^{31}P 1D spectrum of ATP, nor was there any change in the ^{1}H 1D spectrum that showed primarily ATP and lysine. If there were activity, it would be expected that signals in UBact would shift, and there would be definite changes in both sets of 1D experiments as ATP were consumed and lysine was attached[121]. The results were inconclusive, however, as no enzyme could be definitively detected following SEC.

4.3f Alternative Cell Lines

Since the basic BL21 DE3 cell line was showing expression into inclusion bodies, alternative cell lines were explored so try to slow down the production of enzyme so that it might remain in a soluble form, even at the cost of decreased yield.

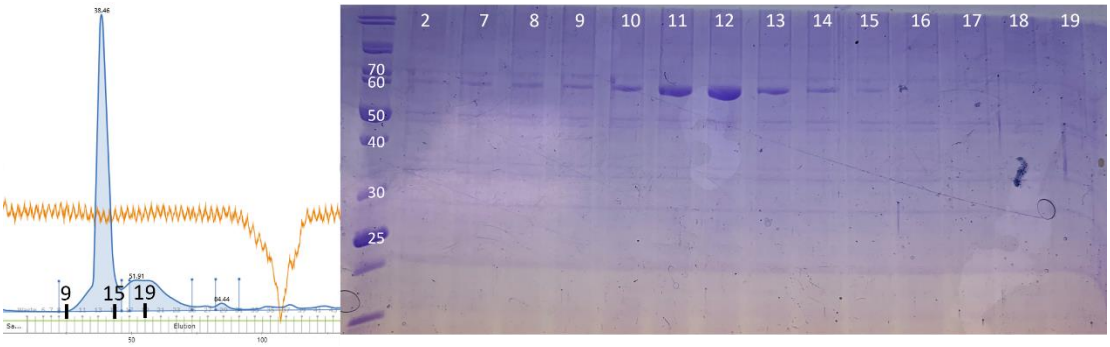

Figure 33- Chromatogram (left) and SDS-PAGE analysis (right) of size exclusion-mediated refolding of His-Ligase. The apparently large peak in the chromatogram has a maximum absorbance of 150 mAbs, and the maximal absorbance is at 38 mL, right at the void volume of the column. Gel analysis reveals that the primary species in this peak is the 60 kDa His-Ligase construct, and that there is essentially nothing in the broad, shallow peak that follows. Fractions showing the 60 kDa band were pooled and concentrated, but no absorbance or ligase activity was subsequently observed.

The pET15b His-Ligase plasmid was transformed into chemically competent BL21 Gold DE3 and BL21 DE3 pLysS cells as detailed above. Successful Gold cell transformation was selected for by growth on ampicillin-containing plates, and pLysS cell transformation was selected for by growth on ampicillin- and chloramphenicol-containing plates. Subcultures and large cultures were grown as above.

The Gold cell cultures yielded exclusively inclusion bodies at 200 µM, 500 µM, and 1 mM IPTG concentrations at 12-18°C. Thus, attempts with Gold cells were abandoned. pLysS cell cultures yielded no visible inclusion bodies at any IPTG concentration tested. Cells were lysed as described above, and clarified lysates were applied to a nickel column. An initial wash step showed a large elution (typically non-specifically bound proteins) that also included the His-Ligase, as determined by SDS-PAGE. A longer gradient showed no elution. Thus, it can be concluded that the N-terminal His_6 tag is at least partially occluded by other protein structures.

Since the nickel affinity purification did not yield pure protein, as might have been suggested by previous studies with homologous enzymes, the nickel flowthrough samples were exchanged into 20 mM tris pH 8.0 and applied to a QFF anion exchange column, which should retain the His-Ligase construct. His-Ligase eluted from the anion column in a broad peak along with numerous other proteins. While these samples were not very pure, they were concentrated and used to run another NMR test as above. Again, there were no changes in the ^{1}H, ^{31}P, or ^{1}H-^{15}N spectra following an overnight time course.

4.3g Lysate Tagging

The *M. tuberculosis* Pup proteasomal pathway has been reconstructed in *E. coli* and demonstrated to tag and degrade native *E. coli* proteins *in vitro*, despite *E. coli* lacking a proteasomal system[125]. In order to test if lysine was simply not a valid model substrate of the UBact ligase enzyme, and if the M17W mutation was significantly affecting ligase activity, and to potentially identify any substrates that might be relatively easy to purify from *E. coli* lysate, exogenous GST-UBact, 10 mM ATP, and 3 mM PMSF were added to a pLysS cell lysate expressing His-Ligase. The lysate was incubated for several hours at room temperature. After 70 hours

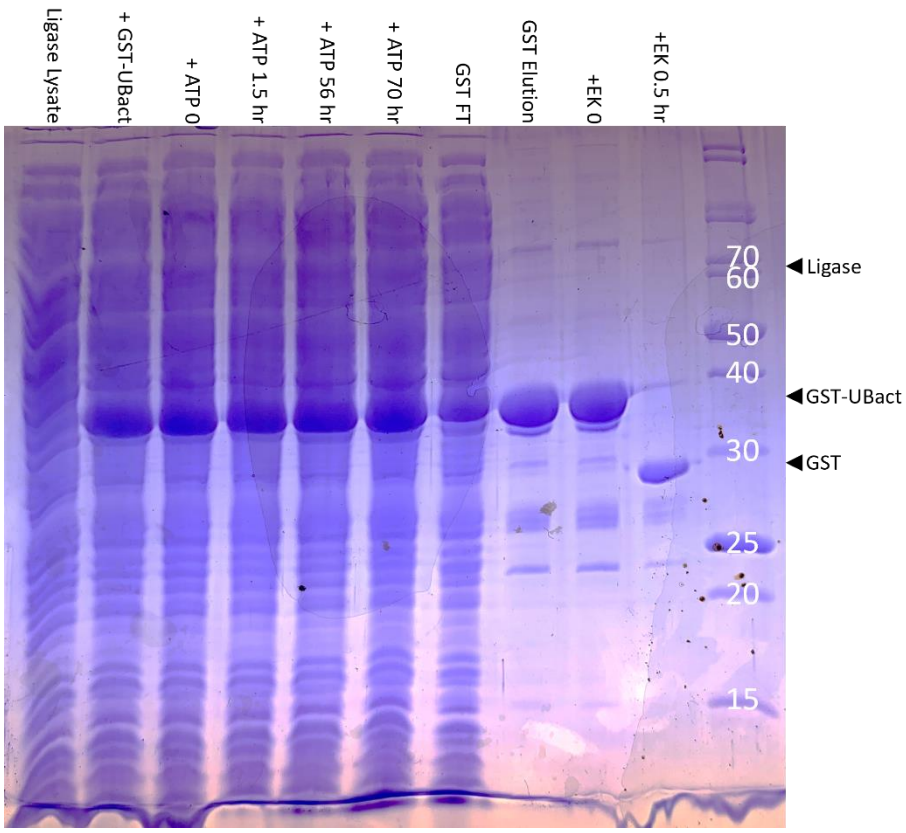

Figure 34- *E. coli* cell lysate tagging using exogenous GST-UBact as the tag and native *E. coli* proteins as substrate. Any soluble Ligase construct should be able to interact with the UBact construct and, assuming the enzyme is active, might tag specific proteins from the lysate. Catalytic amounts of soluble, properly folded enzyme should be enough to see any change in the bands that elute from the GSTrap column. The bands seen in the elution and enterokinase cleavage lanes are commonly seen products of GST interacting with SDS.

incubation, the supernatant was loaded onto a GSTrap column (GE Healthcare) to bind anything that might have been ligated to the C-terminus of UBact. PafA interacts with Pup along the C-terminal portion of Pup[167], and thus it is expected that an N-terminal GST domain will not significantly interact or preclude enzymatic activity in the analogous UBact system. Upon elution with glutathione, the GST domain was removed with enterokinase to provide better resolution of potential tagged substrates. SDS-PAGE revealed no additional eluted bands of protein, beyond small degradation products from the GST fragment (Figure 34). There was also minimal expression of

the His-Ligase construct. The lack of any discernable activity at any pH tested, using the model lysine substrate and ATP in massive molar excess, as well as testing on cell lysate and the resulting lack of tagging, suggests that the enzyme is still not properly folded (even if it is soluble), or does not participate in UBact ligation to a substrate.

In order to test if the His-Ligase construct actually encodes the putative deamidase (and thus the initial assumption of gene identity was incorrect), an E64Q mutation was introduced in UBact M17W to test deamidase activity. A SOFAST HMQC spectrum of ^{15}N UBact M17W/E64Q shows a new signal in the bottom of the spectrum, where a C-terminal signal is generally expected to appear, as well as the disappearance of the E64 signal and small shifts corresponding to other residues at the extreme C-terminus of UBact. Additionally, there are two new sidechain amide signals, from the incorporation of the glutamine sidechain (Figure 35). A fresh batch of His-Ligase was prepared by nickel affinity and anion exchange chromatography,

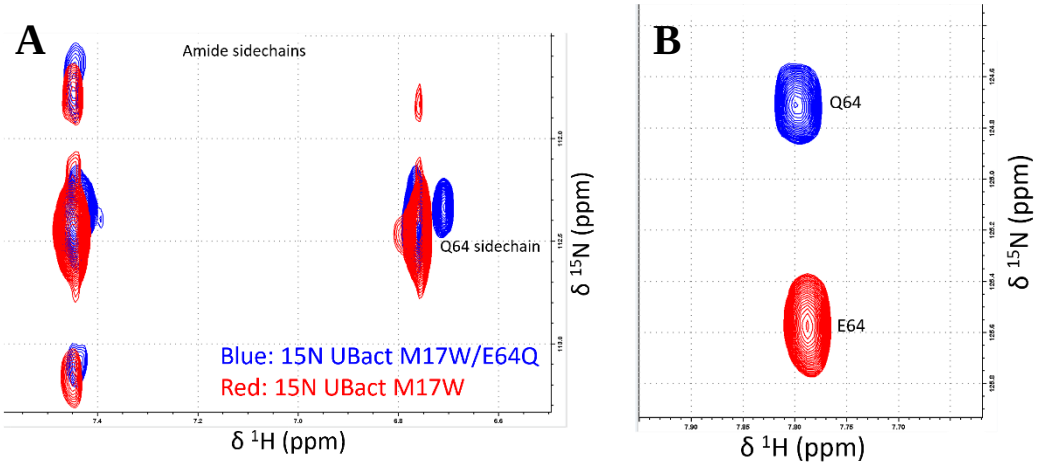

Figure 35- Comparison of the ^{1}H-^{15}N SOFAST HMQC spectra of UBact with (blue) and without (red) the E64Q mutation. A) The sidechain amide region shows the presence of two new signals, which must correspond to Q64. B) The Q64 C-terminal backbone amide signal is well-resolved relative to E64. The presence of unique resonances, and the presence of the additional sidechain signals, is important for tracking any potential enzymatic activity relating to the C-terminus, as Q64 and E64 would not be differentiable by mass spec.

then a small amount of the resulting concentrated solution was added to the sample of UBact M17W/E64Q. If there is deamidase activity, the C-terminal signal of Q64 would be expected to disappear and the E64 signal should reappear. Tests at pH 6.8 and pH 8.0, with and without 1 mM ATP and up to 2 mM lysine with both Q64 and E64 UBact species showed no change in any spectra, even after overnight incubation. Thus, it was concluded that this His-Ligase construct was not active.

4.4 GST-Ligase and GST-Deamidase Construct Tests

4.4a Construct Design

Since the His-Ligase construct was deemed to be inactive in both ligation and deamidation activities, new constructs were designed to ideally express both putative enzymes in a soluble form. GST-enzyme constructs with extended linker regions between the two domains should allow for the extreme solubility of GST to pull the enzymes into solution. The linker region should allow for native-like interaction with substrates, and activity can be tested both with and without the GST domain with the incorporation of specific protease cleavage sites.

E. coli-optimized gene sequences of both UBact ligase (Lig) and UBact deamidase (DEAM; Uniprot accession number A0A0S4L717) were synthesized by Genscript and cloned into the pET41a expression plasmid between the SpeI and XhoI restriction sites. The final construct contained an N-terminal GST domain, followed by a thrombin cleavage site, His_8 tag, enterokinase cleavage site, and the requisite enzyme gene. Each construct would have a molecular weight of ~85 kDa, which is convenient as the gel ladder utilized has a band of exactly 85 kDa, making

identification and tracking of expression and purification more straightforward. The slightly larger size also has the advantage of avoiding the influence of the 60 kDa GroL chaperone protein that is natively expressed in *E. coli* that might have unduly influenced the appearance of expression of the His-Ligase construct.

4.4b Expression and Purification Tests

Plasmids were transformed as above into BL21 DE3 and BL21 DE3 pLysS cells and grown as described above. pLysS cells appeared to display a small amount of protein with an 85 kDa weight by SDS-PAGE, but no material ever stuck to immobilized glutathione, either as beads (in batch or gravity column) or in a pre-packed column. Attempted ammonium sulfate precipitation and subsequent anion exchange chromatography never yielded any enrichment in purity of this protein, and NMR tests using crude fractions with lots of impurities never displayed any evidence of reaction. There was no change in UBact signals (for either E64 or Q64) and no change in ^{31}P or ^{1}H spectra under any conditions sampled. Thus, it was concluded that pLysS repression was too strong under all conditions tested, and that no meaningful amount of protein had been expressed in the cell line.

BL21 DE3 cells did show expression of protein, though the 85 kDa protein was trafficked to inclusion bodies in a similar fashion to the pET15b His-Ligase construct from above. No material from clarified lysate was able to stick to immobilized glutathione, and thus there was no trace amount of soluble material that might be concentrated for study. Thus, the primary goal of utilizing a GST fusion to increase solubility was for naught.

4.4c Refolding Attempts

Refolding attempts began again, with the thought that GST refolding might drag the rest of the protein along a proper folding pathway, or at least into solubility. In general, cells were lysed by heavy sonication similar to above, then the lysate was ultracentrifuged to pellet cell debris and inclusion bodies. The pellet was washed with buffer and 750 mM urea with 2% Triton X-100 to solubilize most of the cell debris, leaving only inclusion bodies after a further ultracentrifugation step.

Inclusion bodies were resuspended and solubilized in 50 mM tris pH 8.0, 500 mM NaCl, 7 M urea, 20 mM $MgCl_2$, 10% glycerol. This served as the starting point for numerous dilution tests. Refolding buffers tested included each combination of the following components: 50 mM tris pH 8.0, 130/500 mM NaCl, 500 mM L-arginine, 10% glycerol, 20 mM $MgCl_2$, 10-50 mM sucrose, 100 mM glutamate, 5 mM glutathione, and 5 mM TCEP. A gravity column was used to add protein urea solution dropwise to the refolding solution, diluting the protein urea solution 150X with stirring at 4°C. There was no apparent precipitation immediately upon dilution. Refolding solutions were concentrated, though this process was very time consuming, at which time there was significant irreversible precipitation.

In order to determine if the process of concentration was simply taking too long and the refolded enzymes were losing stability during that process, which could take several days, refolding solutions were concentrated using an Amicon stirred cell, which uses positive pressure to force the solution through a low molecular weight cutoff filter, in this case retaining anything above 10 kDa, while simultaneously stirring the cell to keep protein components relatively dilute. Retained sample showed

very little protein upon concentration, and the small amount that was present appeared to form a flaky white precipitate that could not be resuspended.

Rather than concentrating the refolding solution again, the solution was loaded onto a GSTrap column, with the thought that properly refolding GST fusion protein would stick to the column, and that properly folded GST might facilitate refolding of the enzyme domain. There was a very small elution from the column, though upon concentration this sample did not show any bands by gel, nor was there any apparent activity or interaction detectable by NMR. As an additional test for activity, samples were checked by mass spec, which again showed no change from the mass of UBact.

Since dilution refolding appeared to yield only precipitated material, N- and C-terminal His tags were incorporated to facilitate on-column refolding. The N-terminal $(HN)_8$ tag was introduced with non-overlapping primer extension mutagenesis. An HN-tag is more able to be expressed on the surface of a protein than the basic His tag and is able to bind to nickel more efficiently[168, 169]. The C-terminal tag was already present in the synthesized plasmid, though it was not expressed as part of the construct following a stop codon. Simply removing the stop codon allowed for addition of the C-terminal His tag. Immobilization on the nickel column could allow for removal of the denaturant over a gradient and allow for refolding followed by a relatively concentrated elution directly into a relevant buffer. Application of either His-tag variety did not appear to confer affinity to a nickel column following removal of any amount of denaturant.

Since urea-mediated denaturation did not yield any active protein, basic pH was utilized to assist in inclusion body solubilization. Basic conditions have been demonstrated to increase solubility without complete denaturation, and in some instances has been shown to allow for persistence of some secondary structure of proteins. In addition, lower concentrations of denaturant can be added to increase the solubilization of inclusion bodies without completely denaturing the resulting proteins.

100 mM tris pH 12.5, 200 mM NaCl was unable to efficiently solubilize the inclusion bodies of GST-Lig or GST-DEAM. The addition of 2 M urea allowed for complete solubilization of the inclusion bodies[170]. Removal of urea at pH 12.5 immediately led to precipitation. Dilution into a more neutral pH buffer immediately led to precipitation. Dialysis to incrementally lower the pH, both in the presence and absence of urea, lead to precipitation after overnight incubation.

Ammonium sulfate precipitation was attempted to salt out protein from urea solutions and facilitate refolding. While the GST-enzyme constructs did differentially precipitate from solution upon addition of increasing concentrations of ammonium sulfate (up to 2 or 2.5 M $(NH_4)_2SO_4$), protein did not resolubilize upon dilution with any of the refolding buffers tested.

Addition of organic solutes has been utilized to refold denatured proteins. Addition of molar concentrations of 1-propanol have been shown to assist in solubilization of some inclusion bodies while maintaining secondary structure elements of the protein[171]. Concentrations up to 6 M n-propanol were unable to solubilize GST-enzyme inclusion in this case. Dimethyl sulfoxide (DMSO) is also

useful for assisting in solubilization of a variety of chemicals, including proteins, yet concentrations up to 100% DMSO did not demonstrate any ability to solubilize the inclusion bodies. Use of three-phase portioning using ammonium sulfate and tert-butanol did not yield any fraction of soluble protein[172, 173]. The interfacial layer that should contain the precipitated protein was unable to be resolubilized using any methods beyond addition of high concentrations of urea (Figure 36).

Sodium dodecyl sulfate (SDS) has been shown to be an effective denaturant and is widely used in biochemistry to denature protein samples. Extremely low concentrations of SDS are required for denaturation, with as low as 0.4% (w/v) being utilized in SDS-PAGE. SDS binds directly to proteins, which can sometimes cause problems when trying to remove SDS from a solution. A 2% (w/v) SDS solution was utilized to assist in inclusion body solubilization at pH 8.0, without the need for high pH or high concentrations of urea. Lower concentrations of SDS did not produce efficient solubilization. Following solubilization, excess SDS was removed from the solution by the addition of 400 mM KCl, which forms insoluble SDS-KCl crystals. While SDS precipitation has been shown to precipitate significant amounts of the resolubilized protein, occasionally some properly refolded protein will remain in solution in the absence of solubilizing factors[174]. Addition of KCl to the SDS-solubilized inclusion body solution did immediately form precipitate, but no material was able to bind to immobilized glutathione following removal of supernatant. 2-methyl-2, 4-pentanediol (MPD) is often utilized as a precipitant in crystallography experiments and is generally thought to stabilize folded protein states by preferential hydration of the peptide backbone. MPD is also able to bind to hydrophobic

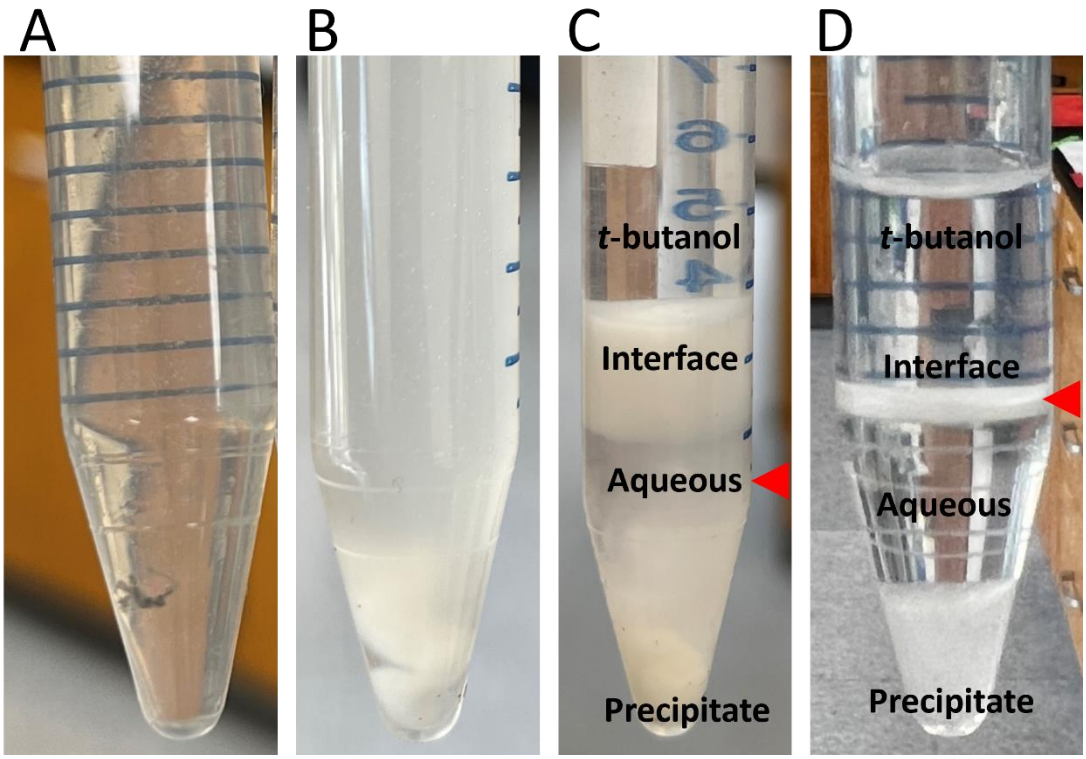

Figure 36-Three phase partitioning of denatured Lig/DEAM solutions. A) 7 M urea, 100 mM DTT solution of solubilized IBs. B) Upon addition of ammonium sulfate and an equivalent volume of *t*-butanol, the solution immediately became turbid, and material precipitated. C) Following shaking and centrifugation, the layers separated. *t*-butanol remained on top, the interfacial precipitate separated the organic solvent from the aqueous layer, and below the aqueous layer was insoluble precipitate. This precipitate could not be resolubilized even with addition of concentrated denaturants. The aqueous layer was removed and used for a second round of portioning. D) Addition of an equivalent volume of *t*-butanol, mixing, and centrifugation leads to well-defined separated layers. The interfacial precipitate has previously been shown to contain refolded, active proteins purified from inclusion bodies. In this case, material in the interfacial precipitate was not resolubilized except in the presence of high concentrations of denaturant. Smaller scale (~1 mL) and larger scale (50 mL) preparations did not yield any better results, not was active protein detected in any fraction.

sidechains, allowing it to act as a weak denaturant[175]. This ability to bind directly to hydrophobic surface area allows for competition between MPD and SDS for protein interaction, and MPD is able to decrease the binding of SDS to a denatured or partially denatured protein and can facilitate refolding following solubilization of inclusion bodies by SDS[176]. In this case, addition of concentrations of MPD up to 2 M

simply caused irreversible precipitation of the solubilized protein. No evidence of protein binding to immobilized glutathione was detected.

4.4dArctic Express Cell Line

Since all attempts at refolding yielded more precipitation and aggregation, another attempt was made with a new cell line, Arctic Express DE3 (Agilent). These *E. coli* cells express chaperones Cpn10 and Cpn60 from *Oleispira antarctica*, which assist in protein folding at lower temperatures than the native GroEL/ES system is capable of functioning. The Cpn10/60 system displayed high refolding activity at temperatures as low as 4°C. The GST-enzyme constructs were transformed into the chemically competent Arctic Express cells, and successful transformation was selected by growth on media containing 20 µg/mL gentamycin and 50 µg/mL kanamycin. Expression tests at 4°C showed no protein production at all, as well as dramatically slowed growth. Tests at 8 and 12°C yielded insoluble inclusions at all IPTG concentrations tested, including simply relying on constitutive leaky expression. An added disadvantage of this cell line was the co-precipitation of Cpn10/60 into the inclusion bodies. Co-purification of these chaperones is a common problem during soluble protein purification, though they can be efficiently removed by incubation with ATP and KCl[177] or low concentrations of urea[178]. However, the presence of the chaperones in the inclusion bodies simply presents another barrier to refolding by any means.

Attempts with all cell lines to utilize 0.5% and 1% (w/v) glucose solutions to suppress constitutive expression of both His-Ligase and GST-Lig/DEAM yielded inclusion bodies as well. Addition of 1M NaCl to large cell cultures did not yield any

soluble protein. Growth in M9 minimal media with 3.3 g/L $(NH_4)_2SO_4$ and 5 g/L D-glucose as the sole source of nitrogen and carbon, respectively, supplemented with 1 mM $MgCl_2$ in order to slow down cell growth and therefore production of protein resulted in reduced overall yield of inclusion bodies.

4.5 Alternative Constructs

4.5a Maltose Binding Protein

Since all attempts at refolding had thus far failed to produce soluble, active enzymes, it was decided to attempt to utilize maltose binding protein (MBP). MBP has been widely demonstrated to increase the solubility of fusion partners and has also been shown to act as a chaperone and assist in the folding of a fusion partner, possibly through interactions along its binding groove. MBP contains a signal peptide that encodes for excretion from *E. coli* cytosol into the periplasmic space, which can both slow down the folding process during transport across the membrane and deliver the protein to an oxidizing environment, allowing structurally important disulfide bonds to form. Deletion of the signal peptide prevents this export, but the high solubility and apparent chaperone activity of MBP are still useful in expression and folding of fusion partners. MBP also presents a single step affinity purification utilizing immobilized amylose, to which MBP preferentially binds. Following a wash

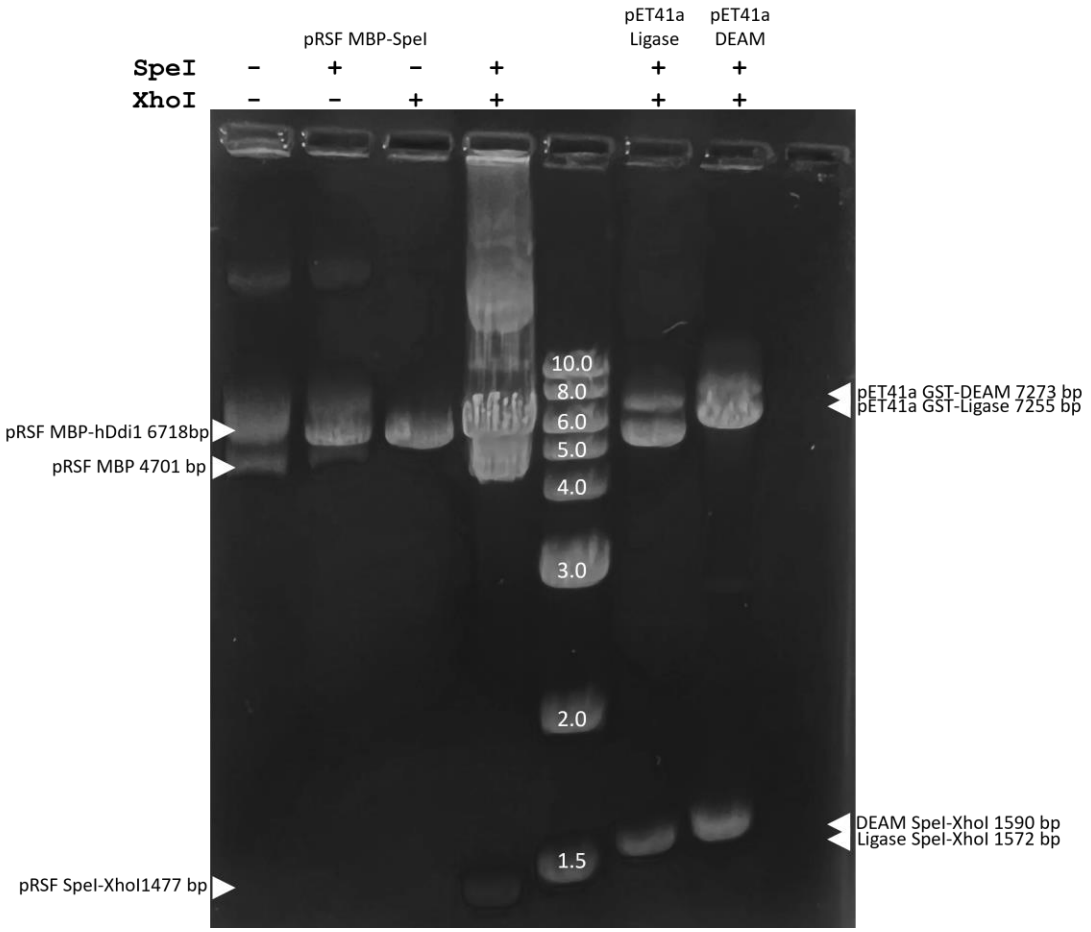

Figure 37- Agarose gel analysis of restriction enzyme cleavage of pRSF-MBP and pET-41 enzyme plasmids. SpeI and XhoI were able to efficiently cleave all three plasmids. The pRSF-MBP fragment and ligase/deamidase genes were well resolved. The relevant bands were excised for extraction and ligation.

step, the bound protein can be eluted with an excess of maltose to disrupt the interaction to the resin. This presents MBP as a desirable fusion partner.

