INVESTIGATIONS INTO NONCANONICAL UBIQUITINATION

by

Kedar Puvar

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Chemistry

West Lafayette, Indiana

May 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Chittaranjan Das, Chair

Department of Chemistry

Dr. Christine Hrycyna

Department of Chemistry

Dr. Daniel Flaherty

Department of Medicinal Chemistry and Molecular Pharmacology

Dr. Angeline Lyon

Department of Chemistry

Approved by:

Dr. Christine Hrycyna

This work is dedicated to Dulesinh Puvar, 1921-2011.

ACKNOWLEDGMENTS

I would first like to acknowledge my advisor Dr. Chittaranjan Das for his mentorship, guidance, and especially his patience over the past five years. He has been a seemingly endless source of knowledge and inspiration throughout my graduate career and I sincerely thank him for that. Also, I would like to thank the members of my committee: Dr. Christine Hrycyna, Dr. Daniel Flaherty, and Dr. Angeline Lyon for their valuable comments and assistance during my time here.

I don't think I would be in this program if it weren't for Dr. Bob Clark and Dr. Laura Rowe, two great mentors who really inspired me to pursue a graduate career.

I also acknowledge my lab mates for their contributions both scientific and non-scientific, whether it be a word of encouragement, a suggestion for an experiment, or a warm home-cooked meal. You all have helped make my time in the lab enjoyable. Thank you especially to Dr. Shalini Iyer and Aya Saleh - it has been fun to work on some great projects together, and I'm going to miss the thoughtful discussions, headaches, stumbles, and harebrained ideas we somehow managed to make work. Thank you to Kristos Negrón Terón, Sebastian Kenny, Zhengrui Zhang, it has been surreal watching you all grow as scientists and I look forward to hearing about your latest discoveries as well as antics. I also would like to acknowledge the newest members of the Das Lab - Gil Gonzalez and Rishi Patel. Get ready for a long but rewarding journey and never stop asking questions. Of course I must acknowledge the undergraduates I had the fortune of working with: Bharathi Ravikumar, Ryan Westman, Rachel Hohe (Good luck with your Ph.D!), Charan Kanthala, and Hannah Rondon. To the past members of the Das Lab: I especially thank Dr. Michael Sheedlo for showing me the ropes early on, as well as Dr. Rashmi Shrestha, Aditya "Raj" Babar, Cameron Wade, John Hausman, and Kwame Brown for all of their valuable help and encouragement throughout.

A big thank you to all the great collaborators I've had the chance to work with, especially Dr. Zhao-Qing Luo, Dr. Peter Brzovic, Dr. Rachel Klevit, Ryan Curtis, Dr. Jean Chmielewski, Dr Yiyang Zhou, Dr. Mary Wirth, and the tireless beamline staff at the Advanced Photon Source. Also a thank you to Dr. Pat Bishop and Dr. Connie Bonham who have been excellent resources on all things instrumentation, and to Laura Mendoza for her great service and the great laughs.

I'm grateful for all of the lasting friendships I've made throughout my time at Purdue as well as before that; you all have played a huge role in helping me make it through graduate school

in one piece. A special shoutout to my Phi Sigma Kappa brothers who have not only given me so much during my undergraduate career but who also gave me the opportunity to return the favor as Advisor during my graduate career.

I would like to acknowledge my parents, Dhiren and Jayshree Puvar, for their unconditional love and support throughout my life and for the foundation they've given me to make all of this possible. I'm always striving to do better thanks to them. Also my sister, Dhara, for her never-ending positivity and can-do spirit that always tends to find its way to me, no matter how grumpy of a mood I might be in. And lastly, Lisa Borosh. Five years across state lines has been far too long. I can hardly wait to come home to you, my dear.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The modification of proteins by the covalent attachment of ubiquitin is a natural process that crucially regulates a wide range of eukaryotic signaling outcomes. This process has been understood as the linking of the C-terminus of ubiquitin to the lysine residue of a target protein via an isopeptide linkage, catalyzed by the coordinated effort by E1, E2, and E3 enzymes. Importantly, ubiquitination has only been observed to be a eukaryotic phenomenon. In recent years though, intracellular bacteria, including human pathogens, have been observed to possess ubiquitin-interacting proteins in their genomes. These proteins serve to subdue and manipulate their hosts' ubiquitin signaling for their own benefit. While some of these proteins act within the eukaryotic context, more recent findings reveal the existence of prokaryotic enzymes that catalyze ubiquitination using mechanisms never before seen in nature. These remarkable processes utilize different cofactors and target different amino acid residues of both ubiquitin as well as substrate protein. The findings reported in this Thesis involve structural and biochemical studies on two new ubiquitinating proteins, the only two proteins known to catalyze ubiquitination outside of the canonical pathway. Both proteins are present in the genome of the intracellular human pathogen *Legionella pneumophila*: the SidE family, which catalyzes ubiquitination via a mechanism combining ADP-ribosylation and phosphodiesterase activities, and MavC, which utilizes a mechanism reminiscent of transglutaminases. Key insights provided in this document include the discovery that SidE enzymes can modify multiple ubiquitin moieties within a ubiquitin chain, and that modified ubiquitin chains are resistant to hydrolytic cleavage from many deubiquitinating enyzmes. Also, the development of a robust, continuous assay for SidE-catalyzed ubiquitination using a synthetic substrate is described. The catalytic action of MavC, which differs from both canonical E1/E2/E3 ubiquitination and SidE ubiquitination is also here elucidated. The crystal structure of MavC in complex with its ubiquitinated product is presented and provides an atomic view into the basis of substrate recognition. These findings bring to light a new dimension of host-pathogen interactions, where pathogenic ubiquitinating enzymes have appeared to arise from convergent evolution. The regulation of these pathogenic enzymes by other effectors is also discussed, as well as biochemical studies of these regulators. Further, these findings describe possible new drug discovery strategies, as well as possible techniques for discovering similar enzymes in organisms besides *Legionella*.

CHAPTER 1. INTRODUCTION

1.1 Ubiquitin Signaling

The small, eukaryotic 76-amino acid protein known as ubiquitin plays a major role in normal cellular function.[1] Ubiquitin signaling begins with the covalent linkage of the C-terminus of ubiquitin to a lysine residue of a target protein, forming an isopeptide bond. Once ubiquitinated, the target protein may undergo further ubiquitination, including on the initial ubiquitin moiety itself. The latter process results in the formation of a ubiquitin chain.[2] Ubiquitin chains can be constructed by way of any of the seven lysine residues of ubiquitin as well as the amino group of the N-terminal methionine. Differently linked chains may possess substantially different shapes and topologies (Figure 1-1).

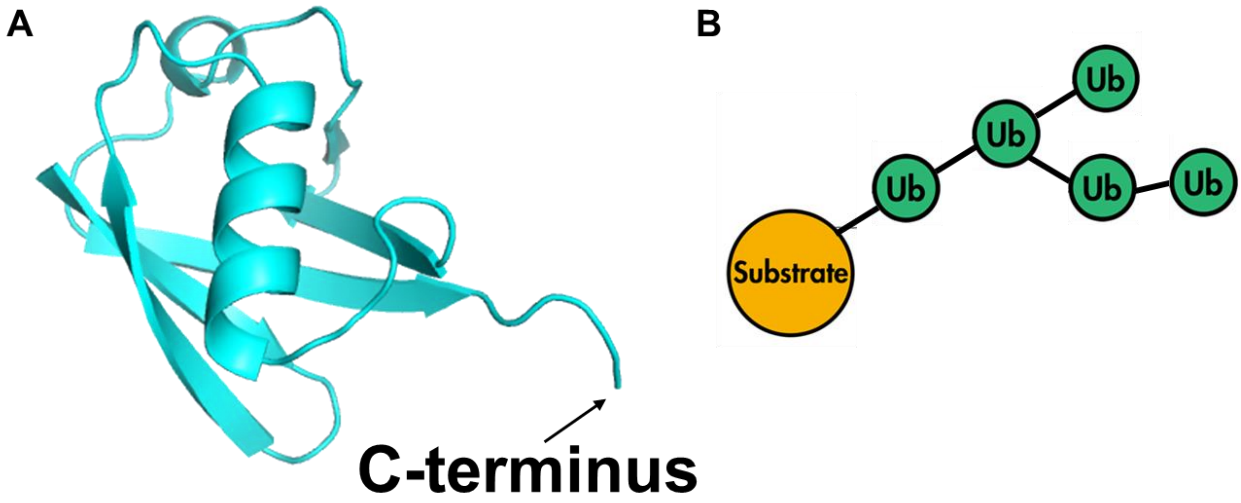


Figure 1-1: Overall structure of ubiquitin and ubiquitin chains. A. Cartoon depiction of a ubiquitin protein. The C-terminus, via which ubiquitin is attached onto substrate proteins, is highlighted. B. Depiction of polyubiquitin chains. This figure shows the presence of mixed linkages.

The ubiquitination of a protein can command a variety of signaling consequences.[3] Among the most widely cited are proteasomal degradation of the protein, endosomal trafficking, autophagy, the DNA damage response, as well as the immune response making ubiquitin signaling indispensable for normal cellular function. The type of ubiquitin chain formed dictates the

signaling pathway. For example, ubiquitination by chains linked via Lys48 of ubiquitin correlates with proteasomal degradation of the substrate protein, as the proteasome's ubiquitin receptors preferentially bind to the topology of Lys48 polyubiquitin chains.[4,5] Lys63-linked chains are well established to play roles in processes such as the trafficking of membrane-associated proteins to the lysosome where they are degraded.[6] In this case, protein complexes important in this trafficking process such as ESCRT bind to Lys63-linked chains, thereby targeting themselves to the appropriate protein substrate.[7] There also exist examples of monoubiquitin-specific signaling outcomes, such as PCNA, which becomes monoubiquitinated as a response to DNA damage.[8] This monoubiquitinated PCNA is recognized by DNA polymerases that are able to carry out repair. Furthermore, there is growing evidence of the presence of mixed-linkage as well as branched chains in cells, adding new layers of complexity to the ubiquitin code.[3,9]

While the isopeptide linkage between ubiquitin and substrate is stable under biological conditions, ubiquitination is a reversible process, where the amide bond may be hydrolyzed by deubiquitinating enzymes, or DUBs.[10,11] There exist over 100 different DUBs in human cells comprising at least seven known families.[12] Six of these families (USP, UCH, OTU, MJD, MINDY, ZUFSP) utilize a cysteine protease mechanism, while one family (JAMM) uses a zinc metalloprotease mechanism. DUBs play important roles in regulating ubiquitination; for example in the context of proteasomal degradation they remove ubiquitin from the substrate protein before it is degraded, allowing it to be reused by the cell. In lysosomal trafficking, the cargo also undergoes deubiquitination prior to lysosomal transport. In the DNA damage response, the deubiquitination of PCNA effectively switches off the damage signal as the repair polymerases no longer associate after ubiquitin is removed.[13] Consistent with the wide variety of Ub chain types, different DUBs recognize different chain topologies which therefore determines their involvement in specific signaling pathways.[14]