The ligase and deamidase constructs were cloned into the pET41a vector between SpeI and XhoI restriction sites. A SpeI restriction site was mutated into a pRSFDuet-1 plasmid at the end of the MBP domain to facilitate insertion of the genes into the pRSF vector. pET41a GST-enzyme constructs and pRSF-SpeI were each incubated with SpeI and XhoI to release the genes and prepare pRSF for ligation. Reactions were separated by agarose gel electrophoresis (Figure 37), and individual

component bands were excised and extracted. Ligase and deamidase genes were mixed with the digested pRSF plasmid with T4 DNA ligase to ligate the overlapping ends of the DNA strands together. The reaction was transformed into chemically competent DH5α *E. coli* cells and successful transformation (and therefore ligation) was selected by growth on LB-agar plate supplemented with 50 µg/mL kanamycin. After numerous attempts, only a small number of colonies ever grew, and the only successful growth came from constructs prepared with the deamidase gene. Miniprepped pRSF-DEAM never returned the high yields of DNA commonly seen with the NEB Monarch system (~30-40 ng/µL pRSF vs. 150-200 ng/µL for other plasmids from the lab). Miniprepped plasmid also was never successfully transformed into any expression strain, suggesting that these growing colonies were some kind of contaminant. Experimental repeats with different concentrations of plasmid for digestion, differing load amounts of ligation, and differing transformation loads never yielded positive results. Fresh reagents did not produce positive results.

4.5b PGP_6 Fusion

Since the PGP_6 construct overexpressed so well, was well folded, and was extremely soluble, it might make a good fusion partner, dragging any fusion partners into solution. While there almost certainly will not be any chaperone activity, the high solubility might help, as well as the extraordinarily easy purification.

The pET28a His-PGP plasmid does not share a common set of restriction enzyme sites with the pET41a enzyme plasmids, but mutation of a single nucleotide will incorporate a BamHI site near the 5’ end of the enzyme genes. During mutagenesis, the primer bound to both the N- and C-terminal His tag sequences,

resulting in looping out the enzyme genes, essentially producing an empty pET41a vector with no multiple cloning region.

To avoid the His tag overlap, a SpeI site was incorporated into the 3' end of the pET28a His-PGP construct. This mutagenesis was successful, although attempted ligation with the excised enzyme genes was again unsuccessful.

4.6 Conclusions and Future Perspectives

All attempts to produce soluble, active *N. nitrosa* UBact ligase and deamidase enzymes were unsuccessful. Protein refolding is notably difficult and conditions producing favorable outcomes cannot be predicted with any certainty. Much of the literature concerning protein refolding conditions centers on the use of model systems such as lysozyme, which are purified in a soluble form and then denatured and refolded, or proteins such as insulin that are expressed in an insoluble form but have been shown to refold under a variety of conditions. Since positive results rely on a certain amount of luck to obtain properly folded, active protein, one can spend years attempting to find the optimal conditions that provide for the correctly folded proteins, a process that is costly in both time and reagents. Since no evidence of folded, active protein was ever found, it was decided to abandon this stage of the project.

In the future, if this project were to be revisited, a new expression system could be explored. First, since the inception of this project, a sister species to *N. nitrosa* has been isolated and cultured. *N. inopinata* cells are available, and growth conditions have been optimized. This species could present a prime opportunity for *in vivo* characterization of the UBact proteasomal operon, as well as provide a platform

for expression of each of the components (including the enzymes) under native conditions. The primary disadvantage of this species is that it is very slow growing, often taking several weeks to generate substantial biomass, and the culture does not grow turbid as the cells grow. Growth can only be indirectly observed by measuring the balance of ammonium, nitrite, and nitrate in solution. Thus, *N. inopinata* shows promise for *in vivo* studies, but does not appear to be a viable source of proteins in the amounts required for structural studies.

Saccharomyces cerevisiae is another model organism for protein expression, possessing the ability to make post-translational modifications that may be vital to protein folding and stability, as well as expressing chaperones capable of facilitating proper folding. *S. cerevisiae* is relatively easy to manipulate and grow and grows only slightly slower than *E. coli*.

A final species that could be utilized would be *Mycobacterium smegmatis*, a non-pathogenic, rapidly dividing cousin of *M. tuberculosis*. As a member of the *Mycobacteria, M. smegmatis* expresses the Pup proteasome system, and thus is already capable of expressing enzymes putatively similar to the UBact ligase and deamidase. Thus, it is not a stretch to imagine that *M. smegmatis* would have little trouble expressing soluble, folded UBact-system enzymes. *M. smegmatis* is used in many labs as a model organism, and protocols are widely available. *M. smegmatis* grows more slowly than *E. coli* but can grow in a variety of media conditions and still reaches high cell density in as little as three days (as compared to several weeks for *M. tuberculosis).*

4.7 Contributions

Sashika Fernando participated in several attempts at expression and ammonium sulfate precipitation of the enzyme constructs. Steven Bonn designed the project, performed the experiments, and analyzed data. David Fushman oversaw the project.

Chapter 5: NMR as a Tool to Study Deubiquitinase Activity

5.1 Objective

Utilize NMR to determine the relative rates and directionality of disassembly of polyubiquitin chains by isopeptide bond cleavage.

5.2 Different Deubiquitinase Enzymes Serve Different Physiological and Pathophysiological Roles

5.2a OTUB1

Otubain 1 (ovarian tumor ubiquitin-aldehyde binding protein 1; OTUB1) is a DUB expressed in a wide range of tissues. Increased OTUB1 activity is associated with a variety of cancers, and a large body of work is dedicated to elucidating its role in cancer pathology[179-181]. Identified in 2002[182], OTUB1 is a DUB with specificity for cleaving K48-linked polyUb chains[183], while other DUBs in the OTU family display different linkage preferences (or little preference at all)[103]. The DUB recognizes two ubiquitin units at distinct sites, with the isopeptide bond located near to the active site to facilitate isopeptidase activity[183]. There are conflicting reports about OTUB1's DUB activity and whether it is able to remove a single ubiquitin moiety or if it must cleave a diubiquitin unit. OTUB1 DUB activity has been identified in numerous pathways, with OTUB1 disassembling K48-polyUb from conjugated substrates and preventing their proteasomal degradation. OTUB1 prevents apoptosis by stabilizing the cellular inhibitor of apoptosis complex by removing polyUb chains that would otherwise encode for its destruction, and thus is able to help cancerous cells avoid

programmed cell death[184]. It can also remove polyUb from the intracellular domain of PD-L1, preventing its degradation and allowing a cancerous cell to escape immune detection[185].

In addition to proteolytic DUB activity, OTUB1 also interacts non-covalently with Ubc13 to inhibit propagation of K63-linked polyubiquitin chains during the DNA damage signaling cascade[186]. OTUB1 appears to bind to a variety of E2-Ub conjugates to prevent ubiquitination by mimicking the product state of K48-polyUb DUB activity in the active site of OTUB1. The DUB inserts itself to disrupt the E2-Ub interface, though Ub remains covalently bound to the E2[187]. Interestingly, several crystal structures of OTUB1-E2~Ub complexes show a free ubiquitin binding to OTUB1 in the same position that a K48-linked distal unit would be positioned, and modeling of a K48 isopeptide bond is readily possible in those structures without disrupting any main chain structures[183, 187]. Functional studies suggest that OTUB1 shows a preference for cleaving from the middle of chains to release a diubiquitin unit[188]. Thus, OTUB1 can play both an active and passive role in downregulating ubiquitination of several varieties. This wide-ranging canonical and non-canonical DUB activity, and OTUB1's implication in numerous cancers makes OTUB1 a prime target for drug development[189].

5.2b IsoT

Isopeptidase T (IsoT; Usp5) was discovered in 1992[190] and thoroughly characterized in 1995[191]. IsoT shows a high preference for cleaving polyUb chains with a free C-terminus on the proximal-most Ub subunit, and any modification of G76, either by attachment to a substrate or by chemistry, dramatically reduces

cleavage activity[191]. IsoT is capable of cleaving all polyUb linkages[192], though it prefers unanchored K48 chains[193]. IsoT is vital to maintaining appropriate levels of free ubiquitin by breaking up these unbound chains, allowing free ubiquitin to be recycled and utilized by the cell rather than repeatedly degraded and re-translated.

IsoT has been implicated in a variety of pathological conditions. The *Drosophila* homologue of IsoT is required for proper development, especially of the eyes[194], and mutation/deletion is eventually lethal due to improper ubiquitin homeostasis[195]. IsoT is required for efficient DNA double stranded break repair, where the enzyme is responsible for disassembling polyUb chains at the site of DNA damage[196]. IsoT has also been identified as a possible therapeutic target in several different types of cancer. IsoT is dramatically overexpressed in hepatocellular carcinoma (HCC), and is able to protect the SLUG transcription factor, which facilitates HCC's ability to migrate and metastasize[197]. Similar to OTUB1, highly elevated IsoT levels lead to deubiquitination and stabilization of PD-L1, facilitating non-small cell lung cancer propogation[198]. Knockdown of elevated IsoT levels using siRNA in pancreatic cancer sensitizes the cell to apoptosis due to cell cycle blockage at G1/S[199]. IsoT has also been associated with the experience of pain and is thus a new target for pain management therapies[200].

5.2c Ubp6

S. cerevisiae Ubp6 (homologue of human Usp14) is a proteasome-associated DUB[201, 202]. Ubp6 contains an N-terminal UBL domain and a C-terminal USP catalytic domain[203]. The UBL domain is responsible and sufficient for association to the proteasome, where it interacts with the base of the 19S RP, where it binds to

Rpn1[204]. Ubp6 also interacts primarily with Rpt1, where it binds and precludes Rpn11 access to ubiquitinated substrates, thereby reducing Rpn1 DUB activity[205].

Recombinant Ubp6/Usp14 display low levels of activity in the absence of the proteasome, despite their active site catalytic triad being positioned for proteolysis. Structural studies reveal that there are two loops, BL1 and BL2, which overlap the active site and preclude ubiquitin access to the catalytic triad[203]. These two loops shift upon binding to the Rpt base of the proteasome, and ubiquitin binding to Ubp6/Usp14 is stabilized in this conformation, which in turn further enhances Ubp6/Usp14 affinity for the proteasome[206, 207]. Full DUB activation is only achieved upon interaction with the proteasome. Interestingly, (poly)Ub or inhibitor-bound Ubp6/Usp14 stimulates proteasomal ATPase activity[205] and activates degradation by facilitating proteasomal gate opening, allowing easier access to the catalytic 20S CP[208]. Conversely, there is evidence that Ubp6 also slows down the processing of ubiquitinated substrates, possibly acting as a timer to allow for removal of ubiquitin units beyond a certain critical length prior to Rpn11 engagement[205]. In the context of binding to the 19S RP, Ubp6 preferentially cleaves the distal ubiquitin units from a polyUb chain[202, 209].

Since Ubp6/Usp14 acts immediately upstream from the final steps of proteasomal degradation, dysregulation is associated with a huge number of disease states. The DUB has roles in gluconeogenesis[210], insulin resistance[211], and other diabetic pathologies[212, 213].

5.2d Usp2

Usp2 is a DUB that has been implicated in facilitating the spread of prostate cancer[214]. Several splice variants have been identified, with N-terminal regulatory

elements ranging from six to 258 amino acids followed by the catalytic 347 amino acid USP domain[215]. Ubiquitin binds inside a defined pocket on Usp2. The bound ubiquitin unit represents the distal unit in a ubiquitin chain, with the C-terminus resting inside the active site[215]. Modelling suggests that deubiquitinase activity requires an open conformation of K48-linked diubiquitin[216, 217]. Involvement of the distal ubiquitin unit in binding, and the general exclusion of the proximal unit suggests that Usp2 might cleave polyUb chains from the distal end inward.

Physiological Usp2 expression is CLOCK-regulated in both liver tissue and the suprachiasmatic nucleus (location in the brain facilitating circadian rhythm), indicating that Usp2 probably plays a role in maintaining circadian rhythm throughout the body[218, 219] by protecting BMAL1 from proteasomal degradation[220].

Usp2 is overexpressed in a variety of cancers. Cyclin D1 degradation can be prevented by Usp2-mediated deubiquitination, allowing unchecked advancement through cell division[221]. Usp2 protects fatty acid synthase (FAS) from ubiquitin-mediated degradation, which allows overexpressed FAS[222] to continue synthesis of fatty acids and assist in prostate cancer propagation and survival[223] by inhibiting apoptosis[224]. Usp2-mediated FAS stabilization was implicated in glioma [225]. Indeed, Usp2 has been identified as a biomarker in some types of cancer[226-228]. Usp2 also influences cellular apoptosis. Knockdown of isoform Usp2-1 by siRNA increased apoptosis in prostate cancer lines[229], while upregulation of isoform Usp2-2 increased apoptosis in HeLa cells[230]. Usp2 knockout mice displayed infertility due to reduced sperm motility and failure to penetrate the ovum, possibly due to proteasomal dysregulation[231]. Usp2's role in such a wide variety of cellular processes, and its

significant relation to cancer pathologies, suggest that it could be a viable chemotherapeutic target for several different cancer types.

5.2e PLpro

The papain-like protease (PLpro) from SARS-CoV 2 (SARS2) is one of two proteases encoded by the viral genome, along with the chymotrypsin-like main protease (3CLpro/Mpro). Part of non-structural protein 3 (Nsp3) of the SARS2 proteome, PLpro participates in viral maturation by cleaving three nonstructural proteins from the larger viral polypeptide. Interestingly, PLpro recognizes the site LXGG|XX for cleavage, which is identical to the C-terminus of ubiquitin and many ubiquitin-like proteins. In addition to its role in viral maturation, PLpro serves a secondary role helping PLpro to avoid frontline immune responses by acting as a deubiquitinase and de-ISGylase. This secondary activity serves to protect SARS2 proteins from degradation by the proteasome by recognition of K48-linked polyUb chains, and to prevent signaling by the attachment of ISG15 (interferon-stimulated gene 15), a protein consisting of two UBL domains that can be attached via isopeptide bond to a substrate lysine residue in a similar fashion to ubiquitin. While the ISG15 signaling pathway has not been definitively explored, it is thought to be involved in immune signaling and serves a protective role in the cell. SARS-CoV2 PLpro is analogous to the PLpro proteins from both SARS CoV and MERS-CoV, with all three PLpro species displaying slightly different preferences for K48-polyUb or ISG15 recognition. While the worldwide push to develop vaccines for SARS2 has been largely successful in stemming the tide of widespread infection, there is still a need to treat acute SARS-Coronavirus disease (Covid) infections. Mpro has been the

primary target of drug discovery efforts, given its importance in cleaving the vast majority of the viral polypeptide into its constituent individual proteins. The approval of Paxlovid from Pfizer as the first oral antiviral treatment for Covid proves that the viral proteases are viable targets for therapeutics. The search for additional treatment modalities continues, with PLpro presenting another promising target. Disrupting both maturation and immune escape (via DUB and de-ISGylase activity) would provide additional protection in the event of acute Covid. As an individual or combination therapy, inhibition of PLpro could provide for a secondary treatment against new strains of SARS2 as it might evolve to avoid either the current suite of vaccines or Paxlovid treatment. In addition, the discovery of effective PLpro inhibitors could provide a head start for treatment discovery of the next coronaviral pandemic, should one occur.

SARS2 PLpro specifically recognizes the diubiquitin moiety across its two recognition sites, with the proximal ubiquitin domain binding to the S1 site, the distal ubiquitin domain binding to the S2 site, and the isopeptide bonded substrate interacting through the active site and across the S1* site. PLpro specifically removes diubiquitin units from viral protein substrates, preventing their recognition and degradation by the proteasomal machinery. Though the understanding of polyUb attachment to SARS2 substrates is incomplete, protection of the viral proteins from proteolysis could play an important role in the viral lifecycle.

5.3 NMR Can Track Multiple Cleavage Sites Simultaneously

5.3a Synthesis of Triubiquitin Chains

For ease of discussion, K48-linked Ub_3 with ^{15}N-labeled middle Ub will be referred to simply as mid-^{15}N-Ub_3 in this chapter. Proximal, middle, and distal are used to denote the relative positions of various ubiquitin units in a chain (Figure 38), where proximal is the unit expected to be closest to a substrate (directly connected via an isopeptide bond), and distal the furthest away in a chain from the substrate.

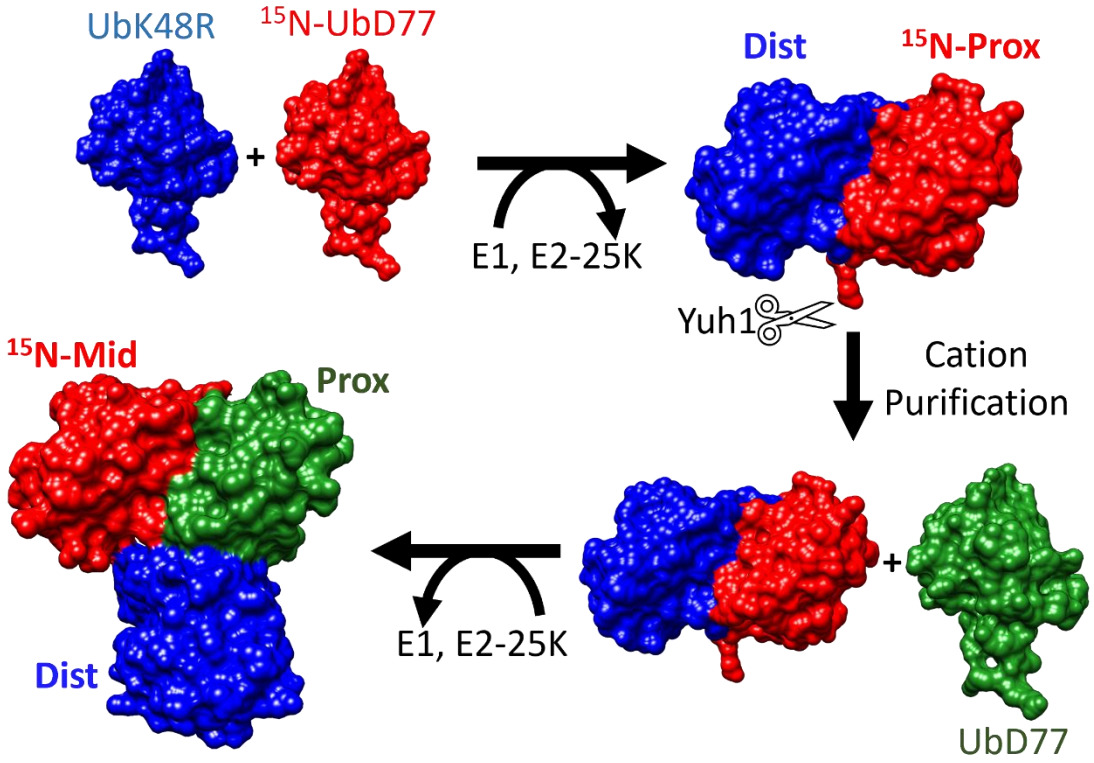

Figure 38-Schematic of mid-^{15}N-Ub_3 chain synthesis. Unlabeled UbK48R (blue) and ^{15}N-UbD77 (red) are conjugated with E1 and E2-25K, the reaction is quenched with DTT, then the dimer is subjected to digested with Yuh1 to remove the C-terminal D77 blocking residue. After incubation, the dimer is purified from the component monomers by cation exchange at pH 4.5. The dimer with free C-terminus is then exchanged into 50 mM pH 8.0 buffer for another round of conjugation with an unlabeled UbD77 (green) moiety. The trimer, with the ^{15}N-labeled ubiquitin unit located at the middle position, is then purified by cation exchange again and exchanged into the reaction buffer and stored for use.

In order to study DUB-mediated cleavage at multiple sites at the same time, mid-^{15}N-Ub$_3$ was synthesized through sequential enzymatic addition of ubiquitin subunits[217]. Briefly, ^{15}N Ub with a C-terminal D77 residue was enzymatically linked to unenriched UbK48R (with free C-terminal G76) with E2-25K. The reaction was quenched with addition of 10 mM DTT, and catalytic amounts of *S. cerevisiae* ubiquitin C-terminal hydrolase 1 (Yuh1) were added to remove the D77 residue. The resulting K48-Ub$_2$ with free G76 was purified by cation exchange, and then was reacted with unenriched UbD77 to form mid-^{15}N-Ub$_3$, which was subsequently purified by cation exchange and concentrated, and buffer exchanged into DUB reaction buffer (50 mM tris pH 7.5, 100 mM NaCl, 1 mM TCEP). Purity was verified by SDS-PAGE.

5.3b Proximal and Distal Isopeptide Bonds can be Tracked

In a ^{1}H-^{15}N SOFAST spectrum of mid-^{15}N-Ub$_3$, reporter groups for the proximal and distal isopeptide linkages are visible. The signal for G76 reflects its isopeptide-bonded position and is shifted upfield in the nitrogen dimension relative to its free (non-isopeptide linked) position, which reports on the proximal isopeptide bond. The distal isopeptide bond is directly observed, as the signal of the ^{15}N ε-amide appears at 7.91 ppm in the ^{1}H dimension and 120.5 ppm in the ^{15}N dimension. Additionally, the isopeptide-bonded K48 backbone signal shows a distinct shift from its unconjugated resonance (Figure 39). R74 also shows a shift upon isopeptide formation, making this a viable secondary reporter group. Upon cleavage of either isopeptide bond, their corresponding reporter signals will decrease, and the free-state signals will appear and increase in intensity as cleavage proceeds. Thus, mid-^{15}N-Ub$_3$

is a prime model substrate to track consecutive cleavage events in one pot, with reporter groups for both decrease of isopeptide-conjugated substrate and appearance of free ubiquitin (Figure 40).

5.4 Quantification of DUB Cleavage and Determination of Directionality

Since NMR is a relatively slow technique, taking tens of minutes to acquire single time points, care was taken to utilize enzyme concentrations that would both be fast enough to observe cleavage while still being slow enough to gather several time points with enough signal to noise to accurately quantify the rates of cleavage. To this end, several tests were run with varying concentrations of the enzymes to be studied

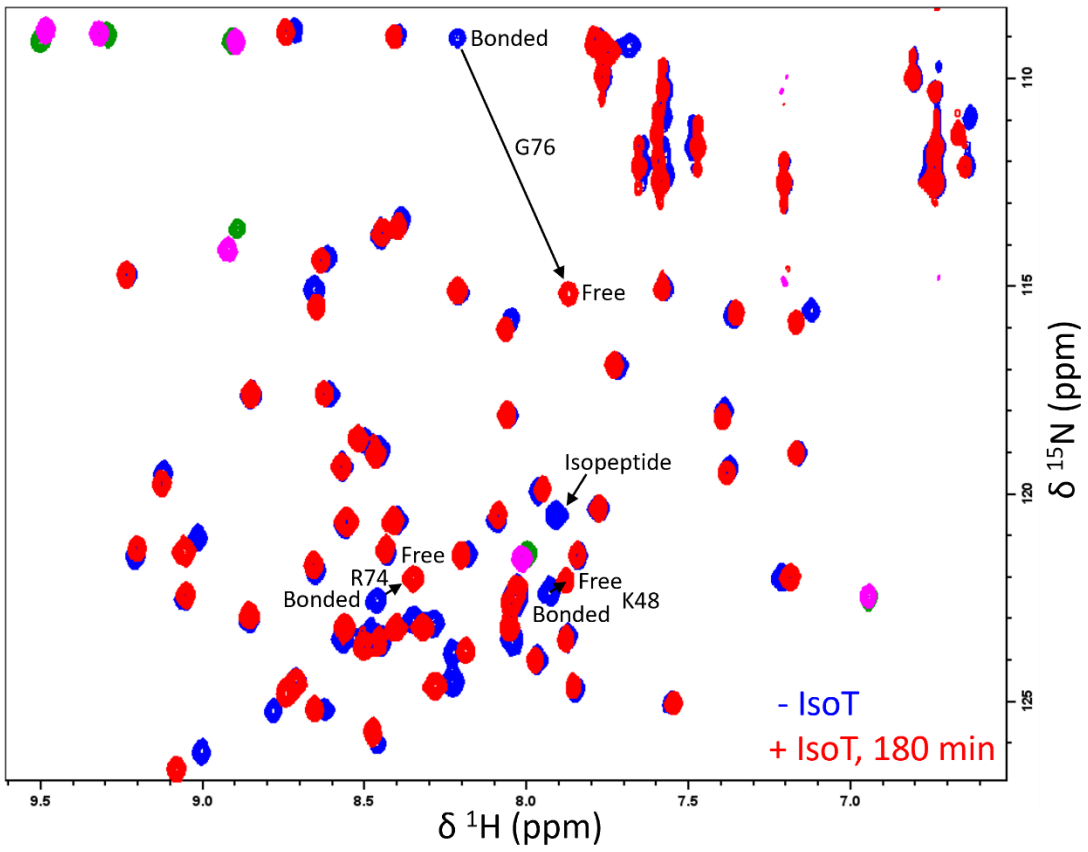

Figure 39- SOFAST HMQC spectra of the start (blue) and finish (red) of IsoT digestion, highlighting the signals used for rate determination. R74 and G76 report on proximal bond cleavage, while K48 and the isopeptide signals report on distal bond cleavage.

to determine optimal conditions. Unlabeled K48-Ub_2 was incubated with the enzymes and time point samples were run on a gel. Enzymatic activity drops over time, so there was a balance to be struck between the long measuring times required of NMR and the lifespan of the enzymes.

In order to quantitatively study DUB cleavage, catalytic amounts of the various DUBs described above were added to 100 µM mid-^{15}N-Ub_3 samples, and SOFAST spectra were acquired at various time points. As expected, reporter signals for both isopeptide linkages decreased over time, while the free signals increased (Figure 41).

Qualitatively, IsoT displayed the fastest cleavage, despite being at the second lowest concentration of any of the DUBs assayed. Ubp6 showed the slowest cleavage, despite having the highest concentration of enzyme utilized. PLpro was the only enzyme that did not appreciably cleave one of the bonds, preferentially cleaving the proximal isopeptide and only cleaving the distal after almost all of the proximal bonds had been cleaved.

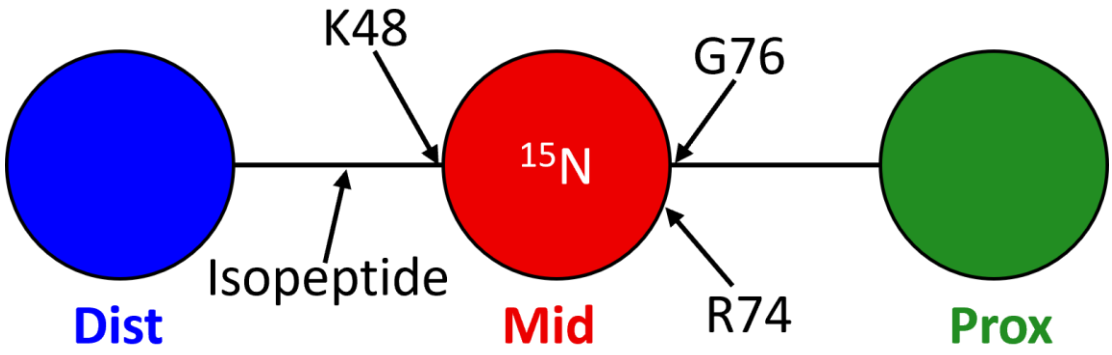

Figure 40- Schematic of residues used to track deubiquitinase progress. The middle unit of the K48-linked Ub_3 is ^{15}N-labeled. Changes in G76 and R74 bonded and free backbone amide resonances track cleavage of the isopeptide bond between the middle and proximal units, while changes in K48 backbone amide and isopeptide resonances track cleavage of the isopeptide bond between the distal and middle units.