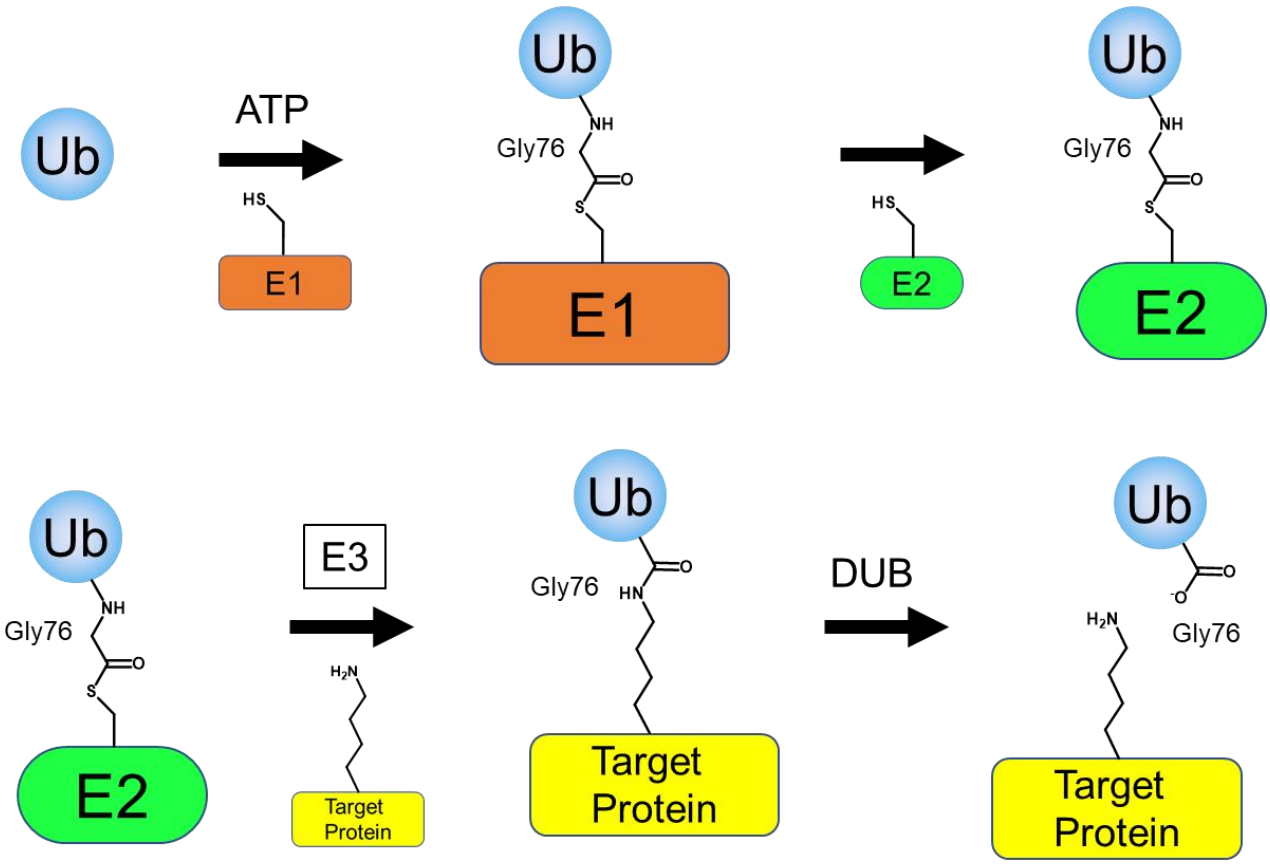


Figure 1-2: Cartoon depicting the canonical E1/E2/E3 ubiquitination mechanism, along with deubiquitination.

1.2 The Canonical Ubiquitination Process

Ubiquitination is accomplished through the combined efforts of three enzymes: E1, or ubiquitin activating enzyme, E2, or ubiquitin conjugating enzyme and E3, or ubiquitin ligase (Figure 1-2). This process begins by E1 adenylating the C-terminus of ubiquitin, requiring the turnover of ATP by release of pyrophosphate. This adenylated, or "activated" ubiquitin undergoes nucleophilic attack by a cysteine residue of E1, generating a thioester linkage between E1 and ubiquitin (E1~Ub).[1] The ubiquitin is then passed from E1~Ub to the catalytic cysteine of an E2 enzyme. This passing of ubiquitin occurs via a transthiolation reaction and generates a new thioester linkage, now between E2 and ubiquitin (E2~Ub). This "charged" E2 is assisted by an E3 to ubiquitinate the substrate protein. There are three known E3 families which operate through different mechanisms: RING (Really Interesting New Gene), HECT (Homologous to the E6-AP Carboxyl Terminus), and RBR (RING-between-RING). RING type ligases function by binding to

both the E2~Ub conjugate as well as the substrate protein, bringing them within close enough proximity for Ub to be transferred to the substrate's lysine residue.[15] In contrast, HECT-type E3 ligases act via a two-step mechanism. They possess a catalytic cysteine residue which serves to first take ubiquitin from the charged E2 in another transthiolation reaction, which generates an E3~Ub species.[16–18] This charged E3 then transfers ubiquitin to a substrate protein. RBR-type ligases combine aspects of both RING and HECT-type E3s in their mechanism and are both structurally distinct and uniquely self-regulating.[19–21]

In humans, only two E1 enzymes exist: UBA1 and UBA6. Meanwhile, approximately 40 E2 enzymes are encoded in the genome, and approximately 600 E3s have been identified.[2] It is evident that the great variety of E3s is required due to the great variety of ubiquitination substrates; nearly all types of proteins in the cell undergo ubiquitination during their lifespans. Each E2 is recognized by a subset of E3s and in turn E3s also can form complexes with different and interchangeable substrate adaptor domains, further allowing for tight control of ubiquitin signaling by the cell. Different E2s and E3s, besides recognizing different substrate proteins, are also known to generate different types of ubiquitin chains. For example, the HECT ligase E6AP generates Lys48-linked chains on different protein substrates,[17] while the heterodimeric E2 pair Ube2N and Uev1A (respectively named Ubc13 and Mms2 in yeast) form a complex that can either elongate Lys63-linked chains or also associate with various E3s to ubiquitinate substrate proteins.[22] The structures of these proteins in complex with ubiquitin have revealed wide differences in ubiquitin binding interfaces, which provide a basis for the observed diversity in Ub chain types generated at both the E2 and E3 level.[9]

1.3 Pathogenic Interference in Host Ubiquitin Signaling

In certain bacteria, some systems exist that use small protein modifiers such as pupylation, where a small, intrinsically disordered protein called Pup is conjugated onto lysine residues of proteins and plays a role in protein degradation.[23] Also, a ubiquitin-like small protein dubbed UBact was also discovered in certain gram-negative bacteria.[24] However, as a rule, ubiquitin is only found in the genomes of eukaryotic species. The evolution of the ubiquitin system is therefore likely to have been an ancient event, and it is believed that the appearance of these enzymes in the genome represents a major step in the transition from prokaryotes to eukaryotes.[25–28]

Despite lacking ubiquitin systems of their own and not utilizing ubiquitin signaling for their own internal cellular processes, an emerging area of interest has been the study of bacterial proteins that interact with ubiquitin and more broadly the ubiquitin signaling machinery of eukaryotes.[29] This is typically understood in the context of host-pathogen interactions; due to the critical role that the ubiquitin system plays in regulating eukaryotic processes, it is ripe for manipulation by pathogens, particularly intracellular pathogens. These are bacteria that carry out their life cycle by becoming phagocytosed and internalized by eukaryotic cells which serve as their host (Figure 1-3). The interior of a cell is a hostile environment for bacteria, as many mechanisms exist to destroy and digest the invading pathogen. In order to avoid the host cell's innate defense mechanisms, these intracellular pathogens employ effectors, which are proteins secreted into the host cell that allow the bacteria to control host processes to ensure the survival of the bacteria. These effectors have a myriad of different functions, but of particular interest are those effectors that target host ubiquitin signaling. Indeed, there is a growing number of already abundant examples of ubiquitin-interacting proteins in the genomes of prokaryotic bacteria.[30]

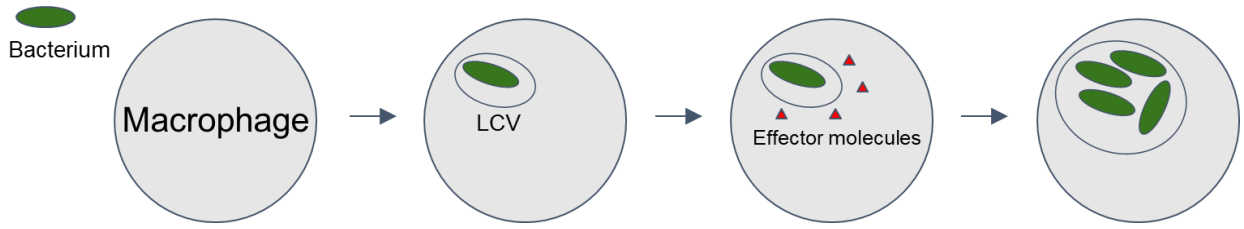


Figure 1-3: Scheme of the life cycle of intracellular bacteria such as *Legionella*. Upon endocytosis by the host cell, effectors are released into the cytoplasm, allowing the pathogen to successfully replicate and carry out its life cycle.

Many of these have been shown to act within the classically known E1/E2/E3 system, with some examples including bacterial E3 ligases like LubX from *Legionella pneumophila* which operates through a RING-like mechanism and causes the ubiquitination and degradation of another effector called SidH, likely to regulate its activity.[31] SopA from *Salmonella enterica*[32] and NleL from *Escherichia coli*, which operate through a HECT-like mechanism to respectively ubiquitinate human TRIM56/TRIM65 E3 ligases,[33] or c-Jun NH2-terminal kinases, with both effectors' activities shown to cause a diminishing of the host's immune response. Bacterial E3 ligases have also been reported to function by promoting proteasomal degradation of proteins for the purpose

of increasing the concentration of free amino acids to promote bacterial growth.[34–36] Also of note is SidC and its paralog SdcA, also expressed by *Legionella pneumophila* and released into the host cell during infection.[37,38] These effectors are E3 ligases that catalyze ubiquitination differently from the RING, HECT, or RBR families. They contain a Cys-His-Asp catalytic triad reminiscent of cysteine proteases and DUBs; however, initial studies did not observe DUB activity; instead, they acted as ligases, associating preferentially with the human E2s UbcH7 (SidC) or UbcH5C (SdcA) to generate polyubiquitin chains.[37] A subsequent study reporting a structural analysis of these effectors in complex with both E2 and E2~Ub conjugates revealed a dynamic conformational motion of the enzyme to accept the Ub in a transthiolation reaction before transfer to the protein substrate.[38] Another example of a novel class of E3 ligases is the IpaH family which is conserved among many bacterial species but structurally unrelated to any known human E3s.[39]

Along with E3 ligases, several pathogenic DUBs have also been described and characterized.[40] Many DUBs have been found to be encoded in viral genomes. Strikingly, some OTU domain-containing DUBs were identified at an early point in viruses via bioinformatics-based analyses and further work found that these DUBs can serve useful functions for viruses, such as immune response modulation.[41] The coronavirus family notably encodes a papain-like protease (PLpro) able to hydrolyze ubiquitin and the ubiquitin-like protein ISG15 from substrate proteins, with examples characterized from both SARS and MERS viruses.[42,43] The novel SARS-CoV-2 possesses a papain-like protease also with predicted ubiquitin-interacting ability but which has yet to be experimentally confirmed.[44]

A wide variety of bacterial DUBs have also been identified. Some noteworthy examples include the effector SseL from *Salmonella*, ChlaDUB1 and ChlaDUB2 from *Chlamydia*, RickCE from *Rickettsia*, ElaD from *E. coli*, the N-terminal domain of the SidE family in *Legionella*,[45] and ShiCE from *Shigella*.[46,47] These proteins are members of the CE clan, a group of proteases that in humans specifically recognize and hydrolyze linkages involving ubiquitin-like proteins such as NEDD8 or SUMO rather than ubiquitin.[48] OTU-domain containing DUBs have also been identified in bacterial species, with at least three having been characterized in *Legionella pneumophila* alone: LotA, RavD, and Ceg23.[49–51] These DUBs have varying linkage specificities, which likely suggests a diversity of roles for DUBs in bacterial pathogenesis.[52]

Along with E3s and DUBs, bacteria have been more recently found to interfere with host ubiquitin signaling via mechanisms never before seen in nature. One example is that of the Cif

family of proteins, which catalyzes deamidation of ubiquitin and the structurally related small protein modifier NEDD8 at the Glu40 position.[53] This has the profound consequence of deamidated ubiquitin being deficient in synthesis of ubiquitin chains, and deamidated NEDD8, an important component in some RING E3 ligase complexes, effectively switching off the ubiquitinating ability of these complexes. The latter activity has been shown to cause macrophage-specific apoptosis of the host cell. Another example of direct covalent attack of the host ubiquitinating machinery is that of the *Shigella* effector OspI.[54,55] This protein was remarkably found to catalyze deamidation of Glu100 of the ubiquitin-conjugating E2 enzyme Ube2N. This deamidation leads to an impairment in ubiquitin-conjugating ability, preventing Ube2N from effectively activating the E3 ligase TRAF6, which plays a role in immune signaling.

1.4 Alternative Ubiquitination Mechanisms: SidE

An unprecedented finding, first reported in 2016 by Qiu and colleagues, has been the discovery of enzymes with the ability to catalyze covalent ubiquitin attachment onto a substrate protein in the absence of E1 or E2.[56] These enzymes consisted of a family of effectors called SidE, found in the genomes of several *Legionella* species. In *Legionella pneumophila*, the SidE family contains four well-conserved members: SdeA, SdeB, SdeC, and SidE. These proteins are relatively large (with SdeA being 1499 amino acids in length) and are secreted by Legionella into the host cell's cytoplasm after the bacterium is phagocytosed. This process is referred to as Type IV secretion and involves the Dot/Icm complex, which mediates translocation of over 300 different protein substrates in *L. pneumophila*.[57] This large arsenal of protein effectors is required, and deletion of the Dot/Icm system results in an inability of bacteria to grow. However, the deletion of individual effectors generally has no effect, with some exceptions, indicating a high degree of functional redundancy among *Legionella*'s molecular arsenal. The importance of the SidE enzymes in virulence had been known for over a decade prior, where a study showed that deletion of the SidE family resulted in a noticeable defect in *Legionella* virulence against the model host organism *Acanthamoeba castellanii*.[58] Subsequent investigations demonstrated the SidE family as being toxic to yeast and mammalian cells when expressed within them, with another effector, SidJ, appearing to regulate SidE-mediated toxicity.[59,60] However, the molecular mechanism of this toxicity remained elusive. Qiu et al. conducted a bioinformatic analysis revealing a conserved domain within the SidE family that resembled that of known bacterial mono-ADP-

ribosyltransferase (mART) enzymes (Figure 1-4).[56] The key motif was an R-S-ExE, a specific hallmark of arginine mARTs. ADP-ribosylation is defined as the transfer of the ADP-ribose (ADPR) group of nicotinamide adenine dinuclotide (NAD^+) onto an amino acid side chain of a protein.[61] Several examples of bacterial mARTs exist in the literature, with IotA toxin and C3 exoenzyme from *Clostridium* species as well as ExoS from *Pseudomonas aeruginosa* all sharing the R-S-ExE motif with the SidE family.[62] These enzymes all acted via the key glutamate residues in this motif, which in prior studies have been shown to be essential for transfer of the ADPR group onto a target arginine of host proteins.[63] Indeed, SdeA mutants where these glutamates were converted into alanine residues became nontoxic to yeast. This indicated that the catalytic activity of SdeA was required for its toxic effect. However, no ADP-ribosylated proteins were found to be generated by this enzyme.[56] Furthermore, some Rab GTPases, common post-translational modification targets of *Legionella* effectors, were found to undergo a shift in molecular weight when SdeA was expressed in mammalian cells which was dependent on the mART motif. This shift was found via mass spectrometry to be a result of ubiquitin linkage. Attempts to conduct this reaction *in vitro* utilizing purified E1s and E2 failed, but interestingly succeeded upon replacing E1/E2/ATP with cell lysate. Moreover, boiled cell lysate also allowed SdeA to catalyze ubiquitination of Rab proteins. Since most proteins were inactivated by the boiling step, this became first reported example of ubiquitination occurring outside of the E1/E2/E3 pathway that has been canonically studied. This study also identified NAD^+ to be a required cofactor as well as Arg42 of ubiquitin as a target site. An mART motif-containing construct of SdeA, $SdeA^{519\text{-}1100}$, was found to not catalyze ubiquitination. When reacted with NAD^+ and ubiquitin, this construct readily formed ADP-ribosylated ubiquitin at Arg42 (Ub-ADPR), indicating that this was likely a reaction intermediate.[56]

Further studies identified the N-terminal region of the domain, (not present in $SdeA^{519\text{-}1100}$) as a phosphodiesterase (PDE) domain that was ultimately responsible for catalyzing the attack of a serine residue of a protein substrate on the β-phosphate to generate a phosphoribose (PR) linkage between Arg42 of ubiquitin and a Ser residue of the target protein.[64,65] A byproduct of this reaction is the generation of phosphoribosylated ubiquitin (Ub-PR), which can be explained as the result of a phosphotransfer to water. *In vitro*, SidE enzymes in conjunction with NAD^+ and protein substrates were found to robustly convert ubiquitin into either PR-linked conjugates or Ub-PR, with the Ub-ADPR species being consumed as an intermediate, and AMP being released in the

phosphodiesterase step. Thus, two separate domains were defined, the mART domain, approximately located between residues 592 and 911 of SdeA, and the PDE domain, approximately located between residues 220 and 590 of SdeA. It is through the coordinated effort of these two domains that ubiquitination of protein substrates occurs (Figure 1-5).

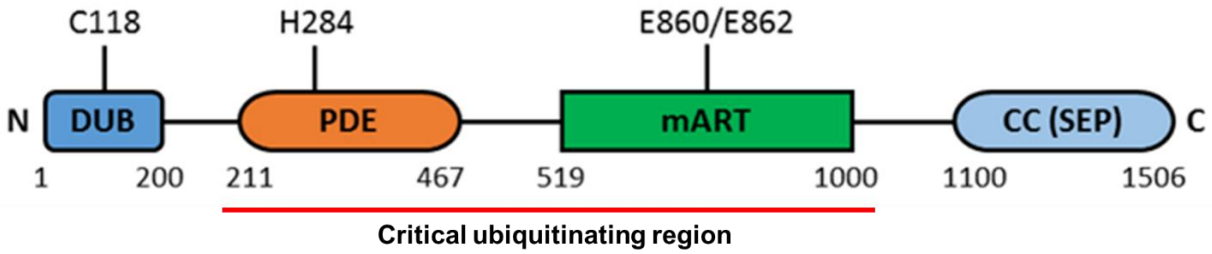


Figure 1-4: Domain organization of SidE enzymes. In this case the domains of SdeA are defined, along with key residues.

The initial findings into this novel ubiquitination reaction spurred five independent investigations that determined in some fashion the molecular structure of SidE family proteins.[66–71] The modular nature of these proteins was highlighted, with a clear domain organization between the PDE and mART domains. Structures of both the PDE and mART domains in complex with ubiquitin identified the basis of substrate recognition. In particular, the cocrystal structure of mART-ubiquitin found that Arg72, instead of Arg42 (mutated in this instance into Ala to prevent reaction during crystallization) of ubiquitin was engaging with NAD^+ and the catalytic Glu residues of SdeA.[66,68] This may represent a state before actual catalysis, where the mART domain initially engages with Arg72 before bringing the nucleophilic Arg42 near the ribose moiety for attack. The actual mode of Ub interaction during ADP-ribosylation by SidE remains to be fully determined.