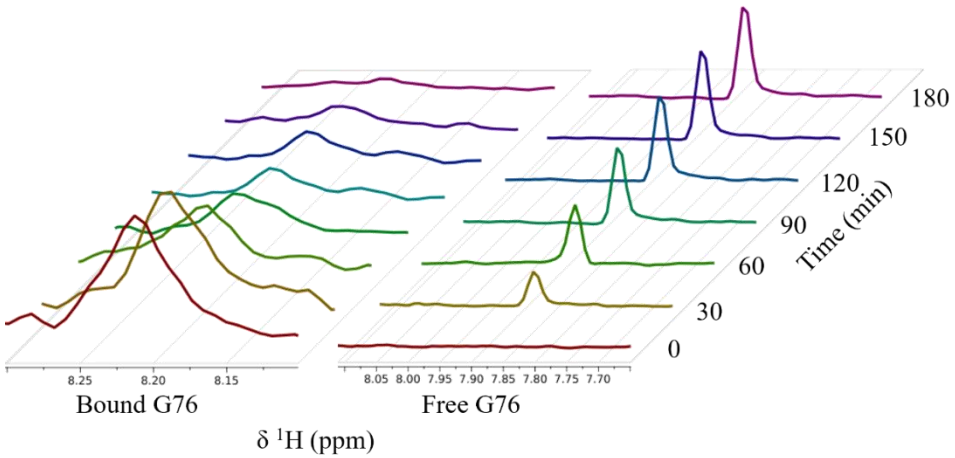

Figure 41- Representative 1D slices of G76 in isopeptide-bonded and free states during digestion with IsoT. Intensities are adjusted for visual simplicity, as the bonded G76 signal intensity is dramatically lower than free G76 intensity.

5.4a Intensity-Derived Results

NMR signal intensities are often used as a stand-in for the number of spins (and therefore molecules) present in a given state, as they are easy to quantify and readily observable by eye. This holds true when the species under observation does not undergo any large changes in T_2 relaxation, as is the case with most small molecules or enzymatic reactions where only specific residues are affected. Large changes in transverse relaxation alter the signal linewidth, and thus the observed intensity, rather than necessarily reflecting a change in population. Smaller molecules tumble faster in solution, producing generally longer T_2 relaxation times (lower relaxation rates), and therefore sharper signals.

Unfortunately, when using mid-^{15}N-Ub_3, DUB cleavage results in a dramatic change in relaxation rates as Ub_3 is cleaved into Ub_2 and just Ub in the end. At each step, molecular tumbling becomes faster, giving rise to slower relaxation rates and thus sharper signals. Thus, upon initial cleavage bonded signals actually increase in

intensity as tumbling speeds up, at the same time the bond the signals report on is actually cleaved.

In order to compensate for this apparent increase in intensity, all signals were scaled using the intensity of Ile30, a well-resolved signal that does not shift upon interaction with a DUB or shift upon cleavage from Ub_3 to Ub_2 to Ub. Ile30 is a buried residue that is spatially distant from site of most Ub interactions, and thus appears unaffected by any factors other than changing relaxation rates due to change in tumbling. Applying this scaling factor appears to compensate for changes in relaxation among the three polyUb species under observation.

5.4b Volume Derived Results

In contrast to signal intensities, NMR integrated signal volumes are directly proportional to the number of spins present in a given state. Assuming all other experimental parameters are identical across a set of experiments, any difference in volume is due exclusively to a change in the population. Even in a scenario where relaxation rates are changing dramatically, as in cleavage of Ub_3, the linewidth difference between each state would not affect the total integrated volume of a signal. Thus, even though a signal might change intensity because linewidth changes, the volume would remain the same between the two signals. This fact makes volume integration a powerful tool for studying kinetics. Integration of 1D signals is a routine part of basic NMR interpretation to determine population sizes and solve the structures of small molecules, but 1D signal integration is problematic in larger molecules that show many overlapping signals at lower resolution, such as proteins. Multidimensional NMR can improve resolution by separating signals into multiple

dimensions, but integration of multidimensional NMR signals is a complex undertaking. While 1D integration can be done mathematically by fitting the lineshape and integrating, performing this calculation in two dimensions is more complicated, and most programs simply use a box sum of intensities to estimate volume. This might be passable in scenarios where relaxation rates are not significantly different between reactant and product, but in the case of Ub_3 degradation, it presents the same problems as simply using calculated intensity, as the box sum of intensities are still influenced by changes in relaxation.

In order to overcome the challenge of multidimensional signal integration, several different programs were utilized to fit the peaks and mathematically integrate the signals. Most programs did not produce reproducible results or results that made sense, or the process of peak selection and integration was too cumbersome and time consuming to be viable for multiple experiments. Notably, Bruker Topspin signal integration gave reproducible results, but at an immense time cost and no straightforward way to export those integrations. One program claiming to integrate signals by mathematically fitting the lineshape produced volumes up to 100-fold lower than the calculated intensity of the same peak.

In the end, NMRFAM-SPARKY[232] proved to be the most useful program for generating reproducible volumes. Spectra from each enzymatic time course were processed identically in Topspin using a Gaussian window function, then were opened in Sparky. Individual peaks were selected and iteratively integrated as Gaussian functions. Sparky includes a feature to remove integrated peak volumes from a spectrum, and thus accurate integration can be directly observed if all apparent

peak volume is removed. By mathematically integrating peaks, Sparky is also able to avoid including contributions from nearby overlapping peaks. Peaks were iteratively integrated until all visible volume was removed from the spectrum, even at lowered contour levels.

Integrated peak volumes from the first five time points collected were fit to a first order exponential decay curve (Figure 42):

$$[Ub_3] = Ae^{-kt} + C \tag{12}$$

where [Ub_3] is the amount of Ub_3 at a given time point, A is the initial concentration of mid-^{15}N-Ub_3, k is the first order rate constant, t is the time, and C is a baseline constant. Notably, there is a disagreement between the disappearance of isopeptide-linked signals and the appearance of free signals, as it appears free signals emerge dramatically faster than bonded signals disappear. In addition, since the total

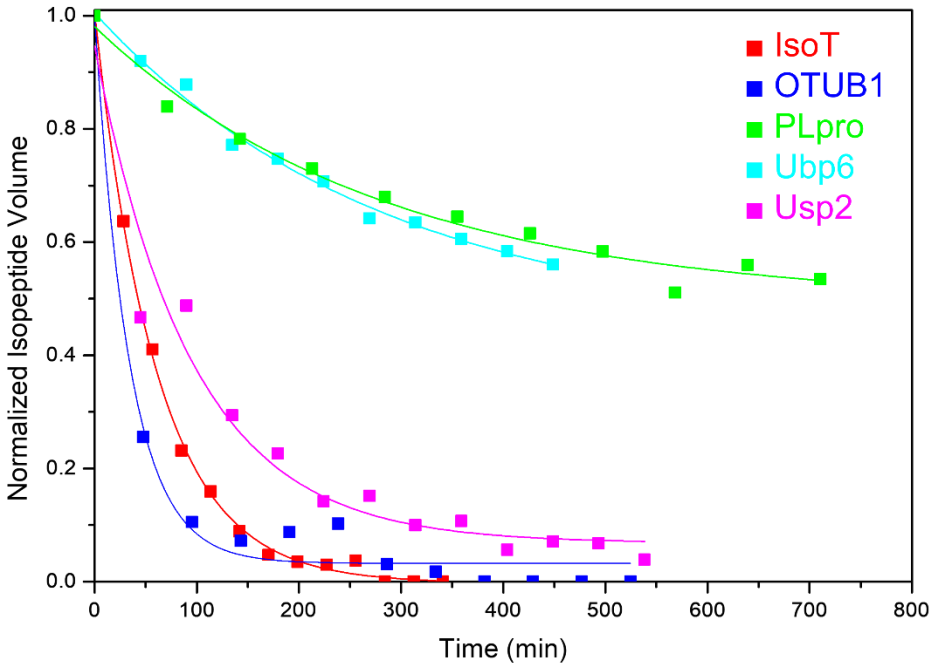

Figure 42- The NMR signal volume decay can be fit to an exponential decay curve. The normalized isopeptide bond signal volume at each time is shown as an example from each DUB experiment, with fit exponential decay curves.

population of the system does not change, only the bonded vs free ratio, the additive sum of the bonded and free volumes should remain constant. Instead, the sum of the volumes increases over time until it reaches a plateau.

This apparent increase in population almost certainly reflects the influence of changing relaxation rate, though this time it is overly influenced by a change in T_1 relaxation. The SOFAST experiment is designed with a shortened recycling delay to allow for more scans per unit time than the comparable HSQC experiment. The disadvantage of this method is that it does not allow for full regeneration of *Z* magnetization between scans, and thus the changing T_1 between Ub_3, Ub_2, and Ub is able to influence the apparent signal. Unfortunately, utilizing a longer recycling delay would dramatically decrease the number of scans possible, and thus render signals too low or time intervals too long to accurately quantify a rate.

To overcome this, the bonded signal volumes were again scaled by the Ile30 volume. This allowed for compensation for changing T_1 relaxation rates, and generally led to agreement between bonded signal disappearance rates and free signal appearance rates (Figure 43 and Appendix Table 3 and 4). The free signals were not scaled, as they are already influenced by the change in relaxation. All signals were normalized to the largest volume signal, and free signals that appeared were rearranged into a decay curve by subtracting the normalized volumes from 1. These values were fit to the exponential decay curve (Equation 12), and their rates plotted as a function of reporter group (Appendix Table 4).

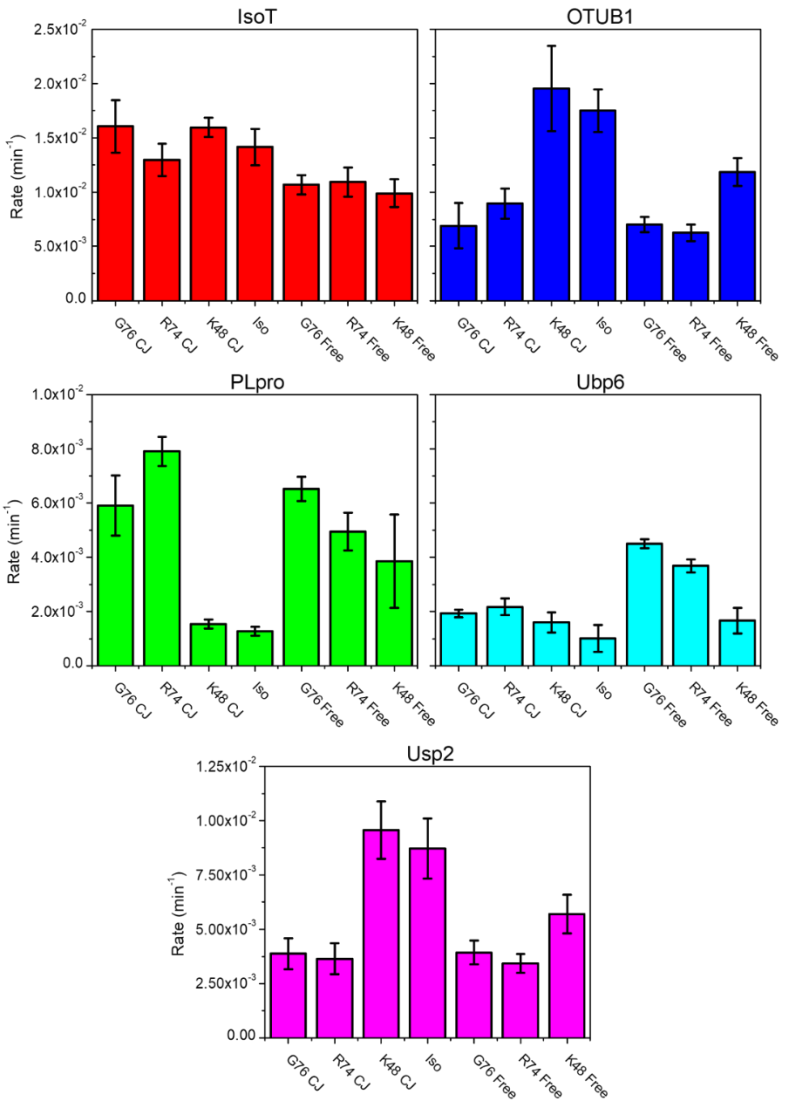

Figure 43- Relative cleavage rates by various DUBs, as measured by time-resolved NMR using volume integration. Importantly, there is agreement in the rates of disappearance of a conjugated signal and the appearance of its corresponding free/unbound signal, showing that both are appropriate reporter groups. Additionally, conjugated R74 disappearance and free R74 appearance rates are consistent with G76 disappearance/appearance rates, proving that although R74 is not directly involved in the cleavage reaction, it still serves as an accurate reporter group, and proves the robustness of the method. CJ stands for conjugated in the X-axis labels. Error was calculated from χ^2.

5.5 Different DUBs Display Different Directional Cleavage Preferences

OTUB1 cleaved the distal domain of Ub_3 approximately twice as fast as cleaving the proximal domain (distal 0.016 min^{-1} vs proximal 0.0073 min^{-1}). This

work and others show that OTUB1 is capable of releasing monoUb species during cleavage, and that there is a clear preference for cleaving from the distal end of a chain, in contrast to a previous report[188]. NMR studies and formation of a covalent adduct show that there are indeed two Ub binding sites on OTUB1[233].

IsoT cleaved both bonds at roughly the same rate (distal 0.0133 min^{-1} vs proximal 0.0127 min^{-1}). Wilkinson's fluorescence experiments suggested that IsoT cleaves the proximal unit from free chains, but that any modification to the C-terminus of that proximal unit decreased processing[191]. The presence of D77 on the proximal unit of the mid-^{15}N-Ub_3 construct presents just such a modification. In this case, the presence of D77 (or any other modification) could simulate ubiquitin conjugated to a substrate, rather than a free-floating chain, and thus physiological cleavage is disrupted, giving rise to this apparent lack of cleavage preference. This is consistent both with the initial study of IsoT suggesting that the enzyme was incapable of removing ubiquitin from substrate[190], and a following structural study showing that the diglycine motif at the C-terminus of the most proximal ubiquitin unit binds deep within the zinc-finger binding domain of IsoT[234]. The enzyme has three more ubiquitin binding sites in addition to the highly selective zinc-finger binding domain[235], and longer chains would probably be expected to bind more tightly and be cleaved even more quickly. It is worth noting that IsoT, despite being present at the second lowest concentration in the reaction among all the DUBs tested, showed the highest average rate of cleavage, consistent with its published 2 µM K_M[236] and very high k_{cat} of 40 min^{-1} for polyUb chains[191].

Ubp6 cleaved the proximal bond twice as fast as the distal (distal 0.0014 min^{-1} vs proximal 0.0031 min^{-1}). This contrasts with existing literature, which suggested that Ubp6 would preferentially cleave more distal units from a chain during a process of slowing down substrate transfer to the 19S RP. This could be due to the absence of 19S RP in the reaction, which has been demonstrated to be required for full Ubp6 activity[203]. Ubp6 showed the lowest cleavage rate despite being present at the highest concentration in the reaction relative to the other DUBs tested, also consistent with the documented low activity of recombinant Ubp6/Usp14 without Rpn1. In fact, at the timepoints analyzed, the proximal bond had been completely cleaved, while the distal bond had not yet been fully digested.

Usp2 cleaved the distal bond approximately twice as fast as the proximal (distal 0.0080 min^{-1} vs proximal 0.0037 min^{-1}). The construct utilized here consisted of only the catalytic USP domain, with any N-terminal regulatory elements having been removed. The shortest isoform of Usp2 consists of the catalytic domain with an N-terminal six amino acid extension, so it is a reasonable assumption that the catalytic Usp domain alone would make a good model for study. Addition of the regulatory elements could alter the apparent preference for cleaving the distal domain from chains.

It is worth noting that PLpro concentrations were ten times lower than all other enzymes, as cleavage of the proximal bond proceeds too quickly to measure by NMR at higher concentrations. PLpro cleaved the proximal isopeptide three times faster than the distal (distal 0.0022 min^{-1} vs proximal 0.0063 min^{-1}), as essentially no distal bond cleavage was detected until a majority of the proximal signal had decayed.

In this case, the appearance of the free K48 signal was excluded from the analysis, as normalizing the signal would give the false impression that cleavage of the distal bond was complete. An additional 16-hour incubation allowed for complete cleavage of the distal isopeptide bond. Cleavage of the proximal isopeptide bond at a much higher rate is consistent with PLpro's recognition of two Ub/UBL domains as the primary high affinity substrate[135], and provides further evidence towards its non-canonical physiological role as a deubiquitinase.

5.6 Conclusion

An NMR-based activity assay to quantify DUB activity was utilized to determine directionality of cleavage for a variety of DUB enzymes. Linkage-specific relative cleavage rates are readily determined from extracted peak volumes. The disappearance of isopeptide-linked signals and the appearance of free signals permits for simultaneous tracking of reactant consumption and product formation, allowing for robust two-sided analysis. Previous assays to probe DUB directionality have required the use of extrinsic labels such as fluorophores to differentiate the different Ub units that would otherwise be essentially equivalent. This work does not rely on modifications beyond isotope labeling, resulting in native isopeptide linkages that can be observed in real time. The results (except for Ubp6, which is removed from its vital physiological context) are consistent with previously published reports, further demonstrating the validity and robustness of this method. Distal preference, proximal preference, and no cleavage preference results are easily determined. This assay can be expanded to utilize substrate-bonded ubiquitin units, either in the native isopeptide linkage or utilizing a model nose-to-tail α-amino linkage, to better model removal of

Ub species from a substrate. It can also be utilized to further explore DUB preference for both linkage type and directionality.

The primary downsides of the experiment beyond NMR initial equipment costs are the lengthy amount of time required to extract NMR signal volumes manually. Accurate volume determination is a prerequisite for accurate analysis, and no methods currently available allow for automated volume extraction using integration of a fit curve. Automated box-sum volume extraction exists through a variety of programs but was inadequate for this analysis due to the required compensation for changing relaxation rates. The design of a robust automated volume extraction method would allow for higher throughput and would expand this method further. However, once volume extraction is completed, the remainder of the analysis is trivial and provides easily comparable results.

5.7 Contributions

Benjamin Lanham assisted in trimer synthesis, and prepared all PLpro-related samples and performed all PLpro-related experiments. Steven Bonn performed trimer synthesis, OTUB1, IsoT, Ubp6, and Usp2 preparation, DUB cleavage assays by gel and NMR, and analyzed data. David Fushman assisted in experimental design and supervised the project.

Chapter 6: Small Molecule Inhibitors of the Ubiquitin Proteasome System Studied by NMR

Chapter 6 is adapted from two manuscripts:

Nakasone MA, *et al*. Structural Basis for the Inhibitory Effects of Ubistatins in the Ubiquitin-Proteasome Pathway. *Structure* 2017; **25(12)**: 1839-1855.

Nawatha M *et al*. De novo macrocyclic peptides that specifically modulate Lys48-linked ubiquitin chains. *Nat Chem* 2019; **11:** 644-652.

6.1 Objective

Utilize NMR as a tool to study the interactions of ubiquitin with small molecule systems that inhibit ubiquitin-mediated proteasomal degradation.

6.2 Ubiquitin as a Druggable Target

Ubiquitin is involved in a multitude of different cellular pathways. The most well-characterized is proteasome-mediated protein degradation. Proteasomal degradation of specific substrates is required for advancement through the cell cycle, and dysregulated control of this system is associated with a variety of cancers. In addition, the aberrant function of the ubiquitin tagging and removal systems is associated with cancers and metabolic disease. The common denominator among these systems is ubiquitin itself. Ubiquitin is highly conserved across all eukaryotes, with only three conservative mutations differentiating yeast Ub from human Ub. This high level of conservation across such vast evolutionary space suggests that ubiquitin is not tolerant of mutations, and that mutations which help escape a therapeutic treatment might be likely to destroy ubiquitin's physiological role[237]. Thus, ubiquitin presents an interesting target for therapeutic development.

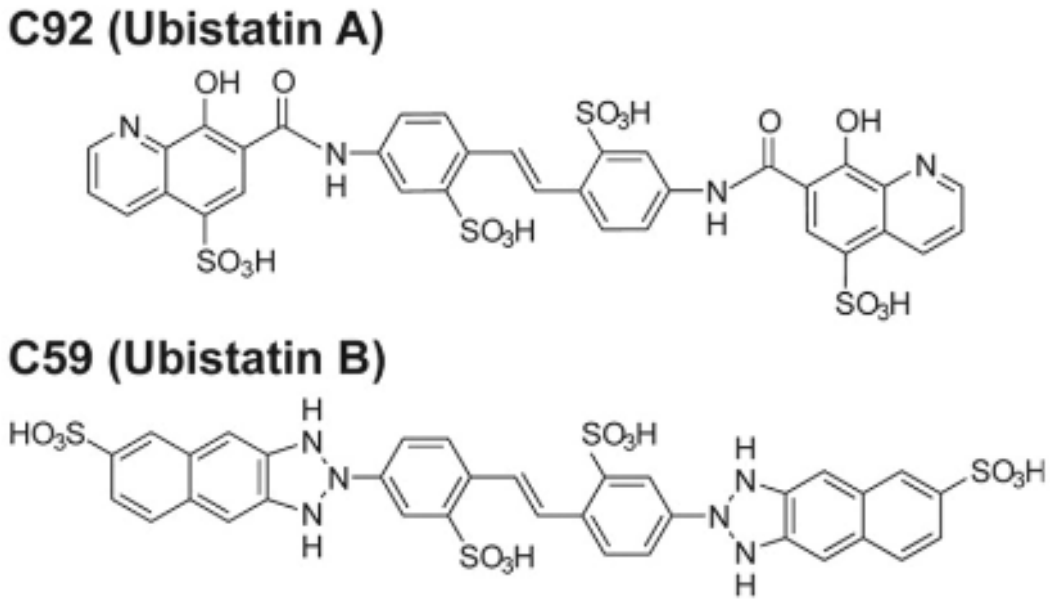

Figure 44- Structure of ubistatin molecules. Reprinted from Verma, R.; Peters, N. R.; D'Onofrio, M.; Tochtrop, G. P.; Sakamoto, K. M.; Varadan, R.; Zhang, M.; Coffino, P.; Fushman, D.; Deshaies, R. J.; et al. Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain. Science **2004**, 306 (5693), 117-120[237]. Used with permission from AAAS.

6.3 Ubistatin

Ubistatins (Figure 44) were identified in a 2004 chemical genetic screen to block mitotic exit by inhibiting cyclin B degradation by the proteasome in *Xenopus* extracts[238]. In addition to arresting the cell cycle, these small molecules prevented ubiquitin-dependent proteasomal degradation of a ubiquitinated substrate Sic1 but did not influence Ub-independent proteasomal degradation. Gel mobility shift assays determined that ubistatins bound specifically to K48-linked polyUb chains, and NMR CSPs verified that they bind at the interface between two K48-linked Ub units[238].

A more recent study in 2017 determined the structure of a hemi-ubistatin upon interaction with ubiquitin[239]. Hemi-ubistatin is effectively half of the full ubistatin molecule, and effectively binds in 1:1 stoichiometry though with an extremely weak

interaction (K_D=590 ± 290 μM) as determined by NMR CSPs[239]. Full ubistatin B binds much more tightly, with K_D on the order of 10-15 μM, as measured by NMR CSPs and fluorescence anisotropy[239]. Fluorescence anisotropy also verified that ubistatin B binds to K48-Ub_2 roughly ten times as tightly as other Ub_2 linkages (K_D=262 ± 23 nM for K48-linked, 2-10 μM for other linkages)[239] Functional studies revealed that ubistatin B binding inhibited proteasomal degradation and DUB activity.

As stated above, many non-conservative mutations in ubiquitin are not tolerated in yeast, and are either fatal or dramatically reduce yeast fitness[237]. In order to fully determine the interactions responsible for the tight ubistatin B binding, mutations to R74 were made. R74 was involved in binding to ubistatin B, as shown by the relatively large CSPs and signal attenuations upon titration of ubistatin B into a K48-Ub_2 solution. R74's positive charge could interact with the sulfonate negative charges, stabilizing ubistatin B's interaction with Ub_2. Making the point mutation R74A would eliminate this charge, disrupt that interaction, and should weaken the overall interaction of ubistatin B with ubiquitin, if not completely abolish it. WT monoUb binds ubistatin B with K_D = 14 ± 1.6 μM, while incorporation of the R74A mutation results in binding that is almost three times weaker, with K_D = 40 ± 15 μM (Figure 45).

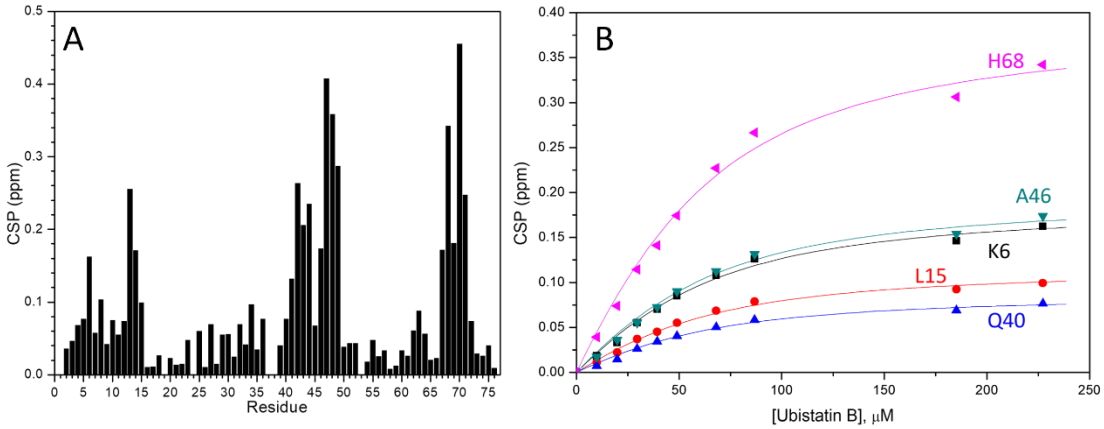

Figure 45- Mutation of R74 to alanine in ubiquitin weakens, but does not eliminate, ubistatin B binding. A) Per-residue CSP plot at saturation with ubistatin B (5 molar excess ubistatin). The largest shifts are located along the hydrophobic patch of ubiquitin, while the tail residues are largely unaffected. B) Binding curves of representative ubiquitin residues involved in ubistatin binding.

6.4 Cyclic Peptides

Cyclic peptides represent a growing area of drug development. This class of molecules typically presents a large surface area for interaction, while reducing conformational mobility and maintaining a relatively small size, leading to higher selectivity for a target[240]. These peptides are typically cell permeable yet non-cytotoxic, since their degradation products are just amino acids[241]. They are relatively easy to synthesize, screen and customize, and incorporation of non-proteinogenic amino acids can allow for autocyclization upon synthesis and enhanced selectivity[242]. Numerous cyclic peptide drugs are available[243, 244], and approximately one new drug per year is approved in the United States.