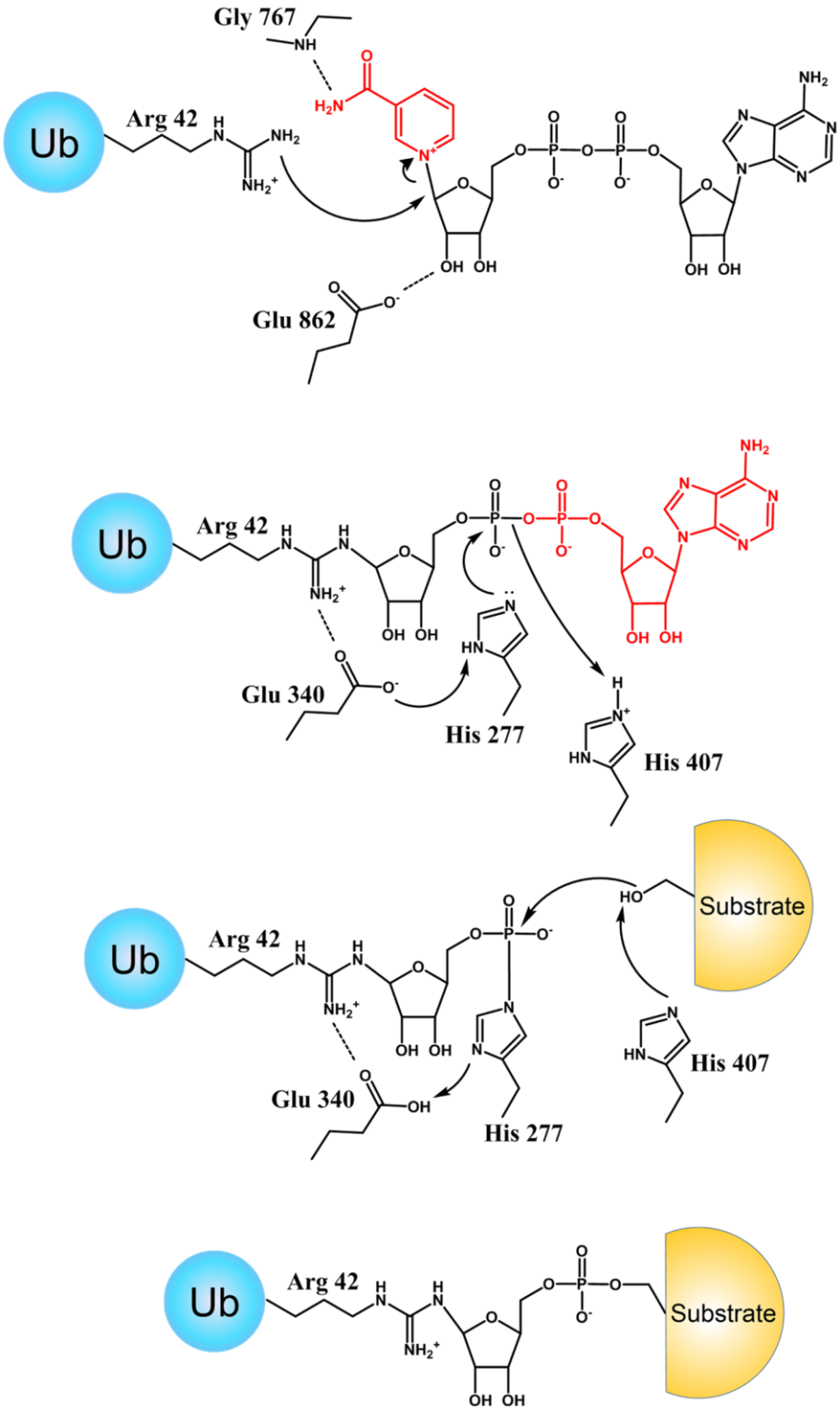


Figure 1-5: Mechanism of SidE-catalyzed ubiquitination. The first step, ADP-ribosylation of Ub at Arg 42, is performed by the mART domain while the second step, phosphotransfer to substrate proteins, is performed by the PDE domain. Key catalytic residues are shown here.

After the mART domain comes a coiled-coiled region (referred to as the CC domain) that had been predicted by bioinformatic analysis (Figure 1-4). This region was found to be important for full mART activity, as well as a possible interface for SidEs to form dimers.[71] While this was an enzyme family with clear distinctions between domains, there were significant inter-domain contacts observed between PDE and mART, as well as mART and CC, that were ultimately required for optimal activity. Interestingly, crystal structures of the mART domain revealed an alpha-helical lobe (AHL) of the mART domain that existed in a different orientation when the CC domain was not included in the protein construct crystallized (Figure 1-6). This suggests that the CC domain's role may also be to stabilize the AHL in a productive position. One study also determined a crystal structure of SdeA with the CC domain bound to two Ub. It is unclear whether this binding is biologically or mechanistically relevant and will require further work to explain.

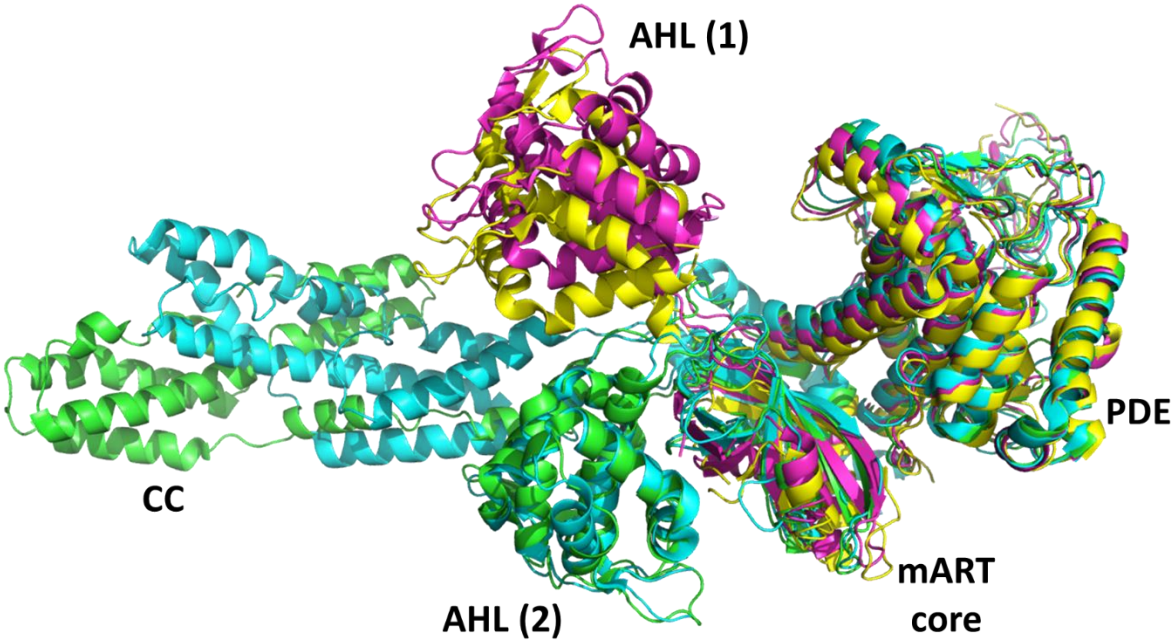


Figure 1-6: Structures of SidE enzyme constructs, with a structural alignment revealing that the AHL region of the mART domain in a different orientation (1) in constructs lacking the CC domain (colored yellow and magenta, PDB codes 6B7Q and 6G0C respectively) compared to constructs containing the CC domain (colored green and cyan, PDB codes 5ZQ2 and 5YIM respectively).

Some insight was also gained into to how SidE enzymes recognize their protein substrates, as it was determined that flexible, unstructured regions of proteins containing Ser residues may be ubiquitinated.[66,69] One study added a flexible tag to non-substrates which resulted in ubiquitination occurring.[66] However, this does not rule out the possibility of a specific fold or folds also being recognized for SidE-catalyzed ubiquitination. Indeed, identification of the most robust protein target as well as the consequence of its PR-linked ubiquitination will be necessary to determine the basis of the importance of SidE enzymes on *Legionella* virulence.

1.5 Alternative Ubiquitination Mechanisms: MavC

Surprisingly, it was revealed that a third mechanism of E1/E2/E3-independent ubiquitination existed, also catalyzed by a *Legionella* pneumophila effector.[72] Gan and colleagues attempted to probe for noncanonical ubiquitination of substrate proteins by expressing an HA-tagged variant of ubiquitin without the C-terminal Gly residues (ΔGG). It was determined that this construct of ubiquitin was detected as part of a higher molecular weight conjugate with Ube2N. a known ubiquitin-conjugating enzyme. This effect was found in cells infected with *L. pneumophila* but not in cells infected with *L. pneumophila* with a defective Dot/Icm system, suggesting that a Dot/Icm effector protein was likely able to catalyze noncanonical ubiquitination of Ube2N (Figure 1-7). The enzyme responsible for this process was then determined to be Lpg2147, hereafter referred to as MavC. This process was determined to require no nucleotide cofactor and mass spectrometric analysis identified the linkage between ubiquitin and Ube2N to have formed between Gln40 of ubiquitin and Lys 92 or Lys 94 of Ube2N (although mutation of Lys94 to alanine did not significantly affect ubiquitination *in vitro*, suggesting Lys92 was the major target residue). See Table 1 for a comparison of these three ubiquitination mechanisms.

Table 1-1: Characteristics of E1/E2/E3, SidE, and MavC-based ubiquitination pathways.

	E1/E2/E3	**SidE**	**MavC**
Cofactor	ATP	NAD^+	None
Ub residue	Gly76	Arg42	Gln40
Target residue	Lys	Ser	Lys
Linker	Isopeptide	Phospho-ribose	Isopeptide

A clear biological purpose was also established in this study: the ubiquitinated Ube2N (Ub-Ube2N) was found to no longer able to perform its catalytic function of transferring ubiquitin to its substrates. Ube2N's role is well known as an E2 that builds K63 polyubiquitin chains on substrate proteins such as NEMO, which is an important step in the activation of the NF-κB signaling pathway.[73] This pathway ultimately activates genes important for the immune response and is a reaction to stimuli such as pathogenic attack. The switching off of Ube2N via noncanonical ubiquitination constitutes an innovative method for *Legionella* to block immune signaling activity of its host cell, allowing the pathogen to remain undisturbed during infection. The mechanism is reminiscent of transglutaminases, enzymes that also crosslink Gln to Lys.[74,75] However, unlike most transglutaminases in the literature, MavC had a clear substrate selectivity, recognizing only ubiquitin (over other ubiquitin-like proteins such as NEDD8), as well as only Ube2N (over other structurally well-conserved E2 enzymes, even those that also contained a Lys residue in a similar location of Lys92 of Ube2N).

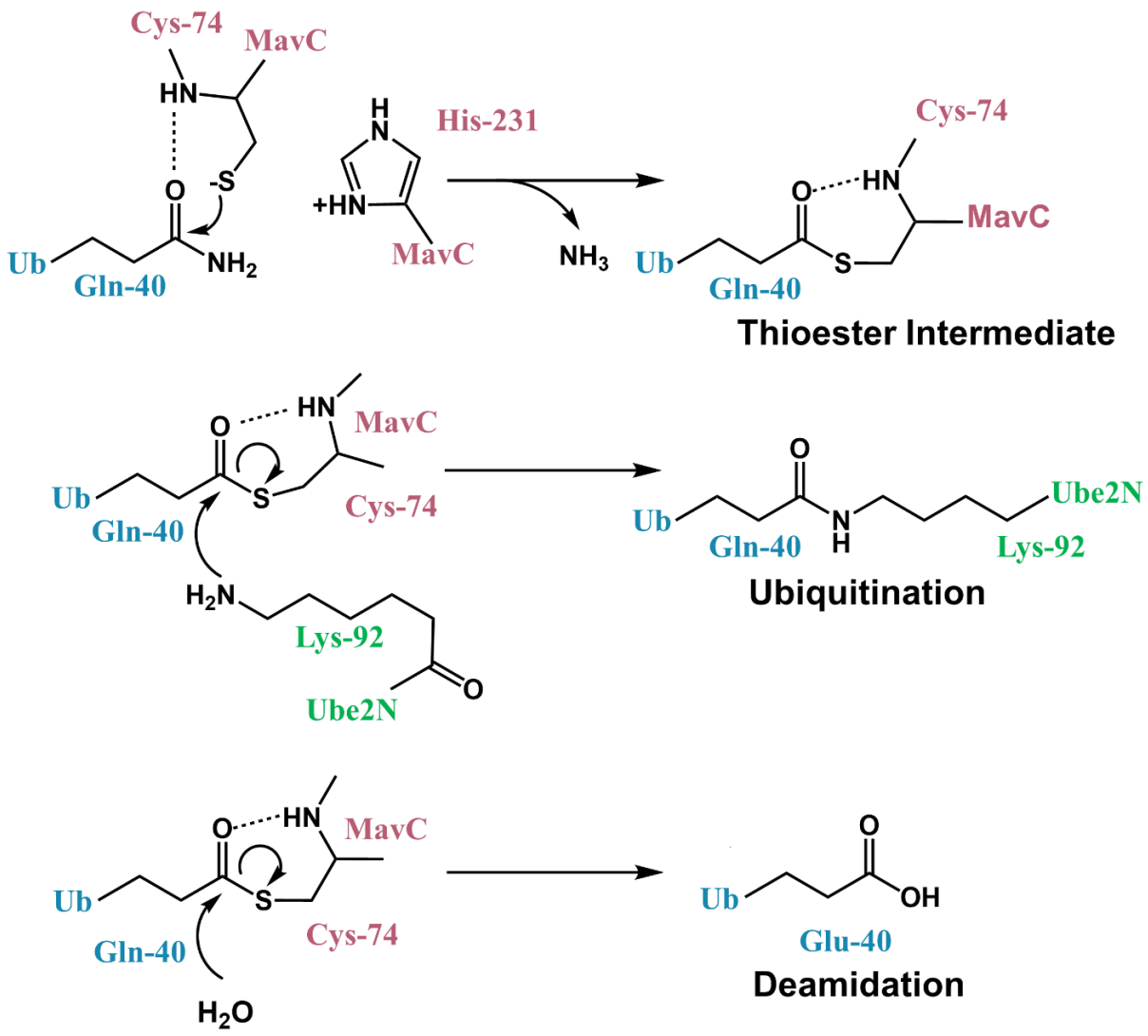


Figure 1-7: Overall scheme of MavC's catalytic activity. Thioester intermediate first forms between Cys74 of MavC and Glu40 of Ub, followed by attack of either Lys92 of Ube2N (ubiquitination) or attack of water (deamidation). Key residues are indicated.

It had been reported earlier that MavC, and its close homolog Lpg2148, or MvcA, possessed *in vitro* ubiquitin deamidating activity, and were also able to bind to Ube2N.[76] The crystal structure of these two enzymes revealed a homology to the Cif family, which was further supported by the fact that MavC and MvcA recognized and deamidated Gln40 of ubiquitin, the same residue targeted by Cif. However, a major difference was observed in that a ~100 residue insertion located approximately in the middle of MavC and MvcA formed a distinct domain from the core Cif-like fold. This domain was further found to be required for Ube2N binding.

The molecular basis of MavC's ubiquitinating activity remained unclear. Less clear was its strict substrate recognition, as well as the nature of the insertion domain. The studies described herein address these gaps by structural and biochemical means, presenting the crystal structure of MavC in complex with Ub-Ube2N. This has revealed the nature of the effector's interactions with ubiquitin and Ube2N and opened the door to a deeper understanding of MavC's true physiological substrate and of its dynamic nature.

CHAPTER 2. BIOCHEMICAL STUDIES OF THE SIDE LIGASE FAMILY

Portions of this chapter have been published in *Biochemistry* (DOI: 10.1021/acs.biochem.7b00664, DOI: 10.1021/acs.biochem.0c00067).[77,78]

2.1 Introduction

Two distinct questions were investigated pertaining to the nature of SidE-catalyzed ubiquitination. The first concerned the action of SidE enzymes against ubiquitin (Ub) chains. It is known that much of the ubiquitin inside of cells exists as part of a polyubiquitin chain instead of the free monoubiquitin form.[79] Further, when *Legionella* infects a cell, it is known that the endosome within which the bacterium resides (*Legionella*-containing vacuole, or LCV) is ubiquitinated with Ub chains.[45] To that end, the extent to which Ub chains could be acted upon by SidE enzymes and the possible effects of such a modification remained elusive. Due to the relative abundance in ubiquitin chains throughout the cell and particularly around the LCV, we hypothesized that SidEs may have the ability to utilize these species as substrates to further control host processes. The aim of the experiments presented here are to evaluate SdeA activity against Ub chains, using diubiquitin (diUb) as a model. This analysis showed that not only is SdeA able to transfer diUb to previously identified substrate proteins, but also to ADP-ribosylate and phosphoribosylate diUb chains of varying linkage types. This suggests that both the mART and PDE domains can act on Ub that is part of a chain and not only free Ub. Furthermore, the effect of Ub chain modification on their ability to be hydrolyzed by DUB enzymes in all major families was also evaluated. Interestingly, with one notable exception, the modified Ub chains proved resistant to DUB cleavage. The basis of this resistance to DUBs was also investigated by careful mutational experiments.

The second major question addressed here is whether it is possible to develop a quantitative and continuous assay for SidE-catalyzed ubiquitination. In order to better understand the enzymology of this effector, efficiently compare different substrates, mutants, and other variants, probe cell lysates for the presence of these enzymes, and also to screen for inhibitors of these important *Legionella* virulence factors, a demand exists to progress beyond the current method of determining ubiquitination activity, which thus far has simply been to conduct a reaction *in vitro*

and separate the components using SDS-PAGE gel electrophoresis. This technique, while indeed able to show ubiquitination as the appearance of a higher molecular weight band above the substrate protein, there exist several limitations. First, this technique cannot give a continuous readout of enzymatic activity. Also, it is relatively difficult to quantify ubiquitination from a stained gel. Further, high amounts of protein are needed to clearly visualize the progress of ubiquitination using this method. This demands a more robust technique for determining the activity of this novel enzymatic family. We present solutions to this problem by using fluorescent analogs of NAD^+ that, upon undergoing ADP-ribosylation, exhibit changes in their emission intensity.[80,81] These analogs allow us to measure the first step of SidE-catalyzed ubiquitination. For the crucial second step, we have accordingly utilized a fluorescent peptide acting as a synthetic substrate to develop an assay that transcends the limitations of the gel-based technique given above. This peptide is derived from the N-terminus of Rab1, a known substrate.[56] By measuring fluorescence polarization (FP) of this synthetic peptide as it becomes ubiquitinated, we are able to quantitatively and continuously measure SidE activity, allowing us to gain various insights into the nature of this process.

2.2 Materials and Methods

2.2.1 Cloning, Expression, and Purification of Recombinant Proteins

Three constructs of SdeA were utilized for ubiquitination and mART activity: the full length construct ($SdeA^{FL}$), a construct spanning residues 181-1000 that contained both the PDE and mART domains and which retained the ability to ubiquitinate protein substrates ($SdeA^{181-1000}$), and a construct spanning residues 519-1100 that contained only the mART domain in full and was therefore only able to perform mART activity ($SdeA^{519-1100}$). $SdeA^{FL}$ was obtained as a construct in the pQE30 vector from the Luo lab. $SdeA^{519-1100}$ and $SdeA^{181-1000}$ were cloned by Dr. Michael Sheedlo as glutathione-S-transferase (GST) N-terminal fusion proteins which used GST as a tag for affinity purification. Briefly, these genes were cloned into the pGEX-6P-1 vector via ligation after restriction digest by the enzymes BamHI and XhoI. Mutant constructs of SdeA were generated via site-directed mutagenesis and the presence of these mutations was confirmed via DNA sequencing. The previously identified SidE protein substrate Rab1 was also cloned into the pGEX 6P-1 vector.

Of the DUB constructs used in these studies: SdeA 1-200 ($SdeA^{DUB}$), AMSH, ChlaDUB1, and UCHL1 were cloned similarly into the pGEX-6P-1 vector. OTUlin was cloned into the pET SUMO vector. Recombinant DUB constructs were transformed into the Rosetta strain of *Escherichia coli*.

Wild type untagged ubiquitin (Ub) were cloned into the pRSET-A vector and mutants used for diUb synthesis (Ub K63R, Ub D77) were generated via site-directed mutagenesis, confirmed by DNA sequencing. A variant of Ub with a tryptophan residue after the C-terminal glycine (Ub-W) was cloned into the pGEX-6P-1 vector. Linear diUb linked via Met1 (M1 diUb) was cloned in the pET-26b vector (no tag), with its mutant constructs purchased from Genscript (Piscataway, NJ).

For expression and purification of recombinant GST-tagged proteins, *Escherichia coli* cells containing the protein of interest which was used to inoculate a starter culture in LB media supplemented with ampicillin. This starter culture was incubated under shaking conditions (~180 rpm) at 37 °C overnight. To express recombinant proteins, 6 L of LB media supplemented with ampicillin was inoculated from the starter culture and grown to an OD600 of approximately 0.4-0.6 and induced with isopropyl thio-D-galactopyranoside (IPTG), added to a final concentration of 350 µM. Following induction, the temperature was lowered to 18 °C and the cultures were further incubated under shaking conditions for 16-18 hours.

All other protein constructs were expressed similarly, with the exception of those cloned into pET-26b, where kanamycin instead of ampicillin was utilized as the antibiotic.

After expression, cells were harvested via centrifugation and resuspended in buffer (1X PBS, 400 mM KCl). This cellular suspension was then lysed by at least two passes through a French Press (Thermo Electron). The resultant lysate was cleared by ultracentrifugation at 100,000 xg for 1 hour at 4 °C. The resulting supernatant was decanted from the pellet and applied to Glutathione-Sepharose resin (GE Life Sciences) equilibrated with the same buffer used for cell resuspension (1X PBS, 400 mM KCl). After flowing the supernatant over the resin, the resin was washed with at least 5 column volumes of resuspension buffer. Following washing, GST-tagged proteins were eluted from the resin with GSH elution buffer (250 mM Tris pH 8, 500 mM KCl, 10 mM reduced glutathione). After elution, the GST fusion tag was removed by incubating the fusion protein with GST-tagged PreScission™ Protease (GE Life Sciences) at 4 °C overnight. The eluate was also dialyzed during this period into fresh resuspension buffer in order to remove glutathione.

The dialyzed and cleaved protein solution was once again passed over the Glutathione-Sepharose resin, which removed free GST and GST-tagged PreScission™ Protease. Proper cleavage of the tag and removal was monitored by SDS-PAGE (Figure 2-1). This solution was then concentrated to less than 4 mL and purified by size exclusion chromatography on a Superdex S200 column (GE Life Sciences) in size exclusion chromatography buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT). Each fraction that was collected was then analyzed for purity on SDS-PAGE.

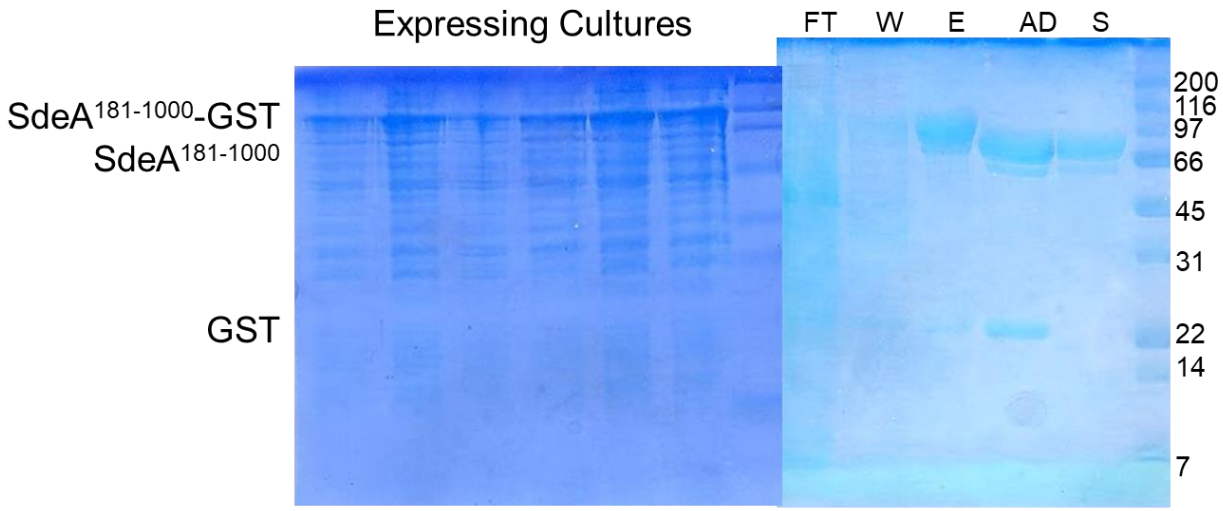


Figure 2-1: Expression and purifiaction of GST-tagged SdeA$^{181\text{-}1000}$. Left panel shows expression of protein in cultures. Right panel shows initial flow-through from GST column (FT), wash (W), eluate (E), sample after dialysis and tag cleavage (AD), and sample after second pass through GST column to subtract GST (S).