Recently, de novo macrocyclic peptides have been discovered that are capable of selectively binding to polyUb chains of varying length and linkage[245-247]. The Ub2ii peptide (Figure 46) was found to selectively and tightly bind K48-Ub_2 by SPR[245]. In order to fully understand this interaction, Ub2ii was titrated into solutions of isotopically enriched K48-Ub_2. Interestingly, the resulting spectra show textbook "slow exchange", where the unbound signals disappear, and bound signals appear (Figure 47). This is consistent with the slow off-rates seen by SPR. In order to determine the extent of CSPs and accurately map the binding interface, a backbone signal assignment was completed of fully saturated K48-Ub_2. The assignment of the fully saturated K48-Ub_2 revealed that Ub2ii interacts across the hydrophobic patch of each ubiquitin unit, with large CSPs across those residues (Figure 48)[248]. The hydrophobic patch (L8-I44-V70) is the interface between the two ubiquitin domains,

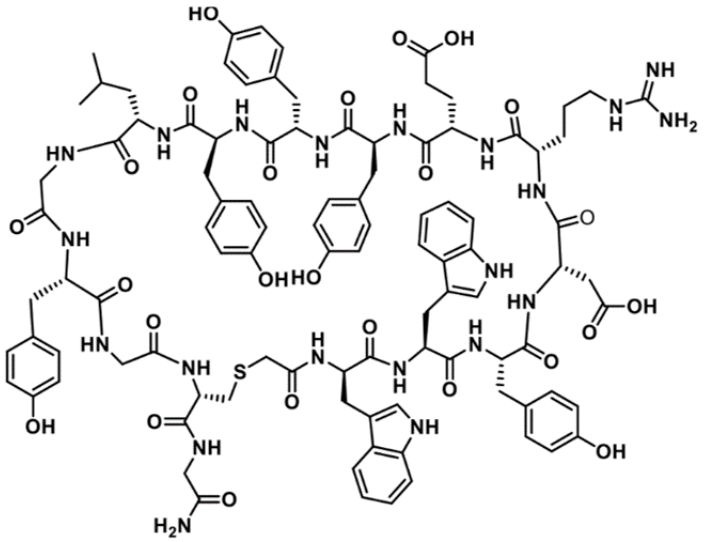

Figure 46-Structure of the Ub2ii peptide that selectively binds to K48-Ub_2. The peptide is very hydrophobic and relatively insoluble in water but is extremely soluble in DMSO.

as well as the site of most ubiquitin-receptor interactions[21, 249], as well as the same location as ubistatin binding[239]. This suggests that Ub2ii is able to intercalate between the two ubiquitin domains. The appearance of a single set of bound signals indicates that Ub2ii binds to a single site and in only one orientation. The particularly large CSPs are most likely due to an induced ring current effect in the many tryptophan moieties of Ub2ii[250]. Importantly, Ub2ii does not display the same interaction with either monoUb or K63-Ub_2, showing only broad attenuations of free signals and no appearance of unbound signals, possibly indicative of induced nonspecific oligomerization.

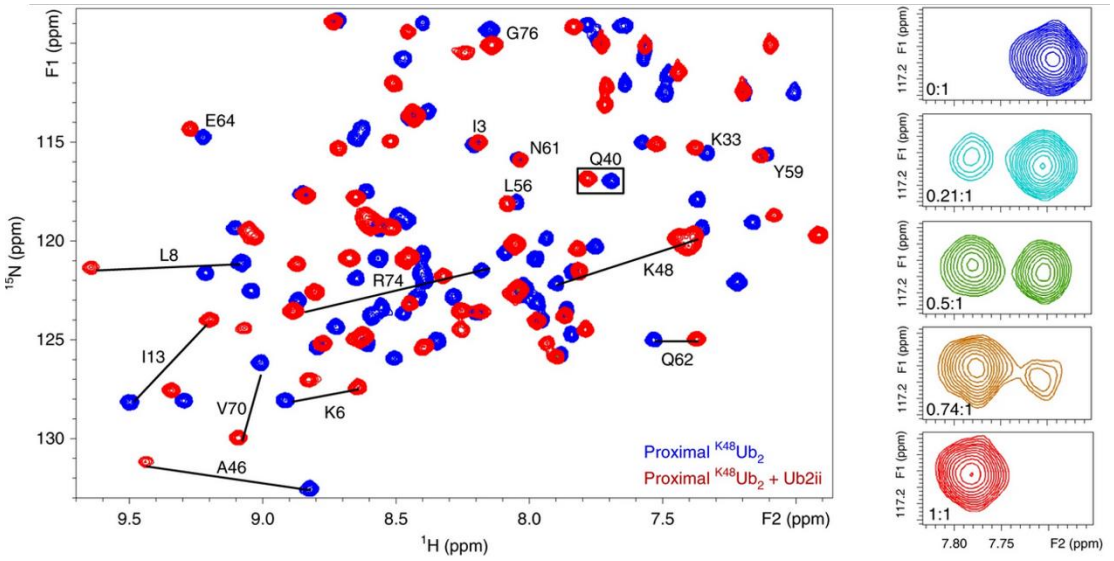

Figure 47- Overlay of ^{1}H-^{15}N SOFAST HMQC spectra of proximal ^{15}N-K48-Ub_2 free (blue) and saturated with 1 molar equivalent of the Ub2ii peptide (red). Selected residues are labeled, and shifts are highlighted. Insert at right highlights the slow exchange pattern of Q40 upon increasing titration of Ub2ii. The initial "free" signal decays as the new "bound" signal appears as the titration progresses. Figure adapted from Nawatha et al[244]. Used with permission.

6.5 Conclusions

Each of the studies above shows complementary *in vitro* and *in vivo* data showing the usefulness of targeting ubiquitin and ubiquitin chains themselves, rather than the associated enzymatic machinery of the ubiquitin proteasome system. The

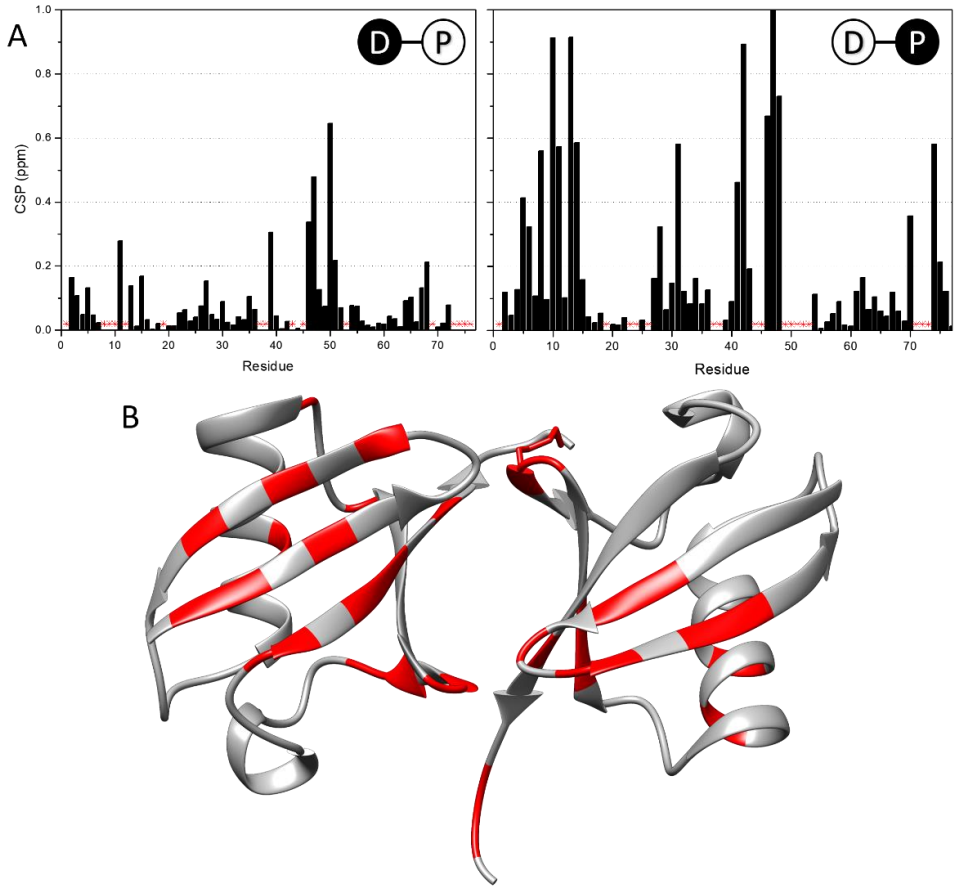

Figure 48-Chemical shift perturbations in K48-Ub_2 upon Ub2ii peptide binding. A) CSP plot showing per-residue shifts at saturation. The distal domain is shown at left, the proximal domain at right. G47 in the proximal domain has a shift of 2.5 ppm. Red asterisks denote that a residue could not be unambiguously assigned either due to attenuations or spectral crowding. B) Affected residues (red) mapped onto the structure of K48-Ub_2 (PDB 1AAR). The distal domain is shown at left, and the proximal domain at right. Residues generally map to the hydrophobic patch and interface between the domains, where the vast majority of ubiquitin interactions occur. Interestingly, the region immediately around the isopeptide bond is strongly affected by Ub2ii binding, suggesting that the peptide can preferentially interact with the isopeptide bond. This could form some of the basis for its selectivity for K48-linked chains. Mapped residues show CSPs above the mean of all assigned CSPs for the given domain (distal mean CSP = 0.09 ppm, proximal mean CSP = 0.25 ppm).

development of small molecules and peptides that can target the ubiquitin proteasome system continues, with further generations of molecules being discovered and synthesized[246, 247]. The high conservation of ubiquitin across eukaryotes and its overall resistance to mutations suggests that interactions with a drug molecule are unlikely to be rescued through random mutations without causing severe harm to the cell[237, 251]. Increasing linkage specificity can present these small molecules as tools for the biochemist to identify more substrates of polyUb chains in addition to their potential application as therapeutics.

Ubistatins and these cyclic peptides present a scaffold upon which further development can proceed, as well as a proof of concept that targeting ubiquitin is indeed a viable therapeutic option. From a drug development point of view, ubistatins show a fairly low affinity for K48-Ub_2, though the molecule can easily be modified to increase the specificity and affinity of the interaction, as well as potentially make it more selective for longer chains. Further development and modification of the cyclic peptides can increase their appeal in the biotechnology space. Indeed, these peptides are already generating some buzz, with nearly $500 million invested in a company dedicated to commercializing these products in the context of drug-resistant multiple myeloma[252].

6.6 Contributions

Steven Bonn performed UbR74E and UbR74A expression and purification, and performed the NMR titrations with ubistatin B and UbR74A and analyzed the data as a rotation project. Christina Camara assisted with all the steps of this process. David Fushman oversaw the entire ubistatin project.

Steven Bonn synthesized isotope enriched monoUb, K48-Ub_2, and K63-Ub_2, performed titrations with Ub2ii, and analyzed the data. David Fushman oversaw the NMR-based aspects of the project.

Chapter 7: Conclusions and Future Perspectives

This work breaks ground exploring the putative UBact proteasomal system. UBact was determined to be a disordered protein, and a backbone resonance assignment was completed. A full suite of NMR experiments as well as circular dichroism did not reveal any evidence of secondary structure, an interesting occurrence for any intrinsically disordered protein. This extreme disorder could assist in its putative physiological role as a degradation tag, where a highly flexible tag could increase the likelihood of interaction with the proteasomal receptor.

UBact does indeed interact with the putative proteasomal receptor, PGP_6, at least the fragment that was utilized in this work. Based on modeling, PGP_6 appears to be homologous to the Rpt proteins from the eukaryotic ubiquitin proteasome system, PAN from the limited examples of Archaeal proteasomes, and Mpa from the Actinobacterial Pup proteasome system. In fact, this work shows that PGP_6 does in fact exist as a hexamer in solution. The observed interaction with UBact further supports this prediction. The N-terminal half of UBact interacts specifically with PGP_6, though it is a fairly low affinity interaction. The UBact-PGP_6 interaction is about 10-times weaker than the interaction of Pup with Mpa, though this could be due to using a truncation of the PGP_6 protein rather than the full sequence. Future studies with full-length protein could shed more light on this interaction. Explorations of the putative UBact proteasome itself could further elucidate how PGP_6 acts as the receptor for UBact, and how it might or might not facilitate interaction with the proteasome.

Despite showing early promise with purifications and characterizations of UBact and PGP_6, ultimately this exploration of the UBact proteasome system came to an end with the inability to purify active ligase and deamidase enzymes. While it is possible that the genes utilized do not encode functional enzymes, the sequence alignment between these UBact operon genes and the *M. tuberculosis* Pup operon genes, with active site residues intact, strongly suggests that these genes to in fact encode enzymes responsible for conjugation and removal of UBact from substrates. Expression strategies utilizing systems other than *E. coli* could provide an alternative towards generating soluble, folded, and active enzymes for further study.

Besides characterizations of the enzymes themselves, functional ligase and deamidase enzymes would have been ideal for identifying possible physiological substrates of the UBact proteasome system, either by proxy in *E. coli* (as has been done with the Pup proteasome system) or directly in *Nitrospira* cells. Identifying substrates of the system, if it indeed does encode a fully functional proteasomal operon, would be a first step toward understanding the evolutionary diversity of this new proteasome system. *Nitrospira* bacteria are not well studied beyond their role in nitrogen cycle, yet sequencing efforts are beginning to discover *Nitrospira* bacteria in habitats around the world. Their role as complete ammonia oxidizers has immediate implications in the treatment of wastewater and other biotechnological applications for those willing to pursue the isolation, growth, and potential modification of these bacteria. Understanding the UBact proteasome system could assist in providing some measure of control over these organisms and could help in understanding the role they play in the environment.

NMR has been used to study kinetics, both in the context of organic reactions and enzyme-mediated reactions. This work further demonstrates the usefulness of NMR for measuring reactions kinetics, this time quantifying deubiquitinase activities. The experimental setup demonstrated herein can be widely applied to any DUB targeting any polyUb linkage, so long as individual ubiquitin units can be specifically isotope labeled. The primary strength of this method is that the native isopeptide linkages are directly observed, with no chemical modifications beyond isotope labeling. Almost all DUB activity assays to date have relied on extrinsic modifications to track cleavage outcomes, or utilized non-native linkages to control chain synthesis or limit which bonds are cleavable. This straightforward isotope-labeled chain synthesis and reaction setup could be made more widely applicable with more efficient signal volume integration. This work generally agrees with published literature, lending support to the robustness of the method. At the same time, there was some discrepancy in the cleavage tendencies of OTUB1 that has been definitively cleared up. OTUB1 is able to cleave individual ubiquitin units from chains, even though it does utilize a diubiquitin unit for efficient binding. This assay can become a widely utilized tool for studying DUBs, especially given many DUB's lack of linkage-specificity and the many physiological implications of dysregulation of deubiquitination.

A more traditional role for NMR is its use in describing interactions of molecules at atomic resolution. The interactions of two small molecules with ubiquitin were explored here. Ubistatins showed promise as highly specific modulators of proteasomal degradation by directly interfering with ubiquitin

recognition at the proteasome, though apparently no further development has come about. Moreover, the Ub2ii peptide demonstrated a similar ability to dramatically slow down the proteasome by binding directly to ubiquitin. While it might be tempting to target ubiquitin itself in evolving therapies, ubiquitin's wide-reaching physiological impact could reduce its applicability as a druggable target. Ubiquitin's ubiquitous appearance across so many varied metabolic pathways raises questions about the off-target effects of directly targeting ubiquitin as a treatment modality.

Perhaps more interesting than directly attempting to develop a drug treatment, these molecules represent scaffolds upon which new biotechnological tools could be built. The search for ubiquitin substrates carries on in labs around the world, and these small molecules could be further developed to have a higher selectivity for specific polyubiquitin linkage types or specific numbers of polyUb units in a chain, enabling pulldown assays and further identification of ubiquitin substrates. Their small size and cell permeability mean that they could be utilized *in vivo* to identify native substrates.

In the end, proteasomes are clearly a vital area of study across all domains of life. Understanding how they function and how species have evolved to utilize them to enhance their fitness is important toward understanding our own evolutionary history and pathologies associated with the proteasomal systems. It is entirely possible that the proteasome, once thought to be unique to eukaryotes, turns out to be relatively common across nature as bioinformatic and microbiological technologies advance.

Chapter 8: Materials and Methods

8.1 General Considerations

All reagents were purchased from Fisher, Millipore Sigma, and New England Biolabs unless otherwise specified. Isotope enriched reagents were purchased from Millipore Sigma or Cambridge Isotope Laboratories.

^{15}N and ^{13}C isotope enrichment was achieved by growing expression cells in M9 minimal media supplemented with 1 g/L $^{15}NH_4Cl$ and 2 g/L ^{13}C-glucose as the sole source of nitrogen and carbon, respectively, in the final large culture. Proteins were expressed and purified as described in their relevant sections.

8.2 Chapter 2 Methods

8.2a UBact Expression and Purification

Expression and purification of *N. nitrosa* UBact from *E. coli* BL21 (DE3) cells is described in detail in Chapter 2. Briefly, GST-UBact is optimally expressed at 18°C with 1 mM IPTG induction. Cells are harvested and lysed, and lysate is clarified by ultracentrifugation. GST-UBact is purified by glutathione affinity chromatography, and UBact is cleaved from the GST construct by enterokinase and purified by cation exchange chromatography at pH 8.0.

8.2b NMR Experiments

NMR samples of 500 µM UBact were prepared in 20 mM sodium phosphate pH 6.8, 0.02% (w/v) sodium azide, 5% D_2O. Spectra were acquired using a Bruker Avance III Ultrashield 600 MHz spectrometer equipped with a TCI cryoprobe and a

Bruker Avance III HD Ascend 800 MHz spectrometer equipped with a CPQCI cryoprobe using Bruker standard pulse seqeunces[253-263] at 23°C. Data was processed using Bruker TopSpin 4.0.8 and analyzed with CCPNMR Analysis 2.4.2[264] and NMRFAM-SPARKY[232]. ^{1}H-^{15}N SOFAST HMQC and HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, and ^{15}N-separated TOCSY experiments were utilized for backbone signal assignments. A ^{13}C-direct detected CON experiment was used for the purposes of proline residue assignment. Steady-state ^{1}H-^{15}N heteronuclear Overhauser enhancement values[265] were obtained as ratios of signal intensities in experiments with and without proton saturation, recorded in a pseudo-3D interleaved fashion using a water flip-back measurement scheme[266] with inter-scan delay of 4.5 s. Errors were calculated by integrating spectral noise. ^{15}N R_1 relaxation experiment was acquired in a pseudo-3D interleaved manner with relaxation delays of 880 and 4 ms. ^{15}N R_2 relaxation experiment was acquired in a pseudo-3D interleaved manner with relaxation delays of 200, 120, and 8 ms. Errors were calculated by quantifying duplicate peaks and Monte Carlo simulation for 500 points. The CEST experiment was acquired in a pseudo-3D manner, with a CEST mixing time of 400 ms and 40 saturated frequencies at 0.5 ppm intervals in the ^{15}N dimension. CPMG relaxation dispersion experiment was acquired in a pseudo-3D interleaved fashion with CPMG frequencies of 40, 100, 203, 307, 413, 630, and 1086 Hz. This data was analyzed with NESSY[267].

8.2c Circular Dichroism

CD experiments were performed on a JASCO J810 spectro-polarimeter with a 25 µM UBact sample in a 1 mm pathlength quartz cuvette in 20 mM sodium

phosphate pH 6.8 at 20°C. Data was smoothed with a 9 point Savitzky-Golay filter and analyzed using the Capito webserver[268].

8.2d Chemical Shift Indexing

Random coil chemical shifts were acquired by entering the amino acid sequence of UBact into the webserver at https://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS/. These values, corrected for temperature, pH, and nearest neighbors, were used to determine secondary chemical shifts for resonances that could be unambiguously assigned.

Assigned resonances were used as the input for TALOS+[143].

8.3 Chapter 3 Methods

8.3a Proteasomal Gate Protein Expression and Purification

The detailed expression and purification of PGP_6 is described in the chapter. Briefly, His-PGP_6 is expressed in *E. coli* BL21 (DE3) at 18°C with 500 µM IPTG induction. Cells are harvested and lysed by sonication, then lysate is clarified by ultracentrifugation. His-PGP_6 is purified by nickel affinity chromatography.

8.3b Circular Dichroism

CD experiments were performed as above with a 2.5 µM sample of PGP_6 in a 1 mm pathlength quartz cuvette in 20 mM sodium phosphate pH 6.8, 130 mM NaF at 20°C. Data were smoothed with a 9 point Savitzky-Golay filter and analyzed using the Dichroweb webserver[150-152, 269].

8.3c NMR Experiments and Relaxation Titration

Spectra were acquired using a Bruker Avance III Ultrashield 600 MHz spectrometer equipped with a TCI cryoprobe and a Bruker Avance III HD Ascend 800 MHz spectrometer equipped with a CPQCI cryoprobe using Bruker standard pulse sequences at 23°C. 83 µM ^{15}N-PGP$_6$ was prepared in a 5 mm NMR tube in 20 mM sodium phosphate pH 6.8, 130 mM NaCl for the ^{1}H-^{15}N TROSY experiment.

For the relaxation titration, several samples of 300 µM ^{15}N UBact M17W were prepared with concentrations of PGP$_6$ of 0, 12, 25, 38, 50, 100, 150, and 300 µM PGP$_6$. ^{15}N T_2 relaxation experiments were acquired in a pseudo-3D interleaved manner with two delays per titration point. Delays were determined by comparing intensities of representative 1D planes of the experiment to achieve enough difference in intensity for accurate rate determination while still maintaining enough signal for accurate signal determination. Delays used are listed below:

[PGP$_6$]	0 µM	12 µM	25 µM	38 µM	50 µM	100 µM	150 µM	300 µM
Delays (ms)	200, 8	160, 8	152, 8	152, 8	160, 8	144, 8	96, 8	64, 8

The intensities of the resulting spectra were picked and were used to determine the relaxation times. ^{15}N T_2 relaxation times were determined with:

$$T_2 = \frac{\Delta t_{delay}}{\ln\left(\frac{I_{short}}{I_{long}}\right)} \tag{13}$$

where T_2 is the relaxation time, I_{short} is the intensity of the signal from the short delay experiment, I_{long} is the intensity from the long delay experiment, and Δt_{delay} is the difference in delay times. These T_2's were used to complete equation 11 in Chapter 3 for input (T2f/T2obs-1) into the in-house Matlab program kd_fit[157]. This program fits the response (in this case the changing relaxation times) to the quadratic equation:

$$P_B = \frac{\left([P_t] + [L_t] + K_D - \sqrt{([P_t] + [L_t] + K_D)^2 - 4[P_t][L_t]}\right)}{(4[P_t])} \qquad (14)$$

Here, $[P_t]$ is the total concentration of UBact, and $[L_t]$ is the total concentration of PGP_6. The error in K_D was determined by fitting individual residues to the binding curve and taking the standard deviation of the individual fits.

8.3d Surface Plasmon Resonance

Handling of the Biacore T200 instrument (GE Healthcare) was done by Guanghui Zong and Connor Donahue. A CM5 chip was activated by washing with 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide, 0.05 M N-hydroxysuccinimide. UBact M17W in 20 mM sodium phosphate pH 6.8, 130 mM NaCl was washed over the activated chip at 20-50 µg/mL concentrations, but no response was ever detected to suggest any protein was immobilized.

8.3e Fluorescence Quenching

Fluorescence experiments were performed on a Horiba Fluoromax 4 fluorimeter, with excitation at 295 nm and bandwidth of 5 nm. Samples of UBact were prepared in a Starna Type 3-3 quartz cuvette. Samples contained 50 µM UBact M17W, and concentrations ranging from 0-50 µM PGP_6. Emission was measured from 310-420 nm with a bandwidth of 5 nm. Spectra from samples containing only PGP_6 were subtracted from titration sample spectra, and the inner filter effect was corrected by measuring absorbance of the titration samples across all wavelengths observed.

8.3f Isothermal Titration Calorimetry

ITC experiments were performed on a MicroCal VP-ITC (Malvern Panalytical) and MicroCal ITC200 (Malvern Panalytical). Proteins were extensively buffer exchanged into 20 mM sodium phosphate pH 6.8, 130 mM NaCl and subsequently dialyzed against the same for at least four hours. Concentrations ranging from 250-500 μM PGP_6 were used in the cell, and 3-5.5 mM UBact M17W was used in the syringe. Experiments were conducted at 20°C. Stirring speed was 750 rpm. 2.1 μL injections of UBact M17W were added to the sample solution of PGP_6 over 4 seconds, with 180 second spacing between injections. A one-site binding model was used with ΔH (enthalpy change in kcal/mol), K_A (association constant), and n (stoichiometry per PGP_6) as variable parameters. Since tests never reached saturation at the conditions tested, no reasonable estimates could be made for these parameters.

8.3g Analytical Ultracentrifugation

AUC experiments were performed on a Beckman Coulter XL-1 with a 4 hole An-60 Ti rotor at 20°C using absorbance detection with a 1.2 cm pathlength. The sedimentation velocity experiment was performed at 42,000 rpm with 100 μM PGP in 20 mM sodium phosphate pH 6.8, 130 mM NaCl in a two-hole cell. Scans were taken every 70 seconds. The data was analyzed using SEDFIT[270, 271] and visualized using Gussi[272] and Origin 8. The sedimentation equilibrium experiment to confirm the molecular weight was run by spinning at a given rpm for 8 hours, acquiring a scan, then spinning for one more hour and acquiring a second scan to test for equilibrium. The speeds utilized were 13,000, 15,000, and 18,000 rpm. 25 μM, 35 μM, and 50 μM PGP were used in 20 mM sodium phosphate pH 6.8, 130 mM NaCl. This experiment

utilized the six-hole cell with a pathlength of 1.2 cm. Data was analyzed using SEDFIT[270, 271] and SEDPHAT[273], and visualized using Gussi[272].

The sedimentation equilibrium experiment to determine the binding of UBact and PGP_6 was performed in six-hole cells with 5 μM UBact-F and PGP_6 concentrations of 0, 5, 10, 15, 25, and 50 μM. Speeds tested were 10,000, 13,000, 15,000, 18,000, 21,000, 24,000, and 27,000 rpm. Samples were spun for 8 hours at each speed, then an additional hour to ensure equilibrium had been reached. Each scan was selected based on the curvature of the signal. In general, the higher speed scans showed most of the sample piled up at the outer wall of the cell, rendering the data useless. No data from the 50 μM PGP_6 was usable for this reason. Data was analyzed with SEDFIT[270, 271] and SEDPHAT[273]. As with the NMR data, the data was best fit when using a binding model of 3 UBact per 1 PGP_6. Error was determined by Monte Carlo simulation over 1000 points. Data was visualized with Gussi[272].

8.4 Chapter 4 Methods

Methods are described in the chapter.

8.5 Chapter 5 Methods

8.5a Polyubiquitin chain synthesis

Ubiquitin monomers UbD77, UbK48R, and ^{15}N-UbD77 were expressed in BL21 (DE3) pJY2 cells from the pET3a plasmid in autoinducing ZYM media[274] (for non-isotope enriched Ub units) or M9 minimal media as described above (for ^{15}N-enriched UbD77). Cells were harvested by centrifugation at 4680 x *g* for 20 minutes, then cells were resuspended in 50 mM tris pH 7.6 supplemented with 0.4 mg/mL

lysozyme, 5 mM $MgCl_2$, DNase, 1% Triton X-100, and 1 mM PMSF. Cells were lysed by sonication, then the lysate was ultracentrifuged at 48,900 x *g* for 30 minutes to pellet cell debris. Clarified lysate supernatant was subjected to acid precipitation with 0.7% perchloric acid, then the solution was ultracentrifuged at 48,900 x *g* for 30 minutes to pellet precipitate. The supernatant containing ubiquitin was dialyzed overnight against 50 mM ammonium acetate pH 4.5 at 4°C. Ubiquitin monomers were purified by cation exchange chromatography at pH 4.5. Individual monomers typically elute around 80-100 mM NaCl and are highly pure upon cation purification. Ubiquitin units are buffer exchanged into a neutral storage buffer for long term storage, or a reaction buffer for chain synthesis.

Mid-^{15}N-Ub_3 was enzymatically synthesized as previously described[217, 275] in the chapter. Human E1 was expressed in BL21 (DE3) Rosetta cells and purified by nickel affinity chromatography. GST-E2-25K was expressed in BL21 (DE3) Rosetta cells and purified by glutathione affinity chromatography. Yeast Yuh1 was expressed in BL21 (DE3) Rosetta 2 pLysS cells and purified by ammonium sulfate purification and anion exchange chromatography. Ubiquitin monomers UbK48R and ^{15}N-UbD77 (up to 40 mg each) were reacted to form a K48-linked Ub_2 in 50 mM tris pH 8.0, 5 mM $MgCl_2$, 10 mM creatine phosphate, 0.6 U/mL inorganic phosphatase, 0.6 U/mL creatine phosphokinase, with 200 nM E1 and 20 µM GST-E2-25k at 37°C overnight. The reaction was quenched with the addition of 10 mM dithiothreitol (DTT), and a catalytic amount of Yuh1 was added to the mixture to remove D77 from the dimer. Following incubation for several hours at 37°C, several drops of glacial acetic acid were added to the reaction mixture to precipitate the enzymes and lower the pH of the

solution. The reaction mixture was then subjected to cation exchange chromatography at pH 4.5 as above. K48-Ub_2 eluted around 150 mM NaCl. Diubiquitin was buffer exchanged back into 50 mM tris pH 8.0. Purity was assessed by SDS-PAGE and mass spec analysis. The K48-Ub_2 was subjected to a second round of enzymatic chain synthesis, with unlabeled UbD77 forming the proximal unit following conjugation. Following quenching of the reaction with DTT, mid-^{15}N-Ub_3 was purified by cation exchange as above and was buffer exchanged into DUB reaction buffer (50 mM tris pH 7.4, 100 mM NaCl, 1 mM TCEP).