His-tagged proteins were harvested, lysed, and ultracentrifuged similarly as above, but the resulting supernatant was then added directly to Ni-NTA resin (Qiagen). After flowing this supernatant over the resin, the resin was washed with at least 5 column volumes of resuspension buffer. Following initial washing, a secondary wash step was performed with resuspension buffer containing 50 mM imidazole to further remove nonspecifically bound proteins before elution with resuspension buffer containing 300 mM imidazole. This solution was then concentrated to less than 4 mL and purified by size exclusion chromatography on a Superdex S200 column (GE Life Sciences) in size exclusion chromatography buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT). Each fraction that was collected was then analyzed for purity on SDS-PAGE.

Non-tagged proteins were harvested as described above and resuspended in cation-exchange buffer (50 mM sodium acetate pH 4.5), then lysed via French press followed by incubation in a water bath of approximately 80 ºC for 30 minutes. This caused heat-sensitive proteins to come out of solution but retained the heat-stable untagged ubiquitin. This lysate was then ultracentrifuged similarly as GST-tagged proteins described above. After centrifugation, the clarified supernatant's pH was adjusted to 4.5 by addition of acetic acid. Precipitated proteins were clarified by further centrifugation. The pH-adjusted supernatant was applied to chromatographic resin. Cation-exchange chromatography was performed using SP Sepharose Fast Flow resin (GE Healthcare). After applying the supernatant, the resin was washed with 5 column volumes of cation exchange buffer. Proteins were eluted using a gradient elution with gradually increasing concentrations of NaCl, up to 1 M. Fractions containing purified proteins were pooled and concentrated.

Other proteins: USP7 was received as a generous gift from Andrew Mesecar. MIY1 was received as a generous gift from Yogesh Kulathu. USP2 was purchased from BostonBiochem (Cambridge, MA). SidJ purified from *Legionella pneumophila* was received from the Luo Lab and prepared as described.

All proteins were used immediately after purification or concentrated and stored at -80 °C.

2.2.2 Preparation of Diubiquitin

DiUb chains were synthesized as previously described.[82] Briefly, K63-linked diUb was generated by combining E1 enzyme with the heterodimeric E2s Ubc13 and Uev1a at 37 °C. Reactions were quenched with MonoS buffer (50 mM Na acetate pH 4.5) and diUb was separated from other proteins by cation exchange chromatography using a MonoS column (GE Healthcare), via a gradient elution using an elution buffer of 50 mM Na acetate pH 4.5, 1 M NaCl.

2.2.3 Mass Spectrometry Analysis of SdeA-Catalyzed Ubiquitin Modification

To modify diUb chains, M1, K11, K48, and K63-linked diUb at 50 µM was incubated with $SdeA^{519\text{-}1100}$ or $SdeA^{181\text{-}1000}$, along with 1 mM NAD^+ in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT). The reaction was allowed to proceed for 1 hour at 37 °C.

Top-down LC-MS separation was performed by Yiyang Zhou (Wirth Lab, Purdue University). Briefly, 0.05 µg of protein from each reaction was injected into a Halo™ ES-C18 reverse phase chromatographic column for separating reaction products. Sample injection and elution was performed by a Thermo Accela UHPLC system and a coupled Thermo LTQ mass spectrometer was used for analysis of the mass of eluted proteins.

Mobile phases A and B were water and 50/50 acetonitrile/water, all containing 0.1 % difluoroacetic acid (DFA) along with 100 µM tris(2-carboxyethyl)phosphine (TCEP) to maintain a reduced oxidation state to prevent heterogeneity of proteins. A gradient elution from 20% B for 2 minutes followed by 20%-100% B for 10 minutes followed by 100 % for 1 minute was used to elute proteins in the injected sample.

2.2.4 Diubiquitin Cleavage Assays

50 µM of diubiquitin was incubated with either $SdeA^{519-1100}$ or $SdeA^{181-1000}$ (5 µM) and NAD^+ (1 mM) for 1 hour at 37 °C in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT) in order to generate modified diUb species. A control was also run where NAD^+ was excluded. These reaction mixtures were then treated with the DUBs specified and reacted for 1 hour. The reaction with M1 diUb mutants and OTUlin was carried out for 16 hrs due to lowered reactivity of the mutated diUb substrate. Reactions were then subjected to analysis by SDS-PAGE and svisualized by staining with Coomassie Blue.

DUBs were added to the modified diUb as follows: OTUlin: 50 nM (vs. M1 WT), 200 nM (vs. M1 mutants), $SdeA^{DUB}$: 100 nM, ChlaDUB1: 100 nM, AMSH: 100 nM, USP7: 100 nM, USP2: 50 nM, UCHL1: 200 nM, SidJ: 500 nM, MIY1: 200 nM

2.2.5 Preparation of Nucleotide Cofactors and Analogues

Nicotinamide 1,N6-ethenoadenine dinucleotide (εNAD^+) was prepared as described.[80] Briefly, NAD^+ (Sigma) was reacted in aqueous chloracetaldehyde at pH 4.5 at room temperature while stirring and away from light for 4 days until UV absorbance of the reaction mixture held constant. The reaction was then purified first by charcoal decolorization followed by ion-exchange chromatography. The purification step was performed by Prasanth Reddy Nyalapatla (Ghosh Lab, Purdue University).

$N^{tz}AD^{+}$ was prepared as described.[81] Briefly, ^{tz}AMP was reacted with activated β-nicotinamide mononucleotide, forming the desired product. This synthesis was performed by Alexander Rovira (Tor Lab, University of California, San Diego).

Purified NAD^{+} analogues were characterized by mass spectrometry (Figure 2-2). Analogues dissolved in water were injected into a Waters Acquity UPLC system utilizing a C18 reversed-phase column. Elution was performed isocratically using a 95% water/5% acetonitrile mixture as the mobile phase. Mass spectrometric data was collected by a coupled Waters SQD2 mass spectrometer.

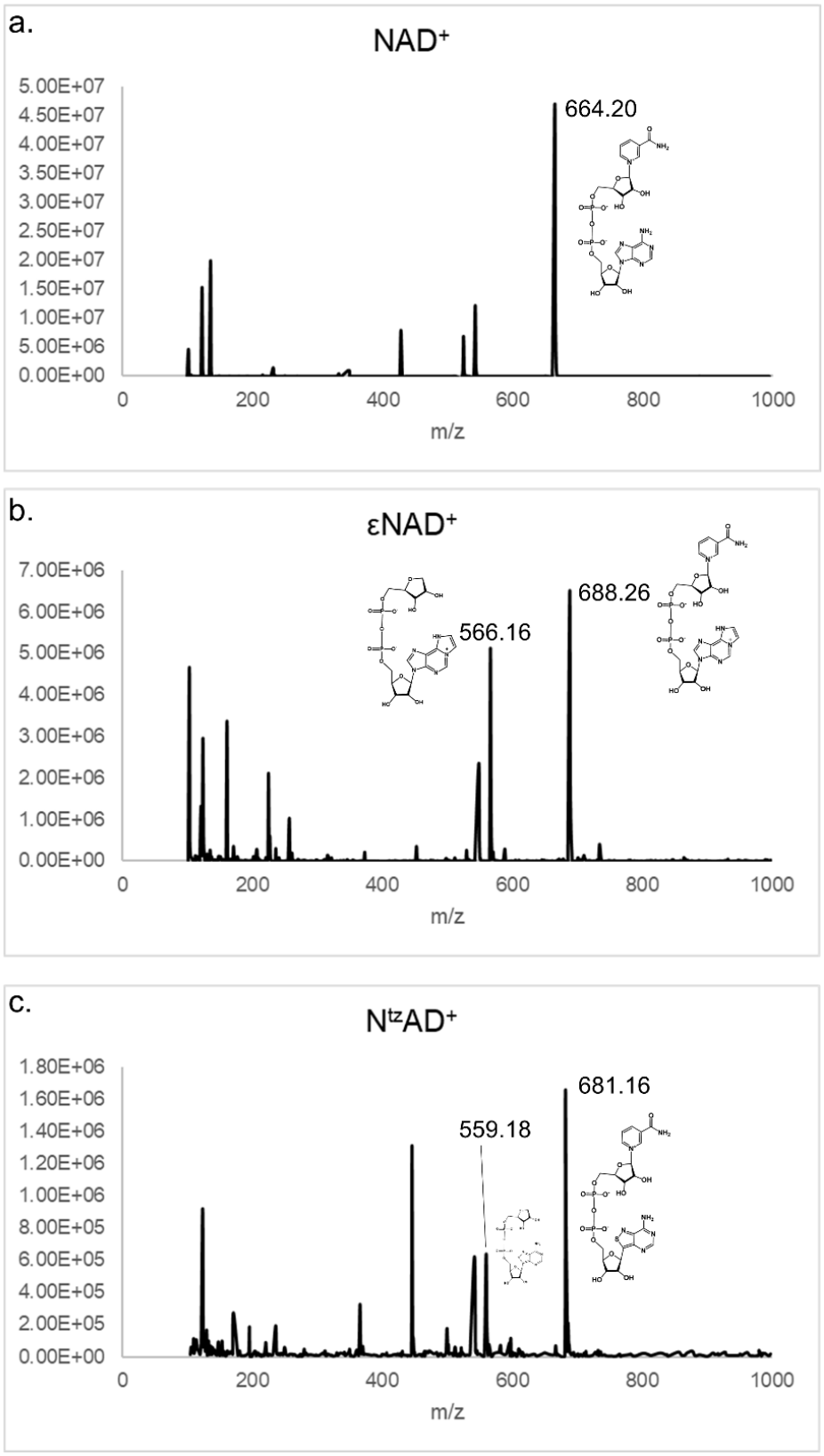


Figure 2-2: Mass spectra of A) NAD^+,B). εNAD^+, C) $N^{tz}AD^+$

2.2.6 Synthesis of Fluorescent Peptide Substrates

Peptides were synthesized using solid-phase peptide synthesis, using a standard Fmoc-based procedure. ChemMatrix Rink Amide resin served as the solid support. The Fmoc-protected amino acid was added to the resin along with the coupling agents HBTU and DIEA. Following this coupling step, the resin was washed with DMF, DCM, MeOH, DCM, and DMF and the Fmoc group was deprotected with 20% piperidine. The wash step was repeated and the whole process was repeated with the next amino acid until all desired amino acids were coupled to the resin. The fluorescein group was added to the N-terminus of the peptides by adding NHS-Fluorescein (Fisher Scientific) to deprotected peptide, along with DIEA and DMF.