8.5b Deubiquitinase Expression Purification, and Testing

IsoT was expressed in BL21 (DE3) cells from the pGEX-6p-1 plasmid, encoding an N-terminal GST fusion tag. OTUB1 was expressed in BL21 (DE3) Rosetta cells from the pRoEx plasmid encoding an N-terminal His_6 tag. PLpro was expressed from BL21 (DE3) Gold cells from a pMCSG53 plasmid, encoding an N-terminal His_6 tag. Ubp6 was expressed in M15 cells from a pQE30 plasmid encoding an N-terminal His_6 tag. The Usp2 catalytic domain (residues 267-599) was expressed in BL21 (DE3) from a pET vector encoding an N-terminal His_6 tag. All cells were grown in 1 L LB cultures at 37°C to OD_{600} = 0.8, where they were induced with 1 mM IPTG at 16° (IsoT, OTUB1, Ubp6, and Usp2) or 200 µM IPTG with 20 mM K_2HPO_4 at 14°C (PLpro). Cells were harvested by centrifugation at 4680 x *g* for 20 minutes, resuspended in cold PBS pH 7.4 supplemented with 0.4 mg/mL lysozyme, DNase, 10 mM $MgCl_2$, and 3 mM PMSF, then cells were lysed by sonication on ice. Lysates were clarified by ultracentrifugation at 48,900 x *g* for 30 minutes to pellet cell debris. Clarified lysate supernatant was filtered through a 0.22 µm membrane.

GST-IsoT was purified by immobilized glutathione affinity in a gravity column at 4°C. Following application of the lysate to the column, unbound material was washed out with 20 column volumes of PBS. Bound protein was eluted with 50 mM tris pH 8.0, 10 mM glutathione. GST-IsoT was concentrated, and buffer exchanged in DUB reaction buffer and stored at -80°C.

His_6-OTUB1, His_6-Ubp6, and His_6-Usp2 were all purified by nickel affinity chromatography at 4°C. After application of the lysate to the column, unbound material was washed out with PBS + 20 mM imidazole until absorbance dropped nearly to 0. Bound enzymes were then eluted with a gradient from 25-500 mM imidazole. All proteins eluted around 100 mM imidazole, as expected for proteins with solvent exposed His_6 tags. Purified proteins were concentrated, and buffer exchanged into DUB reaction buffer and stored at -80°C.

His_6-PLpro was purified by nickel affinity as above. Rather than PBS, PLpro required 50 mM HEPES pH 8.0, 500 mM NaCl, 5% v/v glycerol, 1 mM TCEP, 1 µM $ZnCl_2$.

Prior to NMR studies with the triubiquitin construct, the concentration of enzyme in the reaction had to be optimized. There had to be enough enzyme for cleavage to proceed, but a low enough amount that all polyUb chains would not be consumed rapidly, in order to be accurately quantified by NMR. Test reactions with varying concentrations of enzymes were set up with 100 µM K48-Ub_2 in 85 µL and incubated at room temperature. Reaction time points were quenched with SDS gel loading buffer and visualized by SDS-PAGE.

8.5b NMR Experiments and Data Analysis

^{1}H-^{15}N SOFAST HMQC experiments were utilized for their short recycling delay. A reduced spectral width (19 ppm in ^{15}N centered on 117.5 ppm, 13 ppm in ^{1}H centered on 4.7 ppm) was used to enhance resolution, and no folded-in signals overlapped with the spectral region of interest. Samples of 100 µM mid-^{15}N-Ub_3 were prepared in DUB reaction buffer to a final volume of 150 µL with 10% D_2O in a 5/3 mm tube. An initial spectrum was obtained without DUB enzymes added, and then catalytic amounts of DUB (138 nM IsoT, 192 nM OTUB1, 18 nM PLpro, 667 nM Ubp6, 333 nM Usp2) were added and time course experiments were run. Spectra were processed as Gaussian functions in TopSpin with Gaussian broadening of 0.2 and Lorentzian broadening of -25 Hz.

Peaks were picked using an in-house Matlab program Autopick that also extracts intensities. Peak volumes were integrated using NMRFAM-SPARKY[232], iteratively selecting regions of each peak to serve as the basis for integration until the correct volume could be reproducibly extracted. Peak volumes were normalized and manipulated as described in the chapter and were fit to an exponential decay (Equation 12) to quantify the relative rate of cleavage. Errors were determined directly from the χ^2 measurements. Figures were prepared in Origin 8.

8.6 Chapter 6 Methods

Ubiquitin monomer purification was performed as described above. Ubistatin B was dissolved in 20 mM sodium phosphate pH 7.4. Cyclic peptide Ub2ii was dissolved to a concentration of 20 mM in D_6-DMSO.

The ubistatin titration was carried out using a 50 µM ^{15}N-UbR74A sample in a Shigemi tube (volume 250 µL) in 20 mM sodium phosphate pH 6.8, 7% D_2O. Ubistatin B was added to this solution to give concentrations of 10, 20, 30, 40, 50, 70, 86, 185, and 227 µM ubistatin. ^{1}H-^{15}N SOFAST HMQC experiments were run at each titration point at 23°C. Peaks were picked using the in-house Matlab program Autopick, and CSPs at each point were determined by:

$$CSP = \sqrt{(\Delta\delta_H)^2 + (\frac{\Delta\delta_N}{5})^2} \quad (15)$$

where $\Delta\delta_H$ and $\Delta\delta_N$ are the chemical shift differences for the proton and nitrogen resonances, respectively. The dissociation constant was determined as described above, assuming a 1:1 stoichiometry.

The Ub2ii titration was carried out using a 100 µM sample of either K48-Ub_2 with the proximal or distal Ub unit ^{15}N-labeled in 20 mM sodium phosphate pH 6.8, 7.5% D_2O in a Shigemi tube. Ub2ii was added to 20, 50, 74, and 100 µM concentrations for the experiments with proximally labeled K48-Ub_2, with two additional points at 114 and 129 µM for the distally labeled K48-Ub_2. ^{1}H-^{15}N SOFAST HMQC experiments were run at each point, at 23°C for the proximally labeled K48-Ub_2, and 37°C for the distally labeled K48-Ub_2.

The backbone resonance assignment was completed using Bruker standard and in-house pulse sequences as described above. 200 µM K48-Ub_2 with the proximal Ub unit ^{13}C/^{15}N-labeled with 200 µM Ub2ii, and 330 µM K48-Ub_2 with the distal Ub unit ^{13}C/^{15}N-labeled with 429 µM Ub2ii samples were prepared in 20 mM sodium phosphate pH 6.8, 7.5% D_2O in a Shigemi tube. As above, the proximally labeled

K48-Ub_2 were carried out at 23°C, while the distally labeled K48-Ub_2 experiments were carried out at 37°C.

All NMR data was processed in Bruker TopSpin and analyzed with NMRFAM-SPARKY and CCPNMR. CSPs were calculated using Eq. 15.

Residue	R1 600 MHz	R1 600 Err	R1 800 MHz	R1 800 Err	R2 600 MHz	R2 600 Err	R2 800 mHz	R2 600 Err	hetNOE 600 MHz	hetNOE Err
3	1.06	9.09 E-03	1.07	8.06 E-02	1.48	5.24 E-03	1.27	2.58 E-02	-1.220	0.011
4	1.07	2.70 E-02	1.22	2.01 E-01	1.86	1.51 E-02	2.08	4.49 E-02	-0.806	0.019
5	1.25	5.37 E-03	1.42	6.95 E-02	1.94	3.07 E-03	1.96	1.43 E-02	-0.614	0.007
6	1.36	3.82 E-03	1.38	5.59 E-02	1.96	1.79 E-03	1.83	2.33 E-02	-0.376	0.008
8	1.5	2.40 E-03	1.52	4.06 E-02	2.19	1.39 E-03	2.6	8.26 E-03	-0.220	0.008
9	1.53	5.42 E-03	1.54	6.68 E-02	2.53	3.32 E-03	2.57	1.54 E-02	-0.125	0.007
10	1.51	1.30 E-02	1.68	9.70 E-02	2.7	6.54 E-03	2.63	2.02 E-02	-0.141	0.008
11	1.54	7.69 E-03	1.62	7.27 E-02	2.55	3.25 E-03	2.95	1.45 E-02	-0.074	0.011
12	1.53	9.46 E-03	1.51	3.94 E-02	2.77	2.06 E-03	2.68	9.58 E-03	-0.107	0.004
14	1.44	8.65 E-03	1.69	1.23 E-01	2.59	4.96 E-03	3.17	2.21 E-02	-0.132	0.013
15	1.46	3.26 E-03	1.65	3.95 E-02	2.48	1.80 E-03	2.79	7.51 E-03	-0.039	0.009
17	1.44	4.13 E-03	1.57	6.01 E-02	2.41	2.47 E-03	3.23	1.24 E-02	-0.205	0.010
19	1.47	1.87 E-03	1.5	9.34 E-02	2.45	3.57 E-03	3.36	1.97 E-02	-0.191	0.005
20	1.37	1.16 E-02	1.49	7.97 E-02	2.74	6.59 E-03	3.52	2.06 E-02	-0.146	0.014
22	1.35	6.06 E-03	1.36	4.61 E-02	2.47	3.14 E-03	3.14	1.06 E-02	-0.219	0.007
24	1.51	2.44 E-03	1.43	2.35 E-02	2.47	2.61 E-03	3.08	1.02 E-02	-0.173	0.005
25	1.51	2.80 E-03	1.55	2.43 E-02	2.27	2.52 E-03	2.7	6.24 E-03	-0.167	0.003
26	1.43	2.65 E-03	1.43	3.07 E-02	2.2	1.54 E-03	2.63	1.30 E-02	-0.225	0.007
27	1.34	6.84 E-03	1.38	5.47 E-02	2.04	4.07 E-03	2.63	2.61 E-02	-0.351	0.008
28	1.29	6.46 E-03	1.43	8.65 E-02	2	3.65 E-03	2.66	1.18 E-02	-0.306	0.015
29	1.23	5.18 E-03	1.27	2.86 E-02	1.75	3.48 E-03	1.92	6.61 E-03	-0.391	0.010
31	1.44	3.78 E-03	1.45	4.59 E-02	2.41	2.52 E-03	2.89	9.81 E-03	-0.205	0.010
32	1.52	7.79 E-03	1.5	4.32 E-02	2.59	3.56 E-03	2.73	1.15 E-02	-0.107	0.006
34	1.51	3.69 E-03	1.47	3.52 E-02	2.56	3.34 E-03	3.17	8.44 E-03	-0.112	0.007

35	1.45	1.06 E-02	1.5	5.22 E-02	2.66	6.49 E-03	3.29	1.38 E-02	-0.142	0.016
36	1.38	2.03 E-02	1.31	8.76 E-02	2.35	1.22 E-02	2.99	1.94 E-02	-0.151	0.016
37	1.38	2.30 E-02	1.31	8.41 E-02	2.72	1.39 E-02	3.05	2.16 E-02	-0.080	0.021
39	1.62	6.01 E-03	1.54	4.58 E-02	3.1	3.30 E-03	3.01	1.19 E-02	0.033	0.005
40	1.69	9.77 E-03	1.69	8.46 E-02	3.36	4.87 E-03	3.21	1.78 E-02	0.155	0.007
41	1.79	6.10 E-03	1.69	7.43 E-02	3.66	2.65 E-03	3.54	1.04 E-02	0.210	0.007
42	1.76	7.63 E-03	1.6	4.74 E-02	4.06	4.21 E-03	4.09	9.44 E-03	0.258	0.011
43	1.86	8.41 E-03	1.89	5.53 E-02	4.15	4.44 E-03	3.66	9.92 E-03	0.283	0.007
44	1.89	3.58 E-03	1.68	3.29 E-02	3.89	6.05 E-03	3.94	8.22 E-03	0.278	0.003
46	1.87	5.86 E-03	1.72	5.01 E-02	4.29	3.43 E-03	3.97	1.09 E-02	0.278	0.005
47	1.81	4.94 E-03	1.64	3.07 E-02	3.7	2.40 E-03	3.82	1.04 E-02	0.235	0.004
48	1.8	1.59 E-02	1.77	8.89 E-02	4.26	1.04 E-02	4.27	2.04 E-02	0.235	0.010
49	1.76	2.16 E-02	1.52	7.92 E-02	4.01	1.28 E-02	3.93	1.65 E-02	0.209	0.013
50	1.68	6.70 E-03	1.59	4.13 E-02	3.73	3.37 E-03	3.74	9.49 E-03	0.109	0.005
51	1.64	6.96 E-03	1.53	8.13 E-02	3.03	1.39 E-02	3.73	1.92 E-02	0.075	0.012
53	1.79	6.28 E-03	1.63	5.51 E-02	3.67	2.43 E-03	4.23	1.43 E-02	0.199	0.004
54	1.81	4.75 E-03	1.61	3.44 E-02	3.7	2.42 E-03	3.83	9.59 E-03	0.235	0.004
55	1.72	6.07 E-03	1.68	5.45 E-02	3.2	6.64 E-03	3.07	1.06 E-02	0.147	0.008
56	1.76	5.36 E-03	1.78	5.95 E-02	3.27	2.63 E-03	3.74	1.18 E-02	0.157	0.004
57	1.78	1.03 E-02	1.77	6.86 E-02	3.21	4.55 E-03	3.38	1.43 E-02	0.151	0.006
58	1.81	5.58 E-03	1.7	4.84 E-02	3.5	3.16 E-03	3.52	1.22 E-02	0.194	0.004
59	1.74	7.70 E-03	1.72	6.69 E-02	3.35	4.24 E-03	3.23	1.26 E-02	0.071	0.005
60	1.51	2.78 E-03	1.55	2.35 E-02	2.37	3.08 E-03	2.7	6.06 E-03	-0.167	0.003
62	1.3	1.22 E-02	1.55	8.03 E-02	2.23	8.40 E-03	2.53	1.76 E-02	-0.527	0.022
63	1.12	8.65 E-03	1.16	6.71 E-02	1.66	5.14 E-03	2.02	1.76 E-02	-0.742	0.012
64	0.952585	6.98 E-04	0.927787	9.38 E-03	0.790891	5.78 E-03	0.752357	3.01 E-03	-1.140	0.018

Table 1-Relaxation parameters for UBact acquired at 600 and 800 MHz. R_1 and R_2 are given in s^{-1}. Errors in relaxation rates were determined from duplicate

experiments at each relaxation delay. NOE values have no units. NOE errors were determined by integrated spectral noise.

Residue	0 uM PGP_6	12 uM PGP_6	25 uM PGP_6	38 uM PGP_6	50 uM PGP_6	100 uM PGP_6	150 uM PGP_6	300 uM PGP_6
3	686	439	248	215	169	116	77	45
4	478	314	219	166	146	111	72	19
5	606	308	210	173	145	102	67	31
6	432	249	173	140	117	85	71	60
8	409	246	167	138	116	90	64	46
9	370	265	235	196	178	152	122	84
10	381	205	154	136	116	90	64	58
11	341	208	154	123	108	79	68	--
12	379	192	142	111	105	71	57	59
14	377	201	133	116	101	86	48	20
15	388	149	113	99	62	78	57	29
17	321	173	116	100	84	--	55	--
19	322	212	134	135	111	83	72	174
20	--	--	--	--	--	--	--	--
22	428	204	138	120	101	76	85	131
24	348	142	108	103	117	125	108	60
25	371	215	178	180	179	156	121	73
26	386	203	146	117	123	37	83	5
27	443	229	156	139	109	127	90	153
28	436	242	165	127	117	100	82	173
29	520	233	172	145	149	165	202	190
31	401	146	93	115	132	107	135	87
32	372	129	75	52	37	19	--	38
34	354	180	122	100	89	69	54	33
35	361	243	182	160	274	146	139	141
36	422	278	214	165	171	92	82	95
37	397	294	241	193	176	138	100	61
39	309	246	195	171	156	114	85	57
40	277	215	195	149	142	103	80	46
41	281	16	18	161	143	103	34	56
42	248	204	165	143	130	96	76	50
43	251	201	154	153	140	103	74	50
44	266	203	153	131	119	88	70	46
46	269	212	154	133	123	88	70	42
47	285	221	174	156	143	103	58	53
48	256	203	154	134	119	89	63	35
49	307	198	161	181	150	112	--	22
50	273	204	185	164	151	114	89	50
51	342	265	214	193	165	130	96	52

53	299	265	171	173	166	124	96	48
54	285	221	174	156	143	103	58	53
55	293	250	217	194	184	140	103	57
56	357	290	221	204	--	140	108	57
57	323	253	203	190	182	133	101	53
58	--	--	202	166	--	62	--	51
59	312	264	224	200	184	139	107	59
60	371	215	178	180	179	156	121	73
62	440	405	319	343	330	257	202	113
63	663	592	458	437	427	344	285	168
64	1173	875	700	657	671	615	470	258

Table 2- UBact^{15}N T_2 (in ms) values upon titration with increasing concentrations of PGP_6.

IsoT													
Time (s)	0	1704	3408	5112	8616	8520	10224	11928	13632	15336	17040	18744	20448
B-G76	4.25E+08	4.05E+08	2.55E+08	1.40E+08	1.14E+08	9.79E+07	4.55E+07	3.30E+07	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F-G76	3.93E+08	9.63E+08	1.48E+09	1.85E+09	2.14E+09	2.41E+09	2.47E+09	2.56E+09	2.62E+09	2.61E+09	2.69E+09	2.72E+09	2.66E+09
B-R74	5.01E+08	3.94E+08	3.51E+08	2.22E+08	1.48E+08	1.48E+08	8.47E+07	6.80E+07	5.95E+07	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F-R74	1.24E+08	3.13E+08	4.93E+08	6.62E+08	7.90E+08	8.23E+08	8.30E+08	9.17E+08	9.78E+08	9.36E+08	9.32E+08	9.23E+08	9.30E+08
B-K48	7.95E+08	5.57E+08	3.89E+08	3.16E+08	1.59E+08	8.50E+07	3.01E+07	9.47E+07	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
F-K48	1.53E+08	3.84E+08	5.93E+08	8.30E+08	1.05E+09	1.18E+09	1.18E+09	1.29E+09	1.24E+09	1.31E+09	1.23E+09	1.25E+09	1.22E+09
Iso	1.49E+09	1.19E+09	8.73E+08	5.76E+08	4.20E+08	2.56E+08	1.41E+08	1.08E+08	9.37E+07	1.17E+08	0.00E+00	0.00E+00	0.00E+00
I30	6.89E+08	8.64E+08	9.83E+08	1.15E+09	1.22E+09	1.33E+09	1.38E+09	1.42E+09	1.46E+09	1.45E+09	1.38E+09	1.45E+09	1.47E+09
OTUB1													
Time (s)	0	1706	3412	5118	6824	8530	10236	11942	13648	15354	17060	18766	
B-G76	9.52E+07	2.27E+08	1.20E+08	2.46E+08	1.51E+08	3.74E+07	9.27E+07	9.92E+07	0.00E+00	0.00E+00	0.00E+00	0.00E+00	
F-G76	1.31E+08	1.34E+09	9.66E+08	1.17E+09	2.56E+09	1.44E+09	2.88E+09	3.07E+09	1.56E+09	1.46E+09	3.20E+09	1.63E+09	
B-R74	1.32E+08	4.74E+08	1.20E+08	1.30E+08	3.54E+08	2.86E+07	1.42E+08	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0	
F-R74	3.80E+07	4.16E+08	2.67E+08	3.63E+08	7.45E+08	4.61E+08	9.14E+08	9.11E+08	5.16E+08	3.80E+08	1.01E+09	5.25E+08	
B-K48	2.12E+08	2.13E+08	3.76E+07	5.38E+07	1.22E+08	6.97E+06	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	
F-K48	1.42E+08	1.20E+09	8.14E+08	9.39E+08	1.89E+09	9.28E+08	1.92E+09	2.01E+09	1.04E+09	9.05E+08	2.10E+09	1.06E+09	
Iso	4.07E+08	5.70E+08	1.36E+08	1.00E+08	2.53E+08	1.53E+08	9.51E+07	5.46E+07	0.00E+00	0.00E+00	0.00E+00	0.00E+00	
I30	2.63E+08	1.44E+09	8.36E+08	8.94E+08	1.87E+09	9.63E+08	1.99E+09	2.01E+09	1.02E+09	9.73E+08	2.04E+09	1.03E+09	
Plpro													
Time (s)	4263	8526	12789	17052	21315	25578	29841	34104	38367	42630	46893		
B-G76	1.15E+09	6.99E+08	3.48E+08	3.12E+08	3.42E+08	1.07E+08	1.97E+08	1.89E+08	1.56E+08	8.60E+07	9.68E+07		
F-G76	0	2.67E+09	1.85E+09	4.20E+09	4.74E+09	2.50E+09	5.05E+09	5.31E+09	5.19E+09	5.50E+09	5.54E+09		
B-R74	1.57E+09	1.08E+09	2.70E+08	3.86E+08	2.57E+08	1.00E+08	3.37E+08	3.37E+08	1.73E+08	2.50E+08	2.10E+08		
F-R74	0	9.64E+08	6.82E+08	1.50E+09	1.63E+09	8.54E+08	1.72E+09	1.99E+09	1.80E+09	1.79E+09	2.13E+09		
B-K48	2.80E+09	2.96E+09	1.49E+09	2.73E+09	2.73E+09	1.40E+09	2.67E+09	2.42E+09	2.40E+09	2.32E+09	2.56E+09		
F-K48	0	8.39E+07	5.88E+07	1.42E+08	5.95E+08	2.81E+08	5.48E+08	7.29E+08	8.33E+08	9.16E+08	8.10E+08		
Iso	4.22E+09	4.51E+09	2.23E+09	4.42E+09	4.38E+09	2.09E+09	4.16E+09	3.91E+09	3.70E+09	3.85E+09	3.79E+09		
I30	2.16E+09	2.75E+09	1.46E+09	3.10E+09	3.30E+09	1.66E+09	3.46E+09	3.43E+09	3.71E+09	3.52E+09	3.63E+09		
Ubp6													

Time (s)	2692	5384	8076	10768	13460	16152	18844	21536	24228	26920	29612		
B-G76	7.39E+08	7.12E+08	7.44E+08	7.13E+08	6.43E+08	6.25E+08	6.37E+08	5.83E+08	5.43E+08	4.74E+08	5.34E+08		
F-G76	3.00E+08	5.28E+08	8.04E+08	9.68E+08	1.14E+09	1.25E+09	1.41E+09	1.51E+09	1.66E+09	1.81E+09	1.76E+09		
B-R74	8.85E+08	8.86E+08	8.53E+08	7.75E+08	7.67E+08	7.17E+08	7.85E+08	6.96E+08	6.24E+08	7.54E+08	5.94E+08		
F-R74	9.76E+07	1.93E+08	2.77E+08	3.08E+08	3.56E+08	4.14E+08	5.30E+08	5.42E+08	5.20E+08	7.13E+08	6.56E+08		
B-K48	1.56E+09	1.48E+09	1.48E+09	1.43E+09	1.37E+09	1.34E+09	1.35E+09	1.26E+09	1.20E+09	1.19E+09	1.24E+09		
F-K48	3.96E+06	8.09E+07	1.74E+08	1.41E+08	1.88E+08	5.90E+08	2.40E+08	4.27E+08	6.30E+08	6.73E+08	6.85E+08		
Iso	2.67E+09	2.64E+09	2.74E+09	2.56E+09	2.46E+09	2.40E+09	2.29E+09	2.28E+09	2.22E+09	2.23E+09	2.11E+09		
I30	1.07E+09	1.15E+09	1.25E+09	1.33E+09	1.32E+09	1.36E+09	1.43E+09	1.44E+09	1.47E+09	1.53E+09	1.51E+09		
Usp2													
Time (s)	2692	5384	8076	10768	13460	16152	18844	21536	24228	26920	29612	32304	34996
B-G76	3.18E+08	5.69E+08	5.45E+08	2.86E+08	4.61E+08	3.61E+08	3.61E+08	2.31E+08	1.88E+08	2.59E+08	2.46E+08	2.84E+08	2.81E+08
F-G76	1.87E+08	9.60E+08	9.49E+08	7.08E+08	1.61E+09	2.05E+09	2.04E+09	2.39E+09	2.61E+09	2.53E+09	2.66E+09	2.79E+09	2.71E+09
B-R74	3.99E+08	6.91E+08	7.29E+08	3.29E+08	4.99E+08	5.63E+08	4.16E+08	4.33E+08	3.97E+08	4.30E+08	4.68E+08	3.52E+08	2.50E+08
F-R74	6.13E+07	3.73E+08	4.05E+08	2.51E+08	5.87E+08	7.10E+08	7.78E+08	8.97E+08	9.82E+08	9.92E+08	1.10E+09	1.09E+09	1.11E+09
B-K48	4.56E+08	5.54E+08	5.52E+08	1.72E+08	2.91E+08	1.84E+08	8.21E+07	9.30E+07	6.69E+07	1.86E+08	4.90E+07	3.51E+07	0.00E+00
F-K48	3.07E+08	9.66E+08	9.60E+08	6.11E+08	1.24E+09	1.51E+09	1.56E+09	1.69E+09	1.74E+09	1.62E+09	1.75E+09	1.78E+09	1.82E+09
Iso	8.25E+08	9.46E+08	9.73E+08	3.43E+08	5.76E+08	3.78E+08	4.30E+08	2.82E+08	2.97E+08	1.64E+08	2.16E+08	2.08E+08	1.21E+08
I30	5.29E+08	1.30E+09	1.28E+09	7.49E+08	1.63E+09	1.71E+09	1.82E+09	1.81E+09	1.78E+09	1.86E+09	1.95E+09	1.97E+09	1.99E+09

Table 3-DUB assay volumes compiled. Times are given in seconds.

10^{-3} min^{-1}	IsoT	OTUB1	PLpro	Ubp6	Usp2
G76 CJ	16.1 ± 2.4	6.9 ± 2.1	5.9 ± 1.1	1.9 ± 0.2	3.9 ± 0.7
R74 CJ	13.0 ± 1.4	8.9 ± 1.4	7.9 ± 0.5	2.2 ± 0.3	3.6 ± 0.7
K48 CJ	15.9 ± 0.9	20.0 ± 3.9	1.6 ± 0.2	1.6 ± 0.4	9.5 ± 1.3
Iso	14.2 ± 0.9	17.5 ± 2.0	1.3 ± 0.2	1.0 ± 0.5	8.7 ± 1.4
G76 Free	10.7 ± 0.9	7.0 ± 0.7	6.5 ± 0.4	4.5 ± 0.2	3.9 ± 0.6
R74 Free	10.9 ± 1.3	6.3 ± 0.8	4.9 ± 0.7	3.7 ± 0.2	3.4 ± 0.4
K48 Free	9.9 ± 1.3	11.8 ± 1.3	----------	1.7 ± 0.5	5.7 ± 0.9

Table 4- Deubiquitinase cleavage rates as measured by change in NMR peak volumes. Rates are given in values of min^{-1}. Observed signals for the proximal isopeptide cleavage are shaded in gray, and distal isopeptide reporters are in white. Overall, there is good agreement between rates measured in signals that disappear and signals that appear. CJ stands for conjugated. Units are 10^{-3} $min^{-1.}$ Errors determined as error in the fit.