Fluorescent peptides were then cleaved under acidic conditions from the resin by incubation with 95% TFA, 2.5% TIPS, and 2,5% water. The supernatant, containing cleaved peptides, was subjected to precipitation using cold diethyl ether and isolated by centrifugation. After further drying under reduced pressure, peptide was re-suspended in DMSO to a concentration of 10 mg/mL and purified by reverse-phase HPLC. A Luna C18 semi-prep column running a 60 minute linear solvent gradient from 15%-90% B was utilized for this step(A: water + 0.1% TFA, B: Acetonitrile + 0.1% TFA). Ryan Curtis (Chmielewski Lab, Purdue University) performed the peptide synthesis. See Appendix A for relevant HPLC traces and mass spectra.

2.2.7 Fluorescent Assays for ADP-Ribosylating Activity of SdeA

ADP-ribosylation assays were performed by combining NAD^+ or analogues εNAD^+ or $N^{tz}AD^+$ (100 µM) with ubiquitin in fluorescent assay buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mg/ml BSA). Reaction mixtures were left to equilibrate at room temperature for 3 minutes, then initiated by adding SdeA to a final concentration of 0.5 µM. SdeA constructs used included $SdeA^{FL}$, $SdeA^{181\text{-}1000}$, $SdeA^{519\text{-}1100}$, and $SdeA^{E/A}$. Fluorescence intensity at 410 nm was measured using a Biotek Cytation Multi-Mode Plate Reader using excitation wavelengths of 300 nm (εNAD^+) or 338 nm ($N^{tz}AD^+$) All reactions were conducted at room temperature and at a final concentration of 100 µL. Assays were also performed at least in triplicate.

2.2.8 Generation and Purification of Fluorescent Ub-εADPR

εADP-ribosylated ubiquitin (Ub-εADPR) was generated at a scaled-up amount by incubating 5 μM $SdeA^{519-1100}$ with 100 μM ubiquitin and 1 mM εNAD^+, at a total volume of 20 mL. A sample was analyzed via SDS-PAGE and visualized with UV and Coomassie Blue to confirm reaction (Figure 2-3A). This mixture was allowed to react in the dark under gentle agitation for 3 hours at 25 ºC, then concentrated to 4 mL and subjected to cation-exchange chromatography by passing through a MonoS column (GE Healthcare). Fluorescent fractions corresponding to Ub- εADPR were pooled and collected (Figure 2-3B). Purity was confirmed by LC-MS (Figure 2-3C).

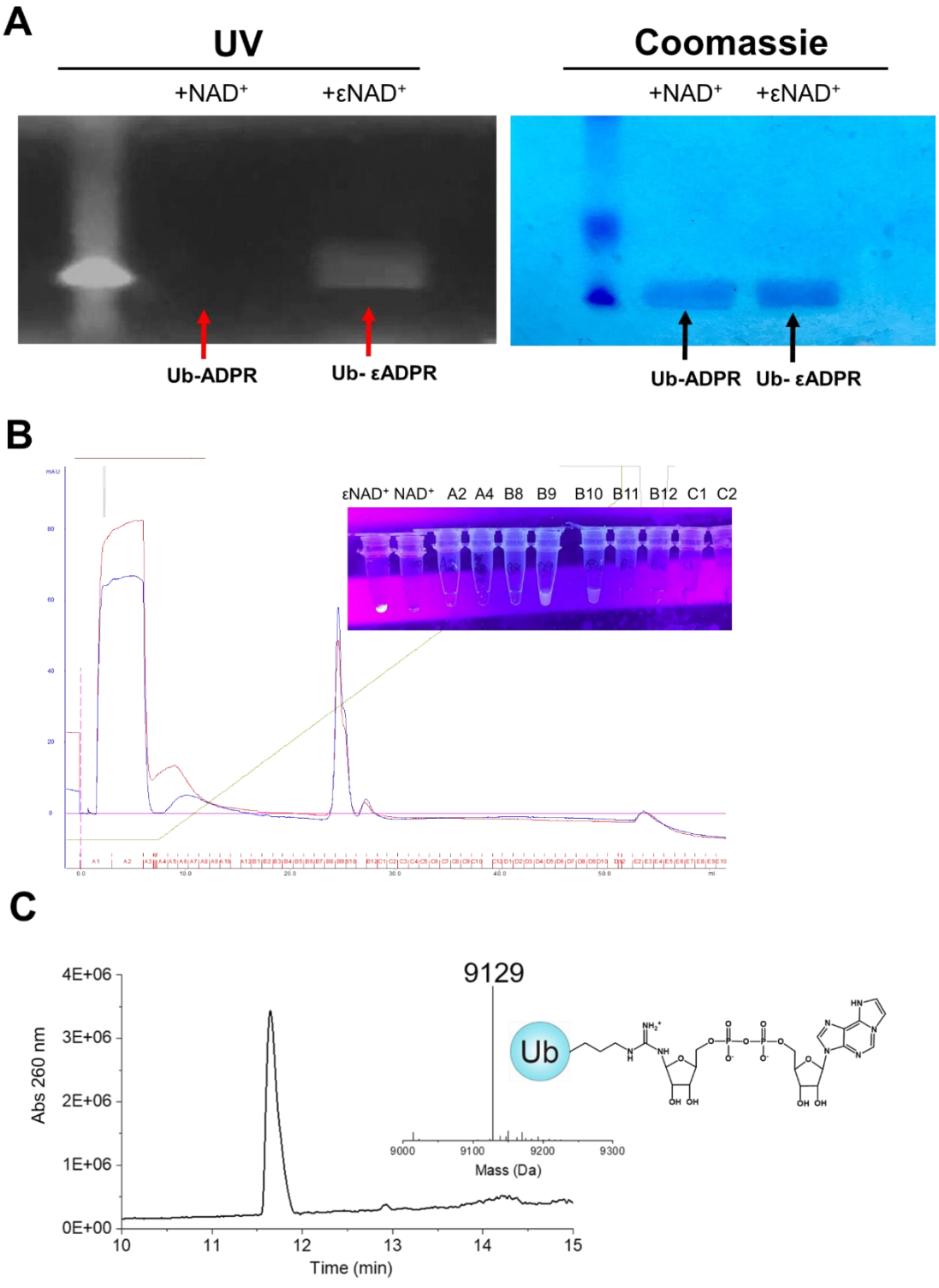


Figure 2-3: Purification of Ub-εADPR. A) SDS-PAGE of SdeA mART reaction using NAD^+ or εNAD^+. B) Ion-exchange chromatogram of scaled-up reaction, with UV analysis of fractions. C) LC-MS of purified Ub-εADPR sample.

2.2.9 Gel-Based Assay of SdeA-Catalyzed Ubiquitination

To evaluate ubiquitination activity of SdeA using diubiquitin as a substrate using a gel-based assay, $SdeA^{181-1000}$ at a concentration of 5 µM was combined with Rab1 (6 µM), and M1-linked diubiquitin (50 µM) in the presence or absence of NAD^+ (1 mM). Reaction was conducted for 1 hour at 37 ºC, quenched with 5X SDS-PAGE loading buffer and analyzed by SDS-PAGE with Coomassie Blue staining.

To determine ubiquitination of the fluorescent peptide substrate, $SdeA^{181-1000}$ (0.25 µM), ubiquitin (100 µM), and fluorescent peptide (MAS, 10 µM) were incubated in the presence or absence of 100 µM NAD^+ for 10 min at room temperature. Reactions were quenched with 5X SDS-PAGE loading buffer and analyzed by SDS-PAGE. Gels were first imaged for in-gel fluorescence under a UV light to visualize emissive protein bands, and then stained with Coomassie Blue to visualize all proteins.

2.2.10 Fluorescence Polarization (FP) Assays of SdeA-Catalyzed Ubiquitination

0.25 µM of enzyme was combined with 100 µM ubiquitin, 10 µM fluorescent peptide, in fluorescent assay buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mg/ml BSA). Reaction mixtures were allowed to equilibrate at room temperature for 3 minutes, then reactions were initiated by the addition of NAD^+ to 100 µM. Fluorescence polarization was measured with a Cytation Multi-Mode Plate Reader (BioTek), using 485 nm excitation and 528 nm emission filters.

Michaelis-Menten kinetic analysis was performed by utilizing the assay above but the concentration of peptide was varied from 5-40 µM. Initial velocity (Vo) was determined by the slope of the initial linear portion of the reaction's progress curve. Vo was plotted versus [peptide] and the curve was fit to the Michaelis-Menten equation.

To assess the ubiquitinating activity of cell lysates, the reaction mixture consisted of *Legionella pneumophila* lysate (Luo Lab), HEK293 cell lysate (Kinzer-Ursem Lab) in the place of SdeA.

To examine the effect of possible inhibitors of SdeA, 0.25 µM of $SdeA^{181-1000}$ was incubated with 5 mM ADPR or AMP (Sigma) for 30 min on ice in fluorescent assay buffer. Fluorescence polarization was then measured after adding ubiquitin to 100 µM, fluorescent peptide to 10 µM, and finally NAD^+ to 100 µM.

2.3 Results

2.3.1 SdeA Modifies Both Ub Groups of a Diubiquitin Chain

In order to test whether Ub chains were recognized by SidE enzymes as substrates, diUb chains of M1, K11, K48, and K63 linkages were used as model substrates and combined with SdeA and the essential cofactor NAD^+. Both $SdeA^{519-1100}$ and $SdeA^{181-1000}$ constructs were used. Mass spectrometry clearly showed a shift in the mass of diUb chains corresponding to either 2 -ADPR groups when incubated with $SdeA^{519-1100}$ or 2 -PR groups when incubated with $SdeA^{181-1000}$ (Figure 2-4, Appendix B, Table 2). This indicates that both the mART and PDE domains are able to recognize both moieties of a diUb chain in various orientations. All diUbs tested appeared, within the parameters of the experiment, to be completely converted into their doubly ADP-ribosylated or doubly phosphoribosylated forms. A notable exception was linear, or M1-linked diubiquitin: when incubated with $SdeA^{181-1000}$, the mass spectrum detected heterogeneity, with ions corresponding to the addition of 1 PR group, 2 PR groups, and also 1 PR with 1 ADPR group. Since incubation with $SdeA^{519-1100}$ revealed complete conversion into a doubly-ADP-ribosylated species, this suggested a slower rate of conversion of linear diUb-2xADPR into diUb-2xPR by the PDE domain. To determine whether one specific Ub group of M1 diUb was more slowly converted, mutant versions of M1 diUb where either the proximal or distal Arg42 was converted to lysine and therefore unable to be modified by SdeA were generated (M1 diUb R42K or the distal mutant, M1diUb R118K or the proximal mutant). Upon conducting the modification reaction and mass spectrometric analysis, it was observed that while both mutants were completely ADP-ribosylated by the mART domain, as expected, the R118K mutant was also completely converted into the corresponding phosphoribosylated species, where in contrast the R42K mutant displayed heterogeneity between the ADP-ribosylated and phosophoribosylated forms (Appendix B, 1C-1E).