(1) Saeki, Y. Ubiquitin recognition by the proteasome. *J Biochem* **2017**, *161* (2), 113-124. DOI: 10.1093/jb/mvw091.
(2) Wilk, S.; Orlowski, M. Evidence that pituitary cation-sensitive neutral endopeptidase is a multicatalytic protease complex. *J Neurochem* **1983**, *40* (3), 842-849. DOI: 10.1111/j.1471-4159.1983.tb08056.x.
(3) Wenzel, T.; Eckerskorn, C.; Lottspeich, F.; Baumeister, W. Existence of a molecular ruler in proteasomes suggested by analysis of degradation products. *FEBS Lett* **1994**, *349* (2), 205-209. DOI: 10.1016/0014-5793(94)00665-2.
(4) Kisselev, A. F.; Akopian, T. N.; Woo, K. M.; Goldberg, A. L. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J Biol Chem* **1999**, *274* (6), 3363-3371. DOI: 10.1074/jbc.274.6.3363.
(5) Förster, A.; Whitby, F. G.; Hill, C. P. The pore of activated 20S proteasomes has an ordered 7-fold symmetric conformation. *EMBO J* **2003**, *22* (17), 4356-4364. DOI: 10.1093/emboj/cdg436.
(6) Bloom, J.; Pagano, M. To be or not to be ubiquitinated? *Cell Cycle* **2004**, *3* (2), 138-140.
(7) Tofaris, G. K.; Layfield, R.; Spillantini, M. G. alpha-synuclein metabolism and aggregation is linked to ubiquitin-independent degradation by the proteasome. *FEBS Lett* **2001**, *509* (1), 22-26. DOI: 10.1016/s0014-5793(01)03115-5.
(8) Choi, W. H.; de Poot, S. A.; Lee, J. H.; Kim, J. H.; Han, D. H.; Kim, Y. K.; Finley, D.; Lee, M. J. Open-gate mutants of the mammalian proteasome show enhanced ubiquitin-conjugate degradation. *Nat Commun* **2016**, *7*, 10963. DOI: 10.1038/ncomms10963.
(9) Tomko, R. J.; Funakoshi, M.; Schneider, K.; Wang, J.; Hochstrasser, M. Heterohexameric ring arrangement of the eukaryotic proteasomal ATPases: implications for proteasome structure and assembly. *Mol Cell* **2010**, *38* (3), 393-403. DOI: 10.1016/j.molcel.2010.02.035.
(10) Verma, R.; Aravind, L.; Oania, R.; McDonald, W. H.; Yates, J. R.; Koonin, E. V.; Deshaies, R. J. Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* **2002**, *298* (5593), 611-615. DOI: 10.1126/science.1075898.
(11) Lander, G. C.; Estrin, E.; Matyskiela, M. E.; Bashore, C.; Nogales, E.; Martin, A. Complete subunit architecture of the proteasome regulatory particle. *Nature* **2012**, *482* (7384), 186-191. DOI: 10.1038/nature10774.
(12) Elsasser, S.; Gali, R. R.; Schwickart, M.; Larsen, C. N.; Leggett, D. S.; Müller, B.; Feng, M. T.; Tübing, F.; Dittmar, G. A.; Finley, D. Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat Cell Biol* **2002**, *4* (9), 725-730. DOI: 10.1038/ncb845.
(13) Rosenzweig, R.; Bronner, V.; Zhang, D.; Fushman, D.; Glickman, M. H. Rpn1 and Rpn2 coordinate ubiquitin processing factors at proteasome. *J Biol Chem* **2012**, *287* (18), 14659-14671. DOI: 10.1074/jbc.M111.316323.
(14) Shi, Y.; Chen, X.; Elsasser, S.; Stocks, B. B.; Tian, G.; Lee, B. H.; Zhang, N.; de Poot, S. A.; Tuebing, F.; Sun, S.; et al. Rpn1 provides adjacent receptor sites for

substrate binding and deubiquitination by the proteasome. *Science* **2016**, *351* (6275). DOI: 10.1126/science.aad9421.
(15) Zhang, D.; Chen, T.; Ziv, I.; Rosenzweig, R.; Matiuhin, Y.; Bronner, V.; Glickman, M. H.; Fushman, D. Together, Rpn10 and Dsk2 can serve as a polyubiquitin chain-length sensor. *Mol Cell* **2009**, *36* (6), 1018-1033. DOI: 10.1016/j.molcel.2009.11.012.
(16) Husnjak, K.; Elsasser, S.; Zhang, N.; Chen, X.; Randles, L.; Shi, Y.; Hofmann, K.; Walters, K. J.; Finley, D.; Dikic, I. Proteasome subunit Rpn13 is a novel ubiquitin receptor. *Nature* **2008**, *453* (7194), 481-488. DOI: 10.1038/nature06926.
(17) Glickman, M. H.; Rubin, D. M.; Fried, V. A.; Finley, D. The regulatory particle of the Saccharomyces cerevisiae proteasome. *Mol Cell Biol* **1998**, *18* (6), 3149-3162. DOI: 10.1128/MCB.18.6.3149.
(18) Lenkinski, R. E.; Chen, D. M.; Glickson, J. D.; Goldstein, G. Nuclear magnetic resonance studies of the denaturation of ubiquitin. *Biochim Biophys Acta* **1977**, *494* (1), 126-130. DOI: 10.1016/0005-2795(77)90140-4.
(19) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of ubiquitin refined at 1.8 A resolution. *J Mol Biol* **1987**, *194* (3), 531-544. DOI: 10.1016/0022-2836(87)90679-6.
(20) Goldstein, G.; Scheid, M.; Hammerling, U.; Schlesinger, D. H.; Niall, H. D.; Boyse, E. A. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proc Natl Acad Sci U S A* **1975**, *72* (1), 11-15.
(21) 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* **1996**, *93* (2), 861-866. DOI: 10.1073/pnas.93.2.861.
(22) Peng, J.; Schwartz, D.; Elias, J. E.; Thoreen, C. C.; Cheng, D.; Marsischky, G.; Roelofs, J.; Finley, D.; Gygi, S. P. A proteomics approach to understanding protein ubiquitination. *Nat Biotechnol* **2003**, *21* (8), 921-926. DOI: 10.1038/nbt849 From NLM.
(23) Cornilescu, G.; Marquardt, J. L.; Ottiger, M.; Bax, A. Validation of Protein Structure from Anisotropic Carbonyl Chemical Shifts in a Dilute Liquid Crystalline Phase. *J Am Chem Soc* **1998**, *120* (27), 6836-6837.
(24) Pickart, C. M.; Fushman, D. Polyubiquitin chains: polymeric protein signals. *Curr Opin Chem Biol* **2004**, *8* (6), 610-616. DOI: 10.1016/j.cbpa.2004.09.009.
(25) Xu, P.; Duong, D. M.; Seyfried, N. T.; Cheng, D.; Xie, Y.; Robert, J.; Rush, J.; Hochstrasser, M.; Finley, D.; Peng, J. Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell* **2009**, *137* (1), 133-145. DOI: 10.1016/j.cell.2009.01.041.
(26) Ohtake, F.; Saeki, Y.; Sakamoto, K.; Ohtake, K.; Nishikawa, H.; Tsuchiya, H.; Ohta, T.; Tanaka, K.; Kanno, J. Ubiquitin acetylation inhibits polyubiquitin chain elongation. *EMBO Rep* **2015**, *16* (2), 192-201. DOI: 10.15252/embr.201439152.
(27) 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* **2018**, *8* (1), 2651. DOI: 10.1038/s41598-018-20860-w.
(28) Linthwaite, V. L.; Pawloski, W.; Pegg, H. B.; Townsend, P. D.; Thomas, M. J.; So, V. K. H.; Brown, A. P.; Hodgson, D. R. W.; Lorimer, G. H.; Fushman, D.; et al.

Ubiquitin is a carbon dioxide-binding protein. *Sci Adv* **2021**, *7* (39), eabi5507. DOI: 10.1126/sciadv.abi5507.
(29) Pawloski, W.; Komiyama, T.; Kougentakis, C.; Majumdar, A.; Fushman, D. Site-Specific Detection and Characterization of Ubiquitin Carbamylation. *Biochemistry* **2022**, *61* (8), 712-721. DOI: 10.1021/acs.biochem.2c00085.
(30) Chen, Z.; Pickart, C. M. A 25-kilodalton ubiquitin carrier protein (E2) catalyzes multi-ubiquitin chain synthesis via lysine 48 of ubiquitin. *J Biol Chem* **1990**, *265* (35), 21835-21842.
(31) Pickart, C. M.; Eddins, M. J. Ubiquitin: structures, functions, mechanisms. *Biochim Biophys Acta* **2004**, *1695* (1-3), 55-72. DOI: 10.1016/j.bbamcr.2004.09.019.
(32) Amerik, A. Y.; Swaminathan, S.; Krantz, B. A.; Wilkinson, K. D.; Hochstrasser, M. In vivo disassembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome. *EMBO J* **1997**, *16* (16), 4826-4838. DOI: https://doi.org/10.1093/emboj/16.16.4826 (acccessed 2023/08/24).
(33) Keusekotten, K.; Elliott, Paul R.; Glockner, L.; Fiil, Berthe K.; Damgaard, Rune B.; Kulathu, Y.; Wauer, T.; Hospenthal, Manuela K.; Gyrd-Hansen, M.; Krappmann, D.; et al. OTULIN Antagonizes LUBAC Signaling by Specifically Hydrolyzing Met1-Linked Polyubiquitin. *Cell* **2013**, *153* (6), 1312-1326. DOI: https://doi.org/10.1016/j.cell.2013.05.014.
(34) Kirisako, T.; Kamei, K.; Murata, S.; Kato, M.; Fukumoto, H.; Kanie, M.; Sano, S.; Tokunaga, F.; Tanaka, K.; Iwai, K. A ubiquitin ligase complex assembles linear polyubiquitin chains. *EMBO J* **2006**, *25* (20), 4877-4887. DOI: 10.1038/sj.emboj.7601360.
(35) Komander, D.; Reyes-Turcu, F.; Licchesi, J. D. F.; Odenwaelder, P.; Wilkinson, K. D.; Barford, D. Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO reports* **2009**, *10* (5), 466-473. DOI: https://doi.org/10.1038/embor.2009.55 (acccessed 2023/08/24).
(36) Rahighi, S.; Ikeda, F.; Kawasaki, M.; Akutsu, M.; Suzuki, N.; Kato, R.; Kensche, T.; Uejima, T.; Bloor, S.; Komander, D.; et al. Specific Recognition of Linear Ubiquitin Chains by NEMO Is Important for NF-κB Activation. *Cell* **2009**, *136* (6), 1098-1109. DOI: 10.1016/j.cell.2009.03.007 (acccessed 2023/08/24).
(37) Fiil, B. K.; Gyrd-Hansen, M. Met1-linked ubiquitination in immune signalling. *The FEBS Journal* **2014**, *281* (19), 4337-4350. DOI: https://doi.org/10.1111/febs.12944 (acccessed 2023/08/24).
(38) Emmerich, C. H.; Ordureau, A.; Strickson, S.; Arthur, J. S.; Pedrioli, P. G.; Komander, D.; Cohen, P. Activation of the canonical IKK complex by K63/M1-linked hybrid ubiquitin chains. *Proc Natl Acad Sci U S A* **2013**, *110* (38), 15247-15252. DOI: 10.1073/pnas.1314715110.
(39) Döffinger, R.; Smahi, A.; Bessia, C.; Geissmann, F.; Feinberg, J.; Durandy, A.; Bodemer, C.; Kenwrick, S.; Dupuis-Girod, S.; Blanche, S.; et al. X-linked anhidrotic ectodermal dysplasia with immunodeficiency is caused by impaired NF-kappaB signaling. *Nat Genet* **2001**, *27* (3), 277-285. DOI: 10.1038/85837 From NLM.
(40) Filipe-Santos, O.; Bustamante, J.; Haverkamp, M. H.; Vinolo, E.; Ku, C. L.; Puel, A.; Frucht, D. M.; Christel, K.; von Bernuth, H.; Jouanguy, E.; et al. X-linked susceptibility to mycobacteria is caused by mutations in NEMO impairing CD40-

dependent IL-12 production. *J Exp Med* **2006**, *203* (7), 1745-1759. DOI: 10.1084/jem.20060085 From NLM.
(41) Gerlach, B.; Cordier, S. M.; Schmukle, A. C.; Emmerich, C. H.; Rieser, E.; Haas, T. L.; Webb, A. I.; Rickard, J. A.; Anderton, H.; Wong, W. W. L.; et al. Linear ubiquitination prevents inflammation and regulates immune signalling. *Nature* **2011**, *471* (7340), 591-596. DOI: 10.1038/nature09816.
(42) Seymour, R. E.; Hasham, M. G.; Cox, G. A.; Shultz, L. D.; Hogenesch, H.; Roopenian, D. C.; Sundberg, J. P. Spontaneous mutations in the mouse Sharpin gene result in multiorgan inflammation, immune system dysregulation and dermatitis. *Genes Immun* **2007**, *8* (5), 416-421. DOI: 10.1038/sj.gene.6364403 From NLM.
(43) Michel, M. A.; Swatek, K. N.; Hospenthal, M. K.; Komander, D. Ubiquitin Linkage-Specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. *Mol Cell* **2017**, *68* (1), 233-246.e235. DOI: 10.1016/j.molcel.2017.08.020.
(44) Durcan, T. M.; Tang, M. Y.; Pérusse, J. R.; Dashti, E. A.; Aguileta, M. A.; McLelland, G.-L.; Gros, P.; Shaler, T. A.; Faubert, D.; Coulombe, B.; et al. USP8 regulates mitophagy by removing K6-linked ubiquitin conjugates from parkin. *EMBO J* **2014**, *33* (21), 2473-2491. DOI: https://doi.org/10.15252/embj.201489729.
(45) Yuan, Y.; Miao, Y.; Qian, L.; Zhang, Y.; Liu, C.; Liu, J.; Zuo, Y.; Feng, Q.; Guo, T.; Zhang, L.; et al. Targeting UBE4A Revives Viperin Protein in Epithelium to Enhance Host Antiviral Defense. *Mol Cell* **2020**, *77* (4), 734-747.e737. DOI: 10.1016/j.molcel.2019.11.003 (acccessed 2023/08/26).
(46) Srivastava, D.; Chakrabarti, O. Mahogunin-mediated α-tubulin ubiquitination via noncanonical K6 linkage regulates microtubule stability and mitotic spindle orientation. *Cell Death Dis* **2014**, *5* (2), e1064-e1064. DOI: 10.1038/cddis.2014.1.
(47) Hong, S.-Y.; Kao, Y.-R.; Lee, T.-C.; Wu, C.-W. Upregulation of E3 Ubiquitin Ligase CBLC Enhances EGFR Dysregulation and Signaling in Lung Adenocarcinoma. *Cancer Res* **2018**, *78* (17), 4984-4996. DOI: 10.1158/0008-5472.CAN-17-3858 (acccessed 8/26/2023).
(48) Zhang, Z.; Wang, D.; Wang, P.; Zhao, Y.; You, F. OTUD1 Negatively Regulates Type I IFN Induction by Disrupting Noncanonical Ubiquitination of IRF3. *J Immunol* **2020**, *204* (7), 1904-1918. DOI: 10.4049/jimmunol.1900305 (acccessed 8/26/2023).
(49) Oh, E.; Akopian, D.; Rape, M. Principles of Ubiquitin-Dependent Signaling. *Annual Review of Cell and Developmental Biology* **2018**, *34* (1), 137-162. DOI: 10.1146/annurev-cellbio-100617-062802 (acccessed 2023/08/26).
(50) Jin, L.; Williamson, A.; Banerjee, S.; Philipp, I.; Rape, M. Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell* **2008**, *133* (4), 653-665. DOI: 10.1016/j.cell.2008.04.012 From NLM.
(51) Boughton, A. J.; Krueger, S.; Fushman, D. Branching via K11 and K48 Bestows Ubiquitin Chains with a Unique Interdomain Interface and Enhanced Affinity for Proteasomal Subunit Rpn1. *Structure* **2020**, *28* (1), 29-43.e26. DOI: 10.1016/j.str.2019.10.008.
(52) Qin, Y.; Zhou, M.-T.; Hu, M.-M.; Hu, Y.-H.; Zhang, J.; Guo, L.; Zhong, B.; Shu, H.-B. RNF26 Temporally Regulates Virus-Triggered Type I Interferon Induction by Two Distinct Mechanisms. *PLOS Pathogens* **2014**, *10* (9), e1004358. DOI: 10.1371/journal.ppat.1004358.

(53) Jin, S.; Tian, S.; Chen, Y.; Zhang, C.; Xie, W.; Xia, X.; Cui, J.; Wang, R.-F. USP19 modulates autophagy and antiviral immune responses by deubiquitinating Beclin-1. *EMBO J* **2016**, *35* (8), 866-880. DOI: https://doi.org/10.15252/embj.201593596 (acccessed 2023/08/26).
(54) Moniz, S.; Bandarra, D.; Biddlestone, J.; Campbell, K. J.; Komander, D.; Bremm, A.; Rocha, S. Cezanne regulates E2F1-dependent HIF2α expression. *Journal of Cell Science* **2015**, *128* (16), 3082-3093. DOI: 10.1242/jcs.168864 (acccessed 8/26/2023).
(55) Castañeda, C. A.; Dixon, E. K.; Walker, O.; Chaturvedi, A.; Nakasone, M. A.; Curtis, J. E.; Reed, M. R.; Krueger, S.; Cropp, T. A.; Fushman, D. Linkage via K27 Bestows Ubiquitin Chains with Unique Properties among Polyubiquitins. *Structure* **2016**, *24* (3), 423-436. DOI: 10.1016/j.str.2016.01.007 From NLM.
(56) Gatti, M.; Pinato, S.; Maiolica, A.; Rocchio, F.; Prato, Maria G.; Aebersold, R.; Penengo, L. RNF168 Promotes Noncanonical K27 Ubiquitination to Signal DNA Damage. *Cell Rep* **2015**, *10* (2), 226-238. DOI: 10.1016/j.celrep.2014.12.021 (acccessed 2023/08/26).
(57) Zhang, Y.; Zhang, H.; Zheng, G.-L.; Yang, Q.; Yu, S.; Wang, J.; Li, S.; Li, L.-F.; Qiu, H.-J. Porcine RING Finger Protein 114 Inhibits Classical Swine Fever Virus Replication via K27-Linked Polyubiquitination of Viral NS4B. *Journal of Virology* **2019**, *93* (21), 10.1128/jvi.01248-01219. DOI: 10.1128/jvi.01248-19 (acccessed 2023/08/26).
(58) Bai, Y.; Li, L.; Shan, T.; Zhang, Y.; Chen, X.; Gao, F.; Jiang, Y.; Zhou, Y.; Li, G.; Yu, L.; et al. Proteasomal degradation of nonstructural protein 12 by RNF114 suppresses porcine reproductive and respiratory syndrome virus replication. *Veterinary Microbiology* **2020**, *246*, 108746. DOI: https://doi.org/10.1016/j.vetmic.2020.108746.
(59) Li, Z.; Fan, S.; Wang, J.; Chen, X.; Liao, Q.; Liu, X.; Ouyang, G.; Cao, H.; Xiao, W. Zebrafish F-box Protein fbxo3 Negatively Regulates Antiviral Response through Promoting K27-Linked Polyubiquitination of the Transcription Factors irf3 and irf7. *J Immunol* **2020**, *205* (7), 1897-1908. DOI: 10.4049/jimmunol.2000305 (acccessed 8/26/2023).
(60) Wu, Y.; Jin, S.; Liu, Q.; Zhang, Y.; Ma, L.; Zhao, Z.; Yang, S.; Li, Y.-P.; Cui, J. Selective autophagy controls the stability of transcription factor IRF3 to balance type I interferon production and immune suppression. *Autophagy* **2021**, *17* (6), 1379-1392. DOI: 10.1080/15548627.2020.1761653.
(61) Nucifora, F. C.; Nucifora, L. G.; Ng, C.-H.; Arbez, N.; Guo, Y.; Roby, E.; Shani, V.; Engelender, S.; Wei, D.; Wang, X.-F.; et al. Ubiqutination via K27 and K29 chains signals aggregation and neuronal protection of LRRK2 by WSB1. *Nature Communications* **2016**, *7* (1), 11792. DOI: 10.1038/ncomms11792.
(62) Zucchelli, S.; Codrich, M.; Marcuzzi, F.; Pinto, M.; Vilotti, S.; Biagioli, M.; Ferrer, I.; Gustincich, S. TRAF6 promotes atypical ubiquitination of mutant DJ-1 and alpha-synuclein and is localized to Lewy bodies in sporadic Parkinson's disease brains. *Human Molecular Genetics* **2010**, *19* (19), 3759-3770. DOI: 10.1093/hmg/ddq290 (acccessed 8/26/2023).
(63) Zucchelli, S.; Marcuzzi, F.; Codrich, M.; Agostoni, E.; Vilotti, S.; Biagioli, M.; Pinto, M.; Carnemolla, A.; Santoro, C.; Gustincich, S.; et al. Tumor Necrosis Factor

Receptor-associated Factor 6 (TRAF6) Associates with Huntingtin Protein and Promotes Its Atypical Ubiquitination to Enhance Aggregate Formation *. *J Biol Chem* **2011**, *286* (28), 25108-25117. DOI: 10.1074/jbc.M110.187591 (acccessed 2023/08/26).
(64) Huang, B.; Baek, S. H. Trim13 Potentiates Toll-Like Receptor 2-Mediated Nuclear Factor. *Mol Pharmacol* **2017**, *91* (4), 307-316. DOI: 10.1124/mol.116.106716.
(65) Nazio, F.; Carinci, M.; Valacca, C.; Bielli, P.; Strappazzon, F.; Antonioli, M.; Ciccosanti, F.; Rodolfo, C.; Campello, S.; Fimia, G. M.; et al. Fine-tuning of ULK1 mRNA and protein levels is required for autophagy oscillation. *Journal of Cell Biology* **2016**, *215* (6), 841-856. DOI: 10.1083/jcb.201605089 (acccessed 8/26/2023).
(66) Tauriello, D. V. F.; Maurice, M. M. The various roles of ubiquitin in Wnt pathway regulation. *Cell Cycle* **2010**, *9* (18), 3724-3733. DOI: 10.4161/cc.9.18.13204.
(67) Fei, C.; Li, Z.; Li, C.; Chen, Y.; Chen, Z.; He, X.; Mao, L.; Wang, X.; Zeng, R.; Li, L. Smurf1-Mediated Lys29-Linked Nonproteolytic Polyubiquitination of Axin Negatively Regulates Wnt/β-Catenin Signaling. *Molecular and Cellular Biology* **2013**, *33* (20), 4095-4105. DOI: 10.1128/MCB.00418-13.
(68) Besche, H. C.; Sha, Z.; Kukushkin, N. V.; Peth, A.; Hock, E.-M.; Kim, W.; Gygi, S.; Gutierrez, J. A.; Liao, H.; Dick, L.; et al. Autoubiquitination of the 26S Proteasome on Rpn13 Regulates Breakdown of Ubiquitin Conjugates. *EMBO J* **2014**, *33* (10), 1159-1176. DOI: https://doi.org/10.1002/embj.201386906 (acccessed 2023/08/26).
(69) Kim, W.; Bennett, Eric J.; Huttlin, Edward L.; Guo, A.; Li, J.; Possemato, A.; Sowa, Mathew E.; Rad, R.; Rush, J.; Comb, Michael J.; et al. Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome. *Mol Cell* **2011**, *44* (2), 325-340. DOI: 10.1016/j.molcel.2011.08.025 (acccessed 2023/08/26).
(70) Lin, M.; Zhao, Z.; Yang, Z.; Meng, Q.; Tan, P.; Xie, W.; Qin, Y.; Wang, R.-F.; Cui, J. USP38 Inhibits Type I Interferon Signaling by Editing TBK1 Ubiquitination through NLRP4 Signalosome. *Mol Cell* **2016**, *64* (2), 267-281. DOI: 10.1016/j.molcel.2016.08.029 (acccessed 2023/08/27).
(71) Liu, S.; Jiang, M.; Wang, W.; Liu, W.; Song, X.; Ma, Z.; Zhang, S.; Liu, L.; Liu, Y.; Cao, X. Nuclear RNF2 inhibits interferon function by promoting K33-linked STAT1 disassociation from DNA. *Nat Immunol* **2018**, *19* (1), 41-52. DOI: 10.1038/s41590-017-0003-0.
(72) Gao, L.; Wang, L.; Dai, T.; Jin, K.; Zhang, Z.; Wang, S.; Xie, F.; Fang, P.; Yang, B.; Huang, H.; et al. Tumor-derived exosomes antagonize innate antiviral immunity. *Nat Immunol* **2018**, *19* (3), 233-245. DOI: 10.1038/s41590-017-0043-5.
(73) Nibe, Y.; Oshima, S.; Kobayashi, M.; Maeyashiki, C.; Matsuzawa, Y.; Otsubo, K.; Matsuda, H.; Aonuma, E.; Nemoto, Y.; Nagaishi, T.; et al. Novel polyubiquitin imaging system, PolyUb-FC, reveals that K33-linked polyubiquitin is recruited by SQSTM1/p62. *Autophagy* **2018**, *14* (2), 347-358. DOI: 10.1080/15548627.2017.1407889.
(74) Heath, R. J.; Goel, G.; Baxt, L. A.; Rush, J. S.; Mohanan, V.; Paulus, G. L. C.; Jani, V.; Lassen, K. G.; Xavier, R. J. RNF166 Determines Recruitment of Adaptor

Proteins during Antibacterial Autophagy. *Cell Rep* **2016**, *17* (9), 2183-2194. DOI: 10.1016/j.celrep.2016.11.005 (acccessed 2023/08/27).
(75) Abe, T.; Umeki, I.; Kanno, S.-i.; Inoue, S.-i.; Niihori, T.; Aoki, Y. LZTR1 facilitates polyubiquitination and degradation of RAS-GTPases. *Cell Death Differ* **2020**, *27* (3), 1023-1035. DOI: 10.1038/s41418-019-0395-5.
(76) Chau, V.; Tobias, J. W.; Bachmair, A.; Marriott, D.; Ecker, D. J.; Gonda, D. K.; Varshavsky, A. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **1989**, *243* (4898), 1576-1583. DOI: 10.1126/science.2538923.
(77) Thrower, J. S.; Hoffman, L.; Rechsteiner, M.; Pickart, C. M. Recognition of the polyubiquitin proteolytic signal. *EMBO J* **2000**, *19* (1), 94-102. DOI: 10.1093/emboj/19.1.94.
(78) Pohl, C.; Dikic, I. Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science* **2019**, *366* (6467), 818-822. DOI: 10.1126/science.aax3769 (acccessed 2023/08/27).
(79) Song, J.; Herrmann, J. M.; Becker, T. Quality control of the mitochondrial proteome. *Nat Rev Mol Cel Biol* **2021**, *22* (1), 54-70. DOI: 10.1038/s41580-020-00300-2.
(80) Hu, H.; Sun, S.-C. Ubiquitin signaling in immune responses. *Cell Res* **2016**, *26* (4), 457-483. DOI: 10.1038/cr.2016.40.
(81) Schwob, E.; Böhm, T.; Mendenhall, M. D.; Nasmyth, K. The B-type cyclin kinase inhibitor $p40^{SIC1}$ controls the G1 to S transition in S. cerevisiae. *Cell* **1994**, *79* (2), 233-244. DOI: 10.1016/0092-8674(94)90193-7 (acccessed 2023/08/27).
(82) Cross, F. R.; Schroeder, L.; Bean, J. M. Phosphorylation of the Sic1 inhibitor of B-type cyclins in Saccharomyces cerevisiae is not essential but contributes to cell cycle robustness. *Genetics* **2007**, *176* (3), 1541-1555. DOI: 10.1534/genetics.107.073494 From NLM.
(83) Zhao, X. C.; Wang, G. Z.; Wen, Z. S.; Zhou, Y. C.; Hu, Q.; Zhang, B.; Qu, L. W.; Gao, S. H.; Liu, J.; Ma, L.; et al. Systematic identification of CDC34 that functions to stabilize EGFR and promote lung carcinogenesis. *EBioMedicine* **2020**, *53*, 102689. DOI: 10.1016/j.ebiom.2020.102689 From NLM.
(84) Eliseeva, E.; Pati, D.; Diccinanni, M. B.; Yu, A. L.; Mohsin, S. K.; Margolin, J. F.; Plon, S. E. Expression and Localization of the CDC34 Ubiquitin-conjugating Enzyme in Pediatric Acute Lymphoblastic Leukemia1. *Cell Growth Differ* **2001**, *12* (8), 427-433. (acccessed 8/27/2023).
(85) Morishima-Kawashima, M.; Hasegawa, M.; Takio, K.; Suzuki, M.; Titani, K.; Ihara, Y. Ubiquitin is conjugated with amino-terminally processed tau in paired helical filaments. *Neuron* **1993**, *10* (6), 1151-1160. DOI: https://doi.org/10.1016/0896-6273(93)90063-W.
(86) Parolini, F.; Tira, R.; Barracchia, C. G.; Munari, F.; Capaldi, S.; D'Onofrio, M.; Assfalg, M. Ubiquitination of Alzheimer's-related tau protein affects liquid-liquid phase separation in a site- and cofactor-dependent manner. *Int J Biol Macromol* **2022**, *201*, 173-181. DOI: https://doi.org/10.1016/j.ijbiomac.2021.12.191.
(87) Hasegawa, M.; Fujiwara, H.; Nonaka, T.; Wakabayashi, K.; Takahashi, H.; Lee, V. M. Y.; Trojanowski, J. Q.; Mann, D.; Iwatsubo, T. Phosphorylated α-Synuclein Is

Ubiquitinated in α-Synucleinopathy Lesions*. *J Biol Chem* **2002**, *277* (50), 49071-49076. DOI: https://doi.org/10.1074/jbc.M208046200.
(88) Nonaka, T.; Iwatsubo, T.; Hasegawa, M. Ubiquitination of alpha-synuclein. *Biochemistry* **2005**, *44* (1), 361-368. DOI: 10.1021/bi0485528 From NLM.
(89) Bhat, K. P.; Yan, S.; Wang, C. E.; Li, S.; Li, X. J. Differential ubiquitination and degradation of huntingtin fragments modulated by ubiquitin-protein ligase E3A. *Proc Natl Acad Sci U S A* **2014**, *111* (15), 5706-5711. DOI: 10.1073/pnas.1402215111 From NLM.
(90) Juenemann, K.; Jansen, A. H. P.; van Riel, L.; Merkx, R.; Mulder, M. P. C.; An, H.; Statsyuk, A.; Kirstein, J.; Ovaa, H.; Reits, E. A. Dynamic recruitment of ubiquitin to mutant huntingtin inclusion bodies. *Sci Rep* **2018**, *8* (1), 1405. DOI: 10.1038/s41598-018-19538-0 From NLM.
(91) Hakim-Eshed, V.; Boulos, A.; Cohen-Rosenzweig, C.; Yu-Taeger, L.; Ziv, T.; Kwon, Y. T.; Riess, O.; Phuc Nguyen, H. H.; Ziv, N. E.; Ciechanover, A. Site-specific ubiquitination of pathogenic huntingtin attenuates its deleterious effects. *P Natl Acad Sci USA* **2020**, *117* (31), 18661-18669. DOI: 10.1073/pnas.2007667117 (acccessed 2023/08/27).
(92) Lowe, J.; Mayer, R. J.; Landon, M. Ubiquitin in neurodegenerative diseases. *Brain Pathol* **1993**, *3* (1), 55-65. DOI: 10.1111/j.1750-3639.1993.tb00726.x From NLM.
(93) Bence, N. F.; Sampat, R. M.; Kopito, R. R. Impairment of the Ubiquitin-Proteasome System by Protein Aggregation. *Science* **2001**, *292* (5521), 1552-1555. DOI: 10.1126/science.292.5521.1552 (acccessed 2023/08/27).
(94) Hofmann, R. M.; Pickart, C. M. Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **1999**, *96* (5), 645-653. DOI: 10.1016/s0092-8674(00)80575-9 From NLM.
(95) Chen, H. T.; Bhandoola, A.; Difilippantonio, M. J.; Zhu, J.; Brown, M. J.; Tai, X.; Rogakou, E. P.; Brotz, T. M.; Bonner, W. M.; Ried, T.; et al. Response to RAG-mediated VDJ cleavage by NBS1 and gamma-H2AX. *Science* **2000**, *290* (5498), 1962-1965. DOI: 10.1126/science.290.5498.1962 From NLM.
(96) Petersen, S.; Casellas, R.; Reina-San-Martin, B.; Chen, H. T.; Difilippantonio, M. J.; Wilson, P. C.; Hanitsch, L.; Celeste, A.; Muramatsuk, M.; Pilch, D. R.; et al. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature* **2001**, *414* (6864), 660-665. DOI: 10.1038/414660a From NLM.
(97) Rogakou, E. P.; Nieves-Neira, W.; Boon, C.; Pommier, Y.; Bonner, W. M. Initiation of DNA fragmentation during apoptosis induces phosphorylation of H2AX histone at serine 139. *J Biol Chem* **2000**, *275* (13), 9390-9395. DOI: 10.1074/jbc.275.13.9390 From NLM.
(98) Rogakou, E. P.; Pilch, D. R.; Orr, A. H.; Ivanova, V. S.; Bonner, W. M. DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139*. *J Biol Chem* **1998**, *273* (10), 5858-5868. DOI: https://doi.org/10.1074/jbc.273.10.5858.
(99) Wang, B.; Elledge, S. J. Ubc13/Rnf8 ubiquitin ligases control foci formation of the Rap80/Abraxas/Brca1/Brcc36 complex in response to DNA damage. *P Natl Acad*

Sci USA **2007**, *104* (52), 20759-20763. DOI: 10.1073/pnas.0710061104 (acccessed 2023/08/28).
(100) Wu, J.; Lu, L. Y.; Yu, X. The role of BRCA1 in DNA damage response. *Protein Cell* **2010**, *1* (2), 117-123. DOI: 10.1007/s13238-010-0010-5 From NLM.
(101) Xia, Z.-P.; Sun, L.; Chen, X.; Pineda, G.; Jiang, X.; Adhikari, A.; Zeng, W.; Chen, Z. J. Direct activation of protein kinases by unanchored polyubiquitin chains. *Nature* **2009**, *461* (7260), 114-119. DOI: 10.1038/nature08247.
(102) Johnston, S. C.; Riddle, S. M.; Cohen, R. E.; Hill, C. P. Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J* **1999**, *18* (14), 3877-3887. DOI: 10.1093/emboj/18.14.3877.
(103) Mevissen, T. E.; Hospenthal, M. K.; Geurink, P. P.; Elliott, P. R.; Akutsu, M.; Arnaudo, N.; Ekkebus, R.; Kulathu, Y.; Wauer, T.; El Oualid, F.; et al. OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. *Cell* **2013**, *154* (1), 169-184. DOI: 10.1016/j.cell.2013.05.046.
(104) Zeng, C.; Zhao, C.; Ge, F.; Li, Y.; Cao, J.; Ying, M.; Lu, J.; He, Q.; Yang, B.; Dai, X.; et al. Machado-Joseph Deubiquitinases: From Cellular Functions to Potential Therapy Targets. *Front Pharmacol* **2020**, *11*, 1311. DOI: 10.3389/fphar.2020.01311.
(105) Abdul Rehman, S. A.; Kristariyanto, Y. A.; Choi, S. Y.; Nkosi, P. J.; Weidlich, S.; Labib, K.; Hofmann, K.; Kulathu, Y. MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes. *Mol Cell* **2016**, *63* (1), 146-155. DOI: 10.1016/j.molcel.2016.05.009.
(106) Shrestha, R. K.; Ronau, J. A.; Davies, C. W.; Guenette, R. G.; Strieter, E. R.; Paul, L. N.; Das, C. Insights into the mechanism of deubiquitination by JAMM deubiquitinases from cocrystal structures of the enzyme with the substrate and product. *Biochemistry* **2014**, *53* (19), 3199-3217. DOI: 10.1021/bi5003162.
(107) de la Peña, A. H.; Goodall, E. A.; Gates, S. N.; Lander, G. C.; Martin, A. Substrate-engaged 26S proteasome structures reveal mechanisms for ATP-hydrolysis-drive translocation. *Science* **2018**, *362* (6418). DOI: 10.1126/science.aav0725.
(108) Benoist, P.; Müller, A.; Diem, H. G.; Schwencke, J. High-molecular-mass multicatalytic proteinase complexes produced by the nitrogen-fixing actinomycete Frankia strain BR. *J Bacteriol* **1992**, *174* (5), 1495-1504. DOI: 10.1128/jb.174.5.1495-1504.1992.
(109) Pouch, M.-N.; Cournoyer, B.; Baumeister, W. Characterization of the 20S proteasome from the actinomycete *Frankia*. *Mol Microbiol* **2002**, *35* (2), 368-377.
(110) Benaroudj, N.; Goldberg, A. L. PAN, the proteasome-activating nucleotidase from archaebacteria, is a protein-unfolding molecular chaperone. *Nat Cell Biol* **2000**, *2* (11), 833-839. DOI: 10.1038/35041081.
(111) Wang, T.; Darwin, K. H.; Li, H. Binding-induced folding of prokaryotic ubiquitin-like protein on the Mycobacterium proteasomal ATPase targets substrates for degradation. *Nat Struct Mol Biol* **2010**, *17* (11), 1352-1357. DOI: 10.1038/nsmb.1918.
(112) Wu, Y.; Hu, K.; Li, D.; Bai, L.; Yang, S.; Jastrab, J. B.; Xiao, S.; Hu, Y.; Zhang, S.; Darwin, K. H.; et al. Mycobacterium tuberculosis proteasomal ATPase Mpa has a β-grasp domain that hinders docking with the proteasome core protease. *Mol Microbiol* **2017**, *105* (2), 227-241. DOI: 10.1111/mmi.13695.

(113) Kim, M.; Otsubo, R.; Morikawa, H.; Nishide, A.; Takagi, K.; Sasakawa, C.; Mizushima, T. Bacterial effectors and their functions in the ubiquitin-proteasome system: insight from the modes of substrate recognition. *Cells* **2014**, *3* (3), 848-864. DOI: 10.3390/cells3030848.
(114) Ranjan, N.; Damberger, F. F.; Sutter, M.; Allain, F. H.; Weber-Ban, E. Solution structure and activation mechanism of ubiquitin-like small archaeal modifier proteins. *J Mol Biol* **2011**, *405* (4), 1040-1055. DOI: 10.1016/j.jmb.2010.11.040.
(115) Humbard, M. A.; Miranda, H. V.; Lim, J. M.; Krause, D. J.; Pritz, J. R.; Zhou, G.; Chen, S.; Wells, L.; Maupin-Furlow, J. A. Ubiquitin-like small archaeal modifier proteins (SAMPs) in Haloferax volcanii. *Nature* **2010**, *463* (7277), 54-60. DOI: 10.1038/nature08659.
(116) Bode, N. J.; Darwin, K. H. The Pup-Proteasome System of Mycobacteria. *Microbiol Spectr* **2014**, *2* (5). DOI: 10.1128/microbiolspec.MGM2-0008-2013.
(117) Pearce, M. J.; Mintseris, J.; Ferreyra, J.; Gygi, S. P.; Darwin, K. H. Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis. *Science* **2008**, *322* (5904), 1104-1107. DOI: 10.1126/science.1163885.
(118) Chen, X.; Solomon, W. C.; Kang, Y.; Cerda-Maira, F.; Darwin, K. H.; Walters, K. J. Prokaryotic ubiquitin-like protein pup is intrinsically disordered. *J Mol Biol* **2009**, *392* (1), 208-217. DOI: 10.1016/j.jmb.2009.07.018.
(119) Barandun, J.; Damberger, F. F.; Delley, C. L.; Laederach, J.; Allain, F. H.; Weber-Ban, E. Prokaryotic ubiquitin-like protein remains intrinsically disordered when covalently attached to proteasomal target proteins. *BMC Struct Biol* **2017**, *17* (1), 1. DOI: 10.1186/s12900-017-0072-1.
(120) Guth, E.; Thommen, M.; Weber-Ban, E. Mycobacterial ubiquitin-like protein ligase PafA follows a two-step reaction pathway with a phosphorylated pup intermediate. *J Biol Chem* **2011**, *286* (6), 4412-4419. DOI: 10.1074/jbc.M110.189282.
(121) Sutter, M.; Damberger, F. F.; Imkamp, F.; Allain, F. H.; Weber-Ban, E. Prokaryotic ubiquitin-like protein (Pup) is coupled to substrates via the side chain of its C-terminal glutamate. *J Am Chem Soc* **2010**, *132* (16), 5610-5612. DOI: 10.1021/ja910546x.
(122) Cerda-Maira, F. A.; Pearce, M. J.; Fuortes, M.; Bishai, W. R.; Hubbard, S. R.; Darwin, K. H. Molecular analysis of the prokaryotic ubiquitin-like protein (Pup) conjugation pathway in Mycobacterium tuberculosis. *Mol Microbiol* **2010**, *77* (5), 1123-1135. DOI: 10.1111/j.1365-2958.2010.07276.x.
(123) Striebel, F.; Imkamp, F.; Sutter, M.; Steiner, M.; Mamedov, A.; Weber-Ban, E. Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. *Nat Struct Mol Biol* **2009**, *16* (6), 647-651. DOI: 10.1038/nsmb.1597.
(124) Imkamp, F.; Striebel, F.; Sutter, M.; Ozcelik, D.; Zimmermann, N.; Sander, P.; Weber-Ban, E. Dop functions as a depupylase in the prokaryotic ubiquitin-like modification pathway. *EMBO Rep* **2010**, *11* (10), 791-797. DOI: 10.1038/embor.2010.119.
(125) Cerda-Maira, F. A.; McAllister, F.; Bode, N. J.; Burns, K. E.; Gygi, S. P.; Darwin, K. H. Reconstitution of the Mycobacterium tuberculosis pupylation pathway

in Escherichia coli. *EMBO Rep* **2011**, *12* (8), 863-870. DOI: 10.1038/embor.2011.109.
(126) Sutter, M.; Striebel, F.; Damberger, F. F.; Allain, F. H.; Weber-Ban, E. A distinct structural region of the prokaryotic ubiquitin-like protein (Pup) is recognized by the N-terminal domain of the proteasomal ATPase Mpa. *FEBS Lett* **2009**, *583* (19), 3151-3157. DOI: 10.1016/j.febslet.2009.09.020.
(127) Striebel, F.; Hunkeler, M.; Summer, H.; Weber-Ban, E. The mycobacterial Mpa-proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. *EMBO J* **2010**, *29* (7), 1262-1271. DOI: 10.1038/emboj.2010.23.
(128) Lehmann, G.; Udasin, R. G.; Livneh, I.; Ciechanover, A. Identification of UBact, a ubiquitin-like protein, along with other homologous components of a conjugation system and the proteasome in different gram-negative bacteria. *Biochem Biophys Res Commun* **2017**, *483* (3), 946-950. DOI: 10.1016/j.bbrc.2017.01.037.
(129) Darwin, K. H.; Ehrt, S.; Gutierrez-Ramos, J. C.; Weich, N.; Nathan, C. F. The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. *Science* **2003**, *302* (5652), 1963-1966. DOI: 10.1126/science.1091176.
(130) Janssen, G. V.; Zhang, S.; Merkx, R.; Schiesswohl, C.; Chatterjee, C.; Darwin, K. H.; Geurink, P. P.; van der Heden van Noort, G. J.; Ovaa, H. Development of Tyrphostin Analogues to Study Inhibition of the Mycobacterium tuberculosis Pup Proteasome System*. *Chembiochem* **2021**, *22* (21), 3082-3089. DOI: 10.1002/cbic.202100333.
(131) Cheng, Y.; Pieters, J. Novel proteasome inhibitors as potential drugs to combat tuberculosis. *J Mol Cell Biol* **2010**, *2* (4), 173-175. DOI: 10.1093/jmcb/mjp053.
(132) Rožman, K.; Alexander, E. M.; Ogorevc, E.; Bozovičar, K.; Sosič, I.; Aldrich, C. C.; Gobec, S. Psoralen Derivatives as Inhibitors of *Mycobacterium tuberculosis* Proteasome. *Molecules* **2020**, *25* (6). DOI: 10.3390/molecules25061305.
(133) Naujokat, C.; Hoffmann, S. Role and function of the 26S proteasome in proliferation and apoptosis. *Lab Invest* **2002**, *82* (8), 965-980. DOI: 10.1097/01.lab.0000022226.23741.37.
(134) Deng, L.; Meng, T.; Chen, L.; Wei, W.; Wang, P. The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal Transduction and Targeted Therapy* **2020**, *5* (1), 11. DOI: 10.1038/s41392-020-0107-0.
(135) Wydorski, P. M.; Osipiuk, J.; Lanham, B. T.; Tesar, C.; Endres, M.; Engle, E.; Jedrzejczak, R.; Mullapudi, V.; Michalska, K.; Fidelis, K.; et al. Dual domain recognition determines SARS-CoV-2 PLpro selectivity for human ISG15 and K48-linked di-ubiquitin. *Nat Commun* **2023**, *14* (1), 2366. DOI: 10.1038/s41467-023-38031-5.
(136) Dzimianski, J. V.; Scholte, F. E. M.; Bergeron, É.; Pegan, S. D. ISG15: It's Complicated. *J Mol Biol* **2019**, *431* (21), 4203-4216. DOI: 10.1016/j.jmb.2019.03.013.
(137) Bonn, S. M.; Fushman, D. Backbone NMR resonance assignment of the intrinsically disordered UBact protein from Nitrospira nitrosa. *Biomolecular NMR Assignments* **2022**, *16* (1), 129-134. DOI: 10.1007/s12104-022-10070-x.
(138) Tompa, P. Intrinsically unstructured proteins. *Trends Biochem Sci* **2002**, *27* (10), 527-533. DOI: https://doi.org/10.1016/S0968-0004(02)02169-2.

(139) LOWRY, O. H.; ROSEBROUGH, N. J.; FARR, A. L.; RANDALL, R. J. Protein measurement with the Folin phenol reagent. *J Biol Chem* **1951**, *193* (1), 265-275.
(140) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **1976**, *72*, 248-254. DOI: 10.1006/abio.1976.9999.
(141) Kjaergaard, M.; Brander, S.; Poulsen, F. M. Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J Biomol NMR* **2011**, *49* (2), 139-149. DOI: 10.1007/s10858-011-9472-x.
(142) Kjaergaard, M.; Poulsen, F. M. Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *J Biomol NMR* **2011**, *50* (2), 157-165. DOI: 10.1007/s10858-011-9508-2.
(143) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **2009**, *44* (4), 213-223. DOI: 10.1007/s10858-009-9333-z.
(144) Majumder, P.; Rudack, T.; Beck, F.; Danev, R.; Pfeifer, G.; Nagy, I.; Baumeister, W. Cryo-EM structures of the archaeal PAN-proteasome reveal an around-the-ring ATPase cycle. *Proc Natl Acad Sci U S A* **2019**, *116* (2), 534-539. DOI: 10.1073/pnas.1817752116.
(145) Guex, N.; Peitsch, M. C.; Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **2009**, *30 Suppl 1*, S162-173. DOI: 10.1002/elps.200900140.
(146) Biasini, M.; Bienert, S.; Waterhouse, A.; Arnold, K.; Studer, G.; Schmidt, T.; Kiefer, F.; Gallo Cassarino, T.; Bertoni, M.; Bordoli, L.; et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **2014**, *42* (Web Server issue), W252-258. DOI: 10.1093/nar/gku340.
(147) Bertoni, M.; Kiefer, F.; Biasini, M.; Bordoli, L.; Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep* **2017**, *7* (1), 10480. DOI: 10.1038/s41598-017-09654-8.
(148) Bienert, S.; Waterhouse, A.; de Beer, T. A.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res* **2017**, *45* (D1), D313-D319. DOI: 10.1093/nar/gkw1132.
(149) Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F. T.; de Beer, T. A. P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **2018**, *46* (W1), W296-W303. DOI: 10.1093/nar/gky427.
(150) Miles, A. J.; Ramalli, S. G.; Wallace, B. A. DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. *Protein Sci* **2022**, *31* (1), 37-46. DOI: 10.1002/pro.4153.
(151) Sreerama, N.; Woody, R. W. A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal Biochem* **1993**, *209* (1), 32-44. DOI: 10.1006/abio.1993.1079.
(152) Sreerama, N.; Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR

methods with an expanded reference set. *Anal Biochem* **2000**, *287* (2), 252-260. DOI: 10.1006/abio.2000.4880.
(153) Drew, E. D.; Janes, R. W. PDBMD2CD: providing predicted protein circular dichroism spectra from multiple molecular dynamics-generated protein structures. *Nucleic Acids Res* **2020**, *48* (W1), W17-W24. DOI: 10.1093/nar/gkaa296.
(154) Vallurupalli, P.; Bouvignies, G.; Kay, L. E. Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. *J Am Chem Soc* **2012**, *134* (19), 8148-8161. DOI: 10.1021/ja3001419.
(155) Fushman, D.; Varadan, R.; Assfalg, M.; Walker, O. Determining domain orientation in macromolecules by using spin-relaxation and residual dipolar coupling measurements. *P Nucl Mag Res Sp* **2004**, (3-4), 189-214. DOI: 10.1016/j.pnmrs.2004.02.001.
(156) Su, X. C.; Jergic, S.; Ozawa, K.; Burns, N. D.; Dixon, N. E.; Otting, G. Measurement of dissociation constants of high-molecular weight protein-protein complexes by transferred 15N-relaxation. *J Biomol NMR* **2007**, *38* (1), 65-72. DOI: 10.1007/s10858-007-9147-9.
(157) Varadan, R.; Assfalg, M.; Haririnia, A.; Raasi, S.; Pickart, C.; Fushman, D. Solution Conformation of Lys^{63}-linked Di-ubiquitin Chain Provides Clues to Functional Diversity of Polyubiquitin Signaling. *J Biol Chem* **2004**, *279* (8), 7055-7063. DOI: 10.1074/jbc.M309184200 (acccessed 2023/08/30).
(158) Battiste, J. L.; Wagner, G. Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with Limited Nuclear Overhauser Effect Data. *Biochemistry* **2000**, *39* (18), 5355-5365. DOI: 10.1021/bi000060h.
(159) Liao, S.; Shang, Q.; Zhang, X.; Zhang, J.; Xu, C.; Tu, X. Pup, a prokaryotic ubiquitin-like protein, is an intrinsically disordered protein. *Biochem J* **2009**, *422* (2), 207-215. DOI: 10.1042/BJ20090738.
(160) Özcelik, D.; Barandun, J.; Schmitz, N.; Sutter, M.; Guth, E.; Damberger, F. F.; Allain, F. H.; Ban, N.; Weber-Ban, E. Structures of Pup ligase PafA and depupylase Dop from the prokaryotic ubiquitin-like modification pathway. *Nat Commun* **2012**, *3*, 1014. DOI: 10.1038/ncomms2009.
(161) Boughton, A. J.; Liu, L.; Lavy, T.; Kleifeld, O.; Fushman, D. A novel recognition site for polyubiquitin and ubiquitin-like signals in an unexpected region of proteasomal subunit Rpn1. *J Biol Chem* **2021**, *297* (3), 101052. DOI: 10.1016/j.jbc.2021.101052.
(162) Jungbauer, A.; Kaar, W.; Schlegl, R. Folding and refolding of proteins in chromatographic beds. *Curr Opin Biotechnol* **2004**, *15* (5), 487-494. DOI: 10.1016/j.copbio.2004.08.009.
(163) Vallejo, L. F.; Rinas, U. Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact* **2004**, *3* (1), 11. DOI: 10.1186/1475-2859-3-11.
(164) Yamaguchi, H.; Miyazaki, M. Refolding techniques for recovering biologically active recombinant proteins from inclusion bodies. *Biomolecules* **2014**, *4* (1), 235-251. DOI: 10.3390/biom4010235.

(165) Batas, B.; Schiraldi, C.; Chaudhuri, J. B. Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography. *J Biotechnol* **1999**, *68* (2-3), 149-158. DOI: 10.1016/s0168-1656(98)00197-7.
(166) Freydell, E. J.; van der Wielen, L. A.; Eppink, M. H.; Ottens, M. Size-exclusion chromatographic protein refolding: fundamentals, modeling and operation. *J Chromatogr A* **2010**, *1217* (49), 7723-7737. DOI: 10.1016/j.chroma.2010.10.038.
(167) Barandun, J.; Delley, C. L.; Ban, N.; Weber-Ban, E. Crystal structure of the complex between prokaryotic ubiquitin-like protein and its ligase PafA. *J Am Chem Soc* **2013**, *135* (18), 6794-6797. DOI: 10.1021/ja4024012.
(168) Chaga, G.; Bochkariov, D. E.; Jokhadze, G. G.; Hopp, J.; Nelson, P. Natural poly-histidine affinity tag for purification of recombinant proteins on cobalt(II)-carboxymethylaspartate crosslinked agarose. *J Chromatogr A* **1999**, *864* (2), 247-256. DOI: 10.1016/s0021-9673(99)01008-0.
(169) Tchaga, G.; Jokhadze, G. Polynucleotides encoding metal ion affinity peptides and related products. United States of America 2001.
(170) Khan, R. H.; Rao, K. B.; Eshwari, A. N.; Totey, S. M.; Panda, A. K. Solubilization of recombinant ovine growth hormone with retention of native-like secondary structure and its refolding from the inclusion bodies of Escherichia coli. *Biotechnol Prog* **1998**, *14* (5), 722-728. DOI: 10.1021/bp980071q.
(171) Singh, S. M.; Sharma, A.; Upadhyay, A. K.; Singh, A.; Garg, L. C.; Panda, A. K. Solubilization of inclusion body proteins using n-propanol and its refolding into bioactive form. *Protein Expr Purif* **2012**, *81* (1), 75-82. DOI: 10.1016/j.pep.2011.09.004.
(172) Gautam, S.; Dubey, P.; Singh, P.; Varadarajan, R.; Gupta, M. N. Simultaneous refolding and purification of recombinant proteins by macro-(affinity ligand) facilitated three-phase partitioning. *Anal Biochem* **2012**, *430* (1), 56-64. DOI: 10.1016/j.ab.2012.07.028.
(173) Raghava, S.; Barua, B.; Singh, P. K.; Das, M.; Madan, L.; Bhattacharyya, S.; Bajaj, K.; Gopal, B.; Varadarajan, R.; Gupta, M. N. Refolding and simultaneous purification by three-phase partitioning of recombinant proteins from inclusion bodies. *Protein Sci* **2008**, *17* (11), 1987-1997. DOI: 10.1110/ps.036939.108.
(174) He, C.; Ohnishi, K. Efficient renaturation of inclusion body proteins denatured by SDS. *Biochem Biophys Res Commun* **2017**, *490* (4), 1250-1253. DOI: 10.1016/j.bbrc.2017.07.003.
(175) Anand, K.; Pal, D.; Hilgenfeld, R. An overview on 2-methyl-2,4-pentanediol in crystallization and in crystals of biological macromolecules. *Acta Crystallogr D Biol Crystallogr* **2002**, *58* (Pt 10 Pt 1), 1722-1728. DOI: 10.1107/s0907444902014610.
(176) Michaux, C.; Pomroy, N. C.; Privé, G. G. Refolding SDS-denatured proteins by the addition of amphipathic cosolvents. *J Mol Biol* **2008**, *375* (5), 1477-1488. DOI: 10.1016/j.jmb.2007.11.026.
(177) Joseph, R. E.; Andreotti, A. H. Bacterial expression and purification of interleukin-2 tyrosine kinase: single step separation of the chaperonin impurity. *Protein Expr Purif* **2008**, *60* (2), 194-197. DOI: 10.1016/j.pep.2008.04.001.
(178) Belval, L.; Marquette, A.; Mestre, P.; Piron, M. C.; Demangeat, G.; Merdinoglu, D.; Chich, J. F. A fast and simple method to eliminate Cpn60 from

functional recombinant proteins produced by E. coli Arctic Express. *Protein Expr Purif* **2015**, *109*, 29-34. DOI: 10.1016/j.pep.2015.01.009.
(179) Baietti, M. F.; Simicek, M.; Abbasi Asbagh, L.; Radaelli, E.; Lievens, S.; Crowther, J.; Steklov, M.; Aushev, V. N.; Martínez García, D.; Tavernier, J.; et al. OTUB1 triggers lung cancer development by inhibiting RAS monoubiquitination. *EMBO Mol Med* **2016**, *8* (3), 288-303. DOI: 10.15252/emmm.201505972.
(180) Lee, B. S.; Kang, S. U.; Huang, M.; Kim, Y. S.; Lee, Y. S.; Park, J. Y.; Kim, C. H. OTUB1 knockdown promotes apoptosis in melanoma cells by upregulating TRAIL expression. *BMB Rep* **2021**, *54* (12), 608-613. DOI: 10.5483/BMBRep.2021.54.12.033.
(181) Zhou, K.; Mai, H.; Zheng, S.; Cai, W.; Yang, X.; Chen, Z.; Zhan, B. OTUB1-mediated deubiquitination of FOXM1 up-regulates ECT-2 to promote tumor progression in renal cell carcinoma. *Cell Biosci* **2020**, *10*, 50. DOI: 10.1186/s13578-020-00408-0.
(182) Borodovsky, A.; Ovaa, H.; Kolli, N.; Gan-Erdene, T.; Wilkinson, K. D.; Ploegh, H. L.; Kessler, B. M. Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family. *Chem Biol* **2002**, *9* (10), 1149-1159. DOI: 10.1016/s1074-5521(02)00248-x.
(183) Edelmann, M. J.; Iphöfer, A.; Akutsu, M.; Altun, M.; di Gleria, K.; Kramer, H. B.; Fiebiger, E.; Dhe-Paganon, S.; Kessler, B. M. Structural basis and specificity of human otubain 1-mediated deubiquitination. *Biochem J* **2009**, *418* (2), 379-390. DOI: 10.1042/BJ20081318.
(184) Goncharov, T.; Niessen, K.; de Almagro, M. C.; Izrael-Tomasevic, A.; Fedorova, A. V.; Varfolomeev, E.; Arnott, D.; Deshayes, K.; Kirkpatrick, D. S.; Vucic, D. OTUB1 modulates c-IAP1 stability to regulate signalling pathways. *EMBO J* **2013**, *32* (8), 1103-1114. DOI: 10.1038/emboj.2013.62.
(185) Zhu, D.; Xu, R.; Huang, X.; Tang, Z.; Tian, Y.; Zhang, J.; Zheng, X. Deubiquitinating enzyme OTUB1 promotes cancer cell immunosuppression via preventing ER-associated degradation of immune checkpoint protein PD-L1. *Cell Death Differ* **2021**, *28* (6), 1773-1789. DOI: 10.1038/s41418-020-00700-z.
(186) Nakada, S.; Tai, I.; Panier, S.; Al-Hakim, A.; Iemura, S.; Juang, Y. C.; O'Donnell, L.; Kumakubo, A.; Munro, M.; Sicheri, F.; et al. Non-canonical inhibition of DNA damage-dependent ubiquitination by OTUB1. *Nature* **2010**, *466* (7309), 941-946. DOI: 10.1038/nature09297.
(187) Juang, Y. C.; Landry, M. C.; Sanches, M.; Vittal, V.; Leung, C. C.; Ceccarelli, D. F.; Mateo, A. R.; Pruneda, J. N.; Mao, D. Y.; Szilard, R. K.; et al. OTUB1 co-opts Lys48-linked ubiquitin recognition to suppress E2 enzyme function. *Mol Cell* **2012**, *45* (3), 384-397. DOI: 10.1016/j.molcel.2012.01.011.
(188) Messick, T. E.; Russell, N. S.; Iwata, A. J.; Sarachan, K. L.; Shiekhattar, R.; Shanks, J. R.; Reyes-Turcu, F. E.; Wilkinson, K. D.; Marmorstein, R. Structural basis for ubiquitin recognition by the Otu1 ovarian tumor domain protein. *J Biol Chem* **2008**, *283* (16), 11038-11049. DOI: 10.1074/jbc.M704398200.
(189) Tan, L.; Shan, H.; Han, C.; Zhang, Z.; Shen, J.; Zhang, X.; Xiang, H.; Lu, K.; Qi, C.; Li, Y.; et al. Discovery of Potent OTUB1/USP8 Dual Inhibitors Targeting Proteostasis in Non-Small-Cell Lung Cancer. *J Med Chem* **2022**, *65* (20), 13645-13659. DOI: 10.1021/acs.jmedchem.2c00408.

(190) Hadari, T.; Warms, J. V.; Rose, I. A.; Hershko, A. A ubiquitin C-terminal isopeptidase that acts on polyubiquitin chains. Role in protein degradation. *J Biol Chem* **1992**, *267* (2), 719-727.
(191) Wilkinson, K. D.; Tashayev, V. L.; O'Connor, L. B.; Larsen, C. N.; Kasperek, E.; Pickart, C. M. Metabolism of the polyubiquitin degradation signal: structure, mechanism, and role of isopeptidase T. *Biochemistry* **1995**, *34* (44), 14535-14546. DOI: 10.1021/bi00044a032.
(192) Reyes-Turcu, F. E.; Shanks, J. R.; Komander, D.; Wilkinson, K. D. Recognition of polyubiquitin isoforms by the multiple ubiquitin binding modules of isopeptidase T. *J Biol Chem* **2008**, *283* (28), 19581-19592. DOI: 10.1074/jbc.M800947200.
(193) Dayal, S.; Sparks, A.; Jacob, J.; Allende-Vega, N.; Lane, D. P.; Saville, M. K. Suppression of the deubiquitinating enzyme USP5 causes the accumulation of unanchored polyubiquitin and the activation of p53. *J Biol Chem* **2009**, *284* (8), 5030-5041. DOI: 10.1074/jbc.M805871200.
(194) Fan, X.; Huang, Q.; Ye, X.; Lin, Y.; Chen, Y.; Lin, X.; Qu, J. Drosophila USP5 controls the activation of apoptosis and the Jun N-terminal kinase pathway during eye development. *PLoS One* **2014**, *9* (3), e92250. DOI: 10.1371/journal.pone.0092250.
(195) Wang, C. H.; Chen, G. C.; Chien, C. T. The deubiquitinase Leon/USP5 regulates ubiquitin homeostasis during Drosophila development. *Biochem Biophys Res Commun* **2014**, *452* (3), 369-375. DOI: 10.1016/j.bbrc.2014.08.069.
(196) Nakajima, S.; Lan, L.; Wei, L.; Hsieh, C. L.; Rapić-Otrin, V.; Yasui, A.; Levine, A. S. Ubiquitin-specific protease 5 is required for the efficient repair of DNA double-strand breaks. *PLoS One* **2014**, *9* (1), e84899. DOI: 10.1371/journal.pone.0084899.
(197) Meng, J.; Ai, X.; Lei, Y.; Zhong, W.; Qian, B.; Qiao, K.; Wang, X.; Zhou, B.; Wang, H.; Huai, L.; et al. USP5 promotes epithelial-mesenchymal transition by stabilizing SLUG in hepatocellular carcinoma. *Theranostics* **2019**, *9* (2), 573-587. DOI: 10.7150/thno.27654.
(198) Pan, J.; Qiao, Y.; Chen, C.; Zang, H.; Zhang, X.; Qi, F.; Chang, C.; Yang, F.; Sun, M.; Lin, S.; et al. USP5 facilitates non-small cell lung cancer progression through stabilization of PD-L1. *Cell Death Dis* **2021**, *12* (11), 1051. DOI: 10.1038/s41419-021-04356-6.
(199) Kaistha, B. P.; Krattenmacher, A.; Fredebohm, J.; Schmidt, H.; Behrens, D.; Widder, M.; Hackert, T.; Strobel, O.; Hoheisel, J. D.; Gress, T. M.; et al. The deubiquitinating enzyme USP5 promotes pancreatic cancer via modulating cell cycle regulators. *Oncotarget* **2017**, *8* (39), 66215-66225. DOI: 10.18632/oncotarget.19882.
(200) Gadotti, V. M.; Caballero, A. G.; Berger, N. D.; Gladding, C. M.; Chen, L.; Pfeifer, T. A.; Zamponi, G. W. Small organic molecule disruptors of Cav3.2 - USP5 interactions reverse inflammatory and neuropathic pain. *Mol Pain* **2015**, *11*, 12. DOI: 10.1186/s12990-015-0011-8.
(201) Verma, R.; Chen, S.; Feldman, R.; Schieltz, D.; Yates, J.; Dohmen, J.; Deshaies, R. J. Proteasomal proteomics: identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. *Mol Biol Cell* **2000**, *11* (10), 3425-3439. DOI: 10.1091/mbc.11.10.3425.

(202) Lam, Y. A.; Xu, W.; DeMartino, G. N.; Cohen, R. E. Editing of ubiquitin conjugates by an isopeptidase in the 26S proteasome. *Nature* **1997**, *385* (6618), 737-740. DOI: 10.1038/385737a0.
(203) Hu, M.; Li, P.; Song, L.; Jeffrey, P. D.; Chenova, T. A.; Wilkinson, K. D.; Cohen, R. E.; Shi, Y. Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14. *EMBO J* **2005**, *24* (21), 3747-3756. DOI: 10.1038/sj.emboj.7600832.
(204) Leggett, D. S.; Hanna, J.; Borodovsky, A.; Crosas, B.; Schmidt, M.; Baker, R. T.; Walz, T.; Ploegh, H.; Finley, D. Multiple associated proteins regulate proteasome structure and function. *Mol Cell* **2002**, *10* (3), 495-507. DOI: 10.1016/s1097-2765(02)00638-x.
(205) Bashore, C.; Dambacher, C. M.; Goodall, E. A.; Matyskiela, M. E.; Lander, G. C.; Martin, A. Ubp6 deubiquitinase controls conformational dynamics and substrate degradation of the 26S proteasome. *Nat Struct Mol Biol* **2015**, *22* (9), 712-719. DOI: 10.1038/nsmb.3075.
(206) Kuo, C.-L.; Goldberg, A. L. Ubiquitinated proteins promote the association of proteasomes with the deubiquitinating enzyme Usp14 and the ubiquitin ligase Ube3c. *P Natl Acad Sci USA* **2017**, *114* (17), E3404-E3413. DOI: 10.1073/pnas.1701734114 (acccessed 2023/08/21).
(207) Zhang, S.; Zou, S.; Yin, D.; Zhao, L.; Finley, D.; Wu, Z.; Mao, Y. USP14-regulated allostery of the human proteasome by time-resolved cryo-EM. *Nature* **2022**, *605* (7910), 567-574. DOI: 10.1038/s41586-022-04671-8.
(208) Peth, A.; Besche, H. C.; Goldberg, A. L. Ubiquitinated proteins activate the proteasome by binding to Usp14/Ubp6, which causes 20S gate opening. *Mol Cell* **2009**, *36* (5), 794-804. DOI: 10.1016/j.molcel.2009.11.015.
(209) Lam, Y. A.; DeMartino, G. N.; Pickart, C. M.; Cohen, R. E. Specificity of the ubiquitin isopeptidase in the PA700 regulatory complex of 26 S proteasomes. *J Biol Chem* **1997**, *272* (45), 28438-28446. DOI: 10.1074/jbc.272.45.28438.
(210) Liu, B.; Zhang, Z.; Hu, Y.; Lu, Y.; Li, D.; Liu, J.; Liao, S.; Hu, M.; Wang, Y.; Zhang, D.; et al. Sustained ER stress promotes hyperglycemia by increasing glucagon action through the deubiquitinating enzyme USP14. *Proc Natl Acad Sci U S A* **2019**, *116* (43), 21732-21738. DOI: 10.1073/pnas.1907288116.
(211) Liu, B.; Jiang, S.; Li, M.; Xiong, X.; Zhu, M.; Li, D.; Zhao, L.; Qian, L.; Zhai, L.; Li, J.; et al. Proteome-wide analysis of USP14 substrates revealed its role in hepatosteatosis via stabilization of FASN. *Nat Commun* **2018**, *9* (1), 4770. DOI: 10.1038/s41467-018-07185-y.
(212) Fu, S.; Zheng, Y.; Sun, Y.; Lai, M.; Qiu, J.; Gui, F.; Zeng, Q.; Liu, F. Suppressing long noncoding RNA OGRU ameliorates diabetic retinopathy by inhibition of oxidative stress and inflammation via miR-320/USP14 axis. *Free Radic Biol Med* **2021**, *169*, 361-381. DOI: 10.1016/j.freeradbiomed.2021.03.016.
(213) Xu, J.; Deng, Y.; Wang, Y.; Sun, X.; Chen, S.; Fu, G. SPAG5-AS1 inhibited autophagy and aggravated apoptosis of podocytes via SPAG5/AKT/mTOR pathway. *Cell Prolif* **2020**, *53* (2), e12738. DOI: 10.1111/cpr.12738.
(214) Graner, E.; Tang, D.; Rossi, S.; Baron, A.; Migita, T.; Weinstein, L. J.; Lechpammer, M.; Huesken, D.; Zimmermann, J.; Signoretti, S.; et al. The

isopeptidase USP2a regulates the stability of fatty acid synthase in prostate cancer. *Cancer Cell* **2004**, *5* (3), 253-261. DOI: 10.1016/s1535-6108(04)00055-8.
(215) Renatus, M.; Parrado, S. G.; D'Arcy, A.; Eidhoff, U.; Gerhartz, B.; Hassiepen, U.; Pierrat, B.; Riedl, R.; Vinzenz, D.; Worpenberg, S.; et al. Structural basis of ubiquitin recognition by the deubiquitinating protease USP2. *Structure* **2006**, *14* (8), 1293-1302. DOI: 10.1016/j.str.2006.06.012.
(216) Bozza, W. P.; Liang, Q.; Gong, P.; Zhuang, Z. Transient kinetic analysis of USP2-catalyzed deubiquitination reveals a conformational rearrangement in the K48-linked diubiquitin substrate. *Biochemistry* **2012**, *51* (50), 10075-10086. DOI: 10.1021/bi3009104.
(217) Varadan, R.; Walker, O.; Pickart, C.; Fushman, D. Structural properties of polyubiquitin chains in solution. *J Mol Biol* **2002**, *324* (4), 637-647. DOI: 10.1016/s0022-2836(02)01198-1.
(218) Oishi, K.; Miyazaki, K.; Kadota, K.; Kikuno, R.; Nagase, T.; Atsumi, G.; Ohkura, N.; Azama, T.; Mesaki, M.; Yukimasa, S.; et al. Genome-wide expression analysis of mouse liver reveals CLOCK-regulated circadian output genes. *J Biol Chem* **2003**, *278* (42), 41519-41527. DOI: 10.1074/jbc.M304564200.
(219) Zhang, R.; Lahens, N. F.; Ballance, H. I.; Hughes, M. E.; Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **2014**, *111* (45), 16219-16224. DOI: 10.1073/pnas.1408886111.
(220) Scoma, H. D.; Humby, M.; Yadav, G.; Zhang, Q.; Fogerty, J.; Besharse, J. C. The de-ubiquitinylating enzyme, USP2, is associated with the circadian clockwork and regulates its sensitivity to light. *PLoS One* **2011**, *6* (9), e25382. DOI: 10.1371/journal.pone.0025382.
(221) Shan, J.; Zhao, W.; Gu, W. Suppression of cancer cell growth by promoting cyclin D1 degradation. *Mol Cell* **2009**, *36* (3), 469-476. DOI: 10.1016/j.molcel.2009.10.018.
(222) Pflug, B. R.; Pecher, S. M.; Brink, A. W.; Nelson, J. B.; Foster, B. A. Increased fatty acid synthase expression and activity during progression of prostate cancer in the TRAMP model. *Prostate* **2003**, *57* (3), 245-254. DOI: 10.1002/pros.10297.
(223) Kuhajda, F. P. Fatty-acid synthase and human cancer: new perspectives on its role in tumor biology. *Nutrition* **2000**, *16* (3), 202-208. DOI: 10.1016/s0899-9007(99)00266-x.
(224) Swinnen, J. V.; Van Veldhoven, P. P.; Timmermans, L.; De Schrijver, E.; Brusselmans, K.; Vanderhoydonc, F.; Van de Sande, T.; Heemers, H.; Heyns, W.; Verhoeven, G. Fatty acid synthase drives the synthesis of phospholipids partitioning into detergent-resistant membrane microdomains. *Biochem Biophys Res Commun* **2003**, *302* (4), 898-903. DOI: 10.1016/s0006-291x(03)00265-1.
(225) Tao, B. B.; He, H.; Shi, X. H.; Wang, C. L.; Li, W. Q.; Li, B.; Dong, Y.; Hu, G. H.; Hou, L. J.; Luo, C.; et al. Up-regulation of USP2a and FASN in gliomas correlates strongly with glioma grade. *J Clin Neurosci* **2013**, *20* (5), 717-720. DOI: 10.1016/j.jocn.2012.03.050.
(226) Jeong, P.; Ha, Y. S.; Yun, S. J.; Yoon, H. Y.; Freeman, M. R.; Kim, J.; Kim, W. J. Assess the expression of ubiquitin specific protease USP2a for bladder cancer diagnosis. *BMC Urol* **2015**, *15*, 80. DOI: 10.1186/s12894-015-0074-x.

(227) Poondla, N.; Chandrasekaran, A. P.; Kim, K. S.; Ramakrishna, S. Deubiquitinating enzymes as cancer biomarkers: new therapeutic opportunities? *BMB Rep* **2019**, *52* (3), 181-189. DOI: 10.5483/BMBRep.2019.52.3.048.
(228) Qu, Q.; Mao, Y.; Xiao, G.; Fei, X.; Wang, J.; Zhang, Y.; Liu, J.; Cheng, G.; Chen, X.; Shen, K. USP2 promotes cell migration and invasion in triple negative breast cancer cell lines. *Tumour Biol* **2015**, *36* (7), 5415-5423. DOI: 10.1007/s13277-015-3207-7.
(229) Priolo, C.; Tang, D.; Brahamandan, M.; Benassi, B.; Sicinska, E.; Ogino, S.; Farsetti, A.; Porrello, A.; Finn, S.; Zimmermann, J.; et al. The isopeptidase USP2a protects human prostate cancer from apoptosis. *Cancer Res* **2006**, *66* (17), 8625-8632. DOI: 10.1158/0008-5472.CAN-06-1374.
(230) Gewies, A.; Grimm, S. UBP41 is a proapoptotic ubiquitin-specific protease. *Cancer Res* **2003**, *63* (3), 682-688.
(231) Sutovsky, P.; Manandhar, G.; McCauley, T. C.; Caamaño, J. N.; Sutovsky, M.; Thompson, W. E.; Day, B. N. Proteasomal interference prevents zona pellucida penetration and fertilization in mammals. *Biol Reprod* **2004**, *71* (5), 1625-1637. DOI: 10.1095/biolreprod.104.032532.
(232) Lee, W.; Tonelli, M.; Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **2015**, *31* (8), 1325-1327. DOI: 10.1093/bioinformatics/btu830.
(233) Wang, T.; Yin, L.; Cooper, E. M.; Lai, M. Y.; Dickey, S.; Pickart, C. M.; Fushman, D.; Wilkinson, K. D.; Cohen, R. E.; Wolberger, C. Evidence for bidentate substrate binding as the basis for the K48 linkage specificity of otubain 1. *J Mol Biol* **2009**, *386* (4), 1011-1023. DOI: 10.1016/j.jmb.2008.12.085.
(234) Reyes-Turcu, F. E.; Horton, J. R.; Mullally, J. E.; Heroux, A.; Cheng, X.; Wilkinson, K. D. The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* **2006**, *124* (6), 1197-1208. DOI: 10.1016/j.cell.2006.02.038.
(235) Amerik AYu; Swaminathan, S.; Krantz, B. A.; Wilkinson, K. D.; Hochstrasser, M. In vivo disassembly of free polyubiquitin chains by yeast Ubp14 modulates rates of protein degradation by the proteasome. *EMBO J* **1997**, *16* (16), 4826-4838. DOI: 10.1093/emboj/16.16.4826.
(236) Falquet, L.; Paquet, N.; Frutiger, S.; Hughes, G. J.; Hoang-Van, K.; Jaton, J. C. A human de-ubiquitinating enzyme with both isopeptidase and peptidase activities in vitro. *FEBS Lett* **1995**, *359* (1), 73-77. DOI: 10.1016/0014-5793(94)01451-6.
(237) Roscoe, B. P.; Thayer, K. M.; Zeldovich, K. B.; Fushman, D.; Bolon, D. N. Analyses of the effects of all ubiquitin point mutants on yeast growth rate. *J Mol Biol* **2013**, *425* (8), 1363-1377. DOI: 10.1016/j.jmb.2013.01.032.
(238) Verma, R.; Peters, N. R.; D'Onofrio, M.; Tochtrop, G. P.; Sakamoto, K. M.; Varadan, R.; Zhang, M.; Coffino, P.; Fushman, D.; Deshaies, R. J.; et al. Ubistatins inhibit proteasome-dependent degradation by binding the ubiquitin chain. *Science* **2004**, *306* (5693), 117-120. DOI: 10.1126/science.1100946.
(239) Nakasone, M. A.; Lewis, T. A.; Walker, O.; Thakur, A.; Mansour, W.; Castañeda, C. A.; Goeckeler-Fried, J. L.; Parlati, F.; Chou, T. F.; Hayat, O.; et al. Structural Basis for the Inhibitory Effects of Ubistatins in the Ubiquitin-Proteasome Pathway. *Structure* **2017**, *25* (12), 1839-1855.e1811. DOI: 10.1016/j.str.2017.10.007.

(240) Horton, D. A.; Bourne, G. T.; Smythe, M. L. Exploring privileged structures: The combinatorial synthesis of cyclic peptides. *Mol Divers* **2000**, *5* (4), 289-304. DOI: 10.1023/A:1021365402751.
(241) Loffet, A. Peptides as Drugs: Is There a Market? *J Pept Sci* **2002**, *8* (1), 1-7. DOI: https://doi.org/10.1002/psc.366.
(242) Yamagishi, Y.; Shoji, I.; Miyagawa, S.; Kawakami, T.; Katoh, T.; Goto, Y.; Suga, H. Natural Product-Like Macrocyclic N-Methyl-Peptide Inhibitors against a Ubiquitin Ligase Uncovered from a Ribosome-Expressed De Novo Library. *Chem Biol* **2011**, *18* (12), 1562-1570. DOI: https://doi.org/10.1016/j.chembiol.2011.09.013.
(243) Wang, L.; Wang, N.; Zhang, W.; Cheng, X.; Yan, Z.; Shao, G.; Wang, X.; Wang, R.; Fu, C. Therapeutic peptides: current applications and future directions. *Signal Transduction and Targeted Therapy* **2022**, *7* (1), 48. DOI: 10.1038/s41392-022-00904-4.
(244) Zhang, H.; Chen, S. Cyclic peptide drugs approved in the last two decades (2001-2021). *RSC Chem Biol* **2022**, *3* (1), 18-31. DOI: 10.1039/d1cb00154j.
(245) Nawatha, M.; Rogers, J. M.; Bonn, S. M.; Livneh, I.; Lemma, B.; Mali, S. M.; Vamisetti, G. B.; Sun, H.; Bercovich, B.; Huang, Y.; et al. De novo macrocyclic peptides that specifically modulate Lys48-linked ubiquitin chains. *Nature Chemistry* **2019**, *11* (7), 644-652. DOI: 10.1038/s41557-019-0278-x.
(246) Vamisetti, G. B.; Saha, A.; Huang, Y. J.; Vanjari, R.; Mann, G.; Gutbrod, J.; Ayoub, N.; Suga, H.; Brik, A. Selective macrocyclic peptide modulators of Lys63-linked ubiquitin chains disrupt DNA damage repair. *Nature Comm* **2022**, *13* (1), 6174. DOI: 10.1038/s41467-022-33808-6.
(247) Rogers, J. M.; Nawatha, M.; Lemma, B.; Vamisetti, G. B.; Livneh, I.; Barash, U.; Vlodavsky, I.; Ciechanover, A.; Fushman, D.; Suga, H.; et al. In vivo modulation of ubiquitin chains by N-methylated non-proteinogenic cyclic peptides. *RSC Chem Biol* **2021**, *2* (2), 513-522, 10.1039/D0CB00179A. DOI: 10.1039/D0CB00179A.
(248) Cook, W. J.; Jeffrey, L. C.; Carson, M.; Chen, Z.; Pickart, C. M. Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). *J Biol Chem* **1992**, *267* (23), 16467-16471.
(249) 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. *Mol Cell* **2005**, *18* (6), 687-698. DOI: 10.1016/j.molcel.2005.05.013.
(250) Perkins, S. J.; Dwek, R. A. Comparisons of ring-current shifts calculated from the crystal structure of egg white lysozyme of hen with the proton nuclear magnetic resonance spectrum of lysozyme in solution. *Biochemistry* **1980**, *19* (2), 245-258. DOI: 10.1021/bi00543a001.
(251) Roscoe, B. P.; Bolon, D. N. Systematic exploration of ubiquitin sequence, E1 activation efficiency, and experimental fitness in yeast. *J Mol Biol* **2014**, *426* (15), 2854-2870. DOI: 10.1016/j.jmb.2014.05.019.
(252) CTech. *This company is fighting against cancer drug resistance*. Calcalist, 2023. (accessed 2023 08/15/2023).
(253) Kay, L.; Xu, G. Y.; Yamazaki, T. Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H2O Saturation. *J Mag Res* **1994**, *109* (1), 129-133.
(254) Schleucher, J.; Sattler, M.; Griesinger, C. Coherence Selection by Gradients without Signal Attenuation: Application to the Three-Dimensional HNCO

Experiment. *Angew Chem Int* **1993**, *32* (10), 1489-1491. DOI: https://doi.org/10.1002/anie.199314891.
(255) Grzesiek, S.; Bax, A. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J Mag Res* **1992**, *96* (2), 432-440. DOI: https://doi.org/10.1016/0022-2364(92)90099-S.
(256) Wittekind, M.; Mueller, L. HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha- and Beta-Carbon Resonances in Proteins. *J Mag Res* **1993**, *101* (2), 201-205. DOI: https://doi.org/10.1006/jmrb.1993.1033.
(257) Clubb, R. T.; Thanabal, V.; Wagner, G. A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue $1^{H}N$, 1^{5N}, and $1^{3C\prime}$ chemical shifts in 1^{5N}-^{13C}-labelled proteins. *J Mag Res* **1992**, *97* (1), 213-217. DOI: https://doi.org/10.1016/0022-2364(92)90252-3.
(258) Muhandiram, D. R.; Kay, L. E. Gradient-Enhanced Triple-Resonance Three-Dimensional NMR Experiments with Improved Sensitivity. *J Mag Res* **1994**, *103* (3), 203-216.
(259) Palmer III, A. G.; Cavanagh, J.; Wright, P. E.; Rance, M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J Magn Reson* **1991**, *93* (1), 151-170.
(260) Kay, L. E.; Keifer, P.; Saarinen, T. Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J Am Chem Soc* **1992**, *114* (26), 10663-10665. DOI: 10.1021/ja00052a088.
(261) Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletzky, O.; Glaser, S. J.; Sørensen, O. W.; Griesinger, C. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J Biomol NMR* **1994**, *4* (2), 301-306. DOI: 10.1007/BF00175254.
(262) Schanda, P.; Brutscher, B. Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds. *J Am Chem Soc* **2005**, *127* (22), 8014-8015. DOI: 10.1021/ja051306e.
(263) Bermel, W.; Bertini, I.; Duma, L.; Felli, I. C.; Emsley, L.; Pierattelli, R.; Vasos, P. R. Complete Assignment of Heteronuclear Protein Resonances by Protonless NMR Spectroscopy. *Angew Chem Int* **2005**, *44* (20), 3089-3092. DOI: https://doi.org/10.1002/anie.200461794 (acccessed 2023/08/09).
(264) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **2005**, *59* (4), 687-696. DOI: 10.1002/prot.20449.
(265) Hall, J. B.; Fushman, D. Characterization of the overall and local dynamics of a protein with intermediate rotational anisotropy: Differentiating between conformational exchange and anisotropic diffusion in the B3 domain of protein G. *J Biomol NMR* **2003**, *27* (3), 261-275. DOI: 10.1023/a:1025467918856.
(266) Grzesiek, S.; Bax, A. The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. *J Am Chem Soc* **1993**, *115* (26), 12593-12594.

(267) Bieri, M.; Gooley, P. R. Automated NMR relaxation dispersion data analysis using NESSY. *BMC Bioinformatics* **2011**, *12* (1), 421. DOI: 10.1186/1471-2105-12-421.
(268) Wiedemann, C.; Bellstedt, P.; Görlach, M. CAPITO--a web server-based analysis and plotting tool for circular dichroism data. *Bioinformatics* **2013**, *29* (14), 1750-1757. DOI: 10.1093/bioinformatics/btt278.
(269) Sreerama, N.; Venyaminov, S. Y.; Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. *Anal Biochem* **2000**, *287* (2), 243-251. DOI: 10.1006/abio.2000.4879.
(270) Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **2000**, *78* (3), 1606-1619. DOI: 10.1016/s0006-3495(00)76713-0 From NLM.
(271) Brown, P. H.; Schuck, P. Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys J* **2006**, *90* (12), 4651-4661. DOI: 10.1529/biophysj.106.081372 From NLM.
(272) Brautigam, C. A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. *Methods Enzymol* **2015**, *562*, 109-133. DOI: 10.1016/bs.mie.2015.05.001 From NLM.
(273) Zhao, H.; Piszczek, G.; Schuck, P. SEDPHAT--a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* **2015**, *76*, 137-148. DOI: 10.1016/j.ymeth.2014.11.012 From NLM.
(274) Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **2005**, *41* (1), 207-234. DOI: 10.1016/j.pep.2005.01.016 From NLM.
(275) Pickart, C. M.; Raasi, S. Controlled synthesis of polyubiquitin chains. *Methods Enzymol* **2005**, *399*, 21-36. DOI: 10.1016/s0076-6879(05)99002-2 From NLM.