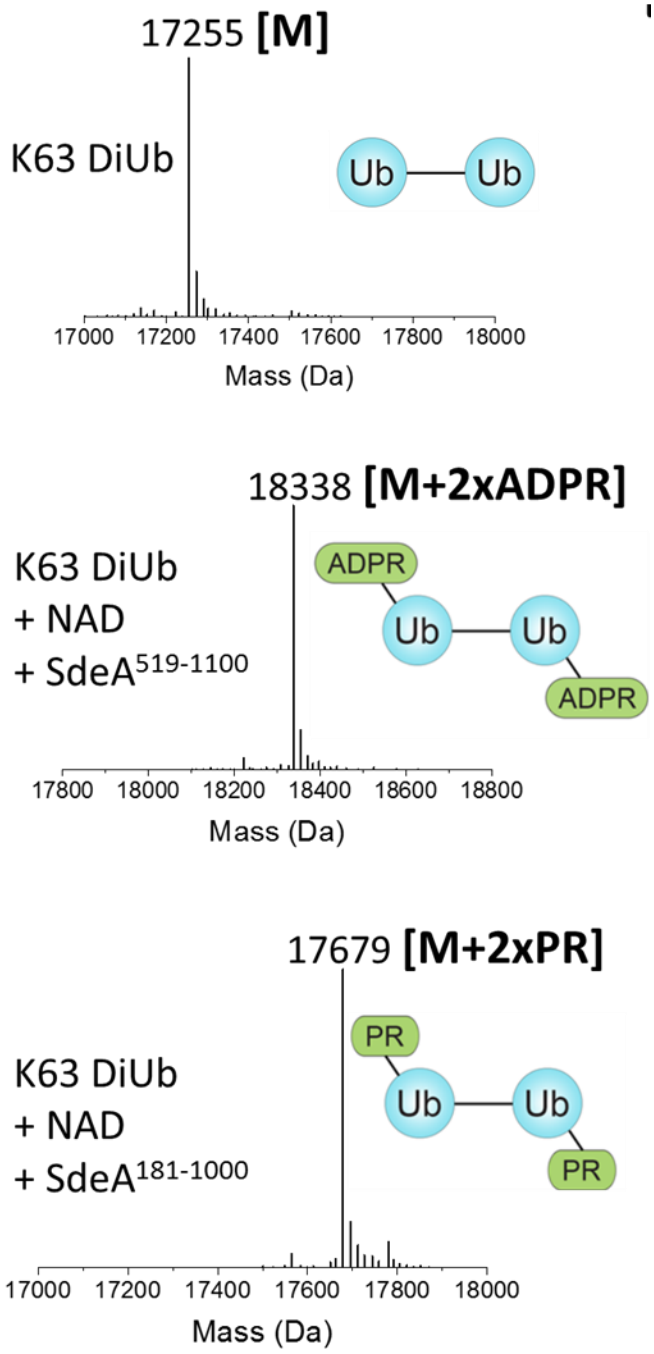


Figure 2-4: Mass spectrometric analysis of SdeA-catalyzed modification of K63 diUb. See Appendix B for full LC-MS data for each chain type and Table 2 for a summary of results.

Table 2-1: LC-MS of modified versus unmodified Ub chains

	Theoretical	Observed	Difference	ΔMass
K11 DiUb	17111.7 WT+WT	17112	0.3	0
K11 DiUb-2xADPR	18195.7	18194	-1.7	-2
K11 DiUb-2xPR	17535.7	17536	0.3	0
K48 DiUb	17284.78 K48R K11R+D77	17283	-1.78	0
K48 DiUb-2xADPR	18368.78	18366	-2.78	-1
K48 DiUb-2xPR	17708.78	17707	-1.78	0
K63 DiUb	17254.78 K63R+D77	17255	0.22	0
K63 DiUb-2xADPR	18338.78	18338	-0.78	-1
K63 DiUb-2xPR	17678.78	17679	0.22	0
M1 DiUb	16997.57 ΔG75-76	16997	-0.57	0
M1 DiUb-2xADPR	18081.57	18080	-1.57	-1
M1 DiUb-2xPR	17421.57	17421	-0.57	0
M1 DiUb R118K	17381.02 GPLGS, ΔG75-76	17377	-4.02	0
M1 R118K-ADPR	17923.02	17919	-4.02	0
M1 R118K-PR	17593.02	17590	-3.02	1
M1 DiUb R42K	17381.02 GPLGS, ΔG75-76	17378	-3.02	0
M1 R42K-ADPR	17923.02	17920	-3.02	0
M1 R42K-PR	17593.02	17591	-2.02	1

To determine whether diUb could be transferred to an actual protein substrate, the product of a ubiquitination including the previously identified substrate Rab1 was conducted and analyzed by SDS-PAGE. A higher molecular weight band corresponding to a ~17 kDa shift was observed along with a reduction in the intensity of the unmodified Rab1 band (Figure 2-5). Furthermore, an upward smear in the SdeA band was observed, indicative of autoubiquitination, a known characteristic of SdeA. This result suggests that diUb chains may also be transferred to protein substrates as well as modified.

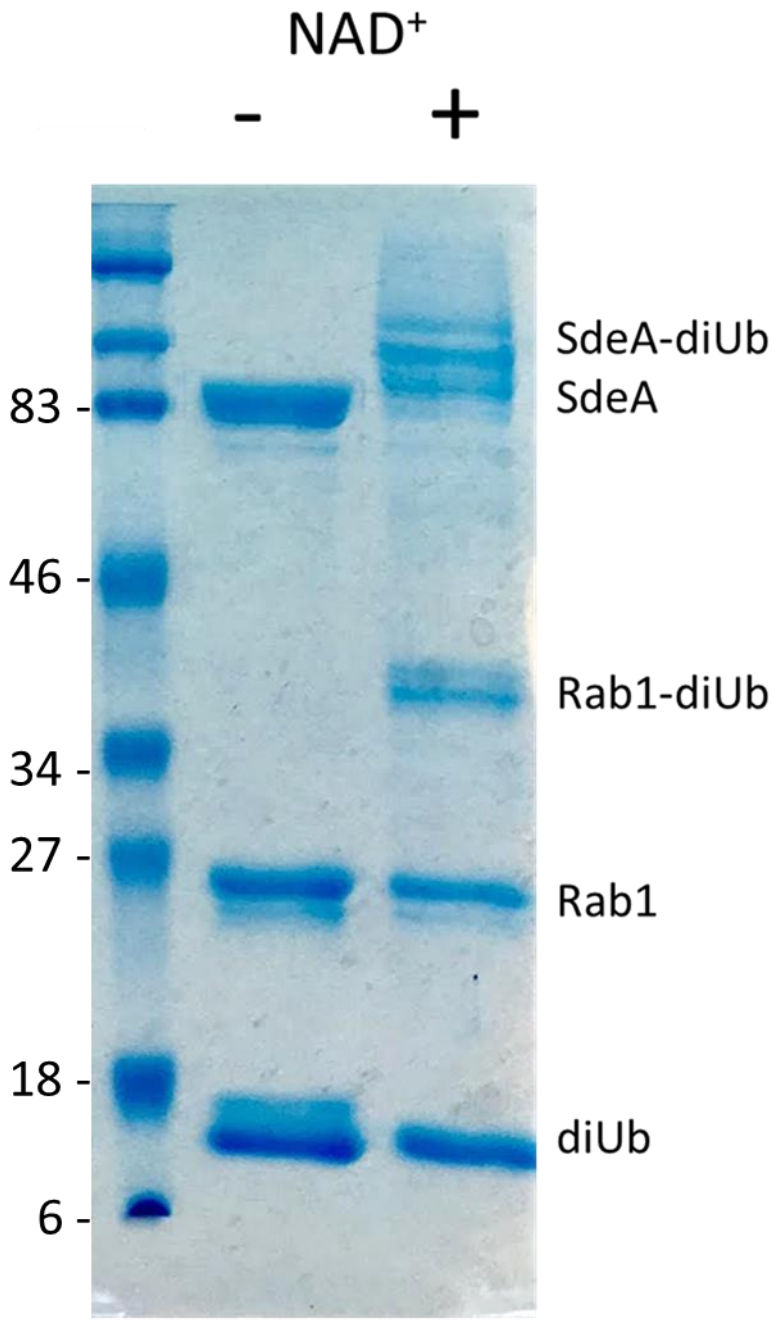
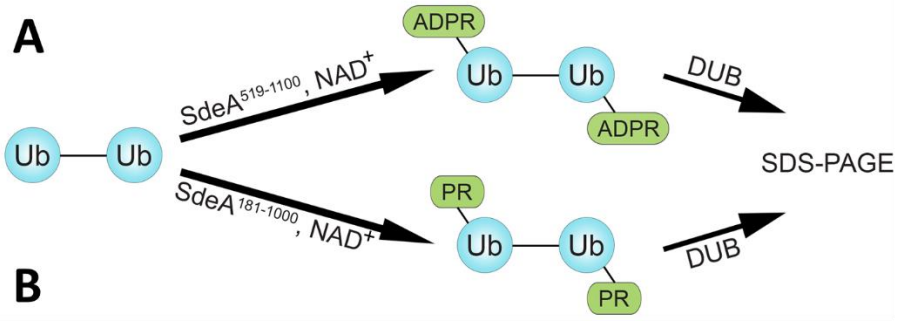


Figure 2-5: SdeA is able to transfer diUb to a protein susbtrate. SDS-PAGE gel of ubiquitination reaction is displayed here.

2.3.2 Modified Ubiquitin Chains are Resistant to DUB Hydrolysis

To further examine the properties of SdeA-modified Ub chains we conducted deubiquitination assays to compare the ability of these chains to become processed by DUBs.

Representative DUBs from a variety of families were tested (Figure 2-6). Assays were conducted in a 2-step fashion, where diUbs were (1) incubated with either the 519-1100 or 181-1000 construct of SdeA and NAD^+, generating the desired modified diUb species, then (2) treated with a DUB that is able to act on the unmodified diUb chain. DUB activity was indicated by the conversion of diUb into monoUb on an SDS-PAGE gel. Interestingly, modification of ubiquitin chains resulted in an inhibition of nearly every DUB tested, with a notable exception being that of AMSH. AMSH belongs to the unique JAMM family of DUBs, which is unique in that it utilizes a zinc metalloprotease mechanism rather than a cysteine protease mechanism which all other DUB families employ.[83]



DUB tested	Family	Mechanism
USP2	USP	Cys protease
USP7	USP	Cys protease
UCHL1	UCH	Cys protease
OTUlin	OTU	Cys protease
MIY1	MINDY	Cys protease
AMSH	JAMM	Zn protease
$SdeA^{DUB}$	Bacterial	Cys protease
ChlaDUB1	Bacterial	Cys protease
SidJ	Bacterial	?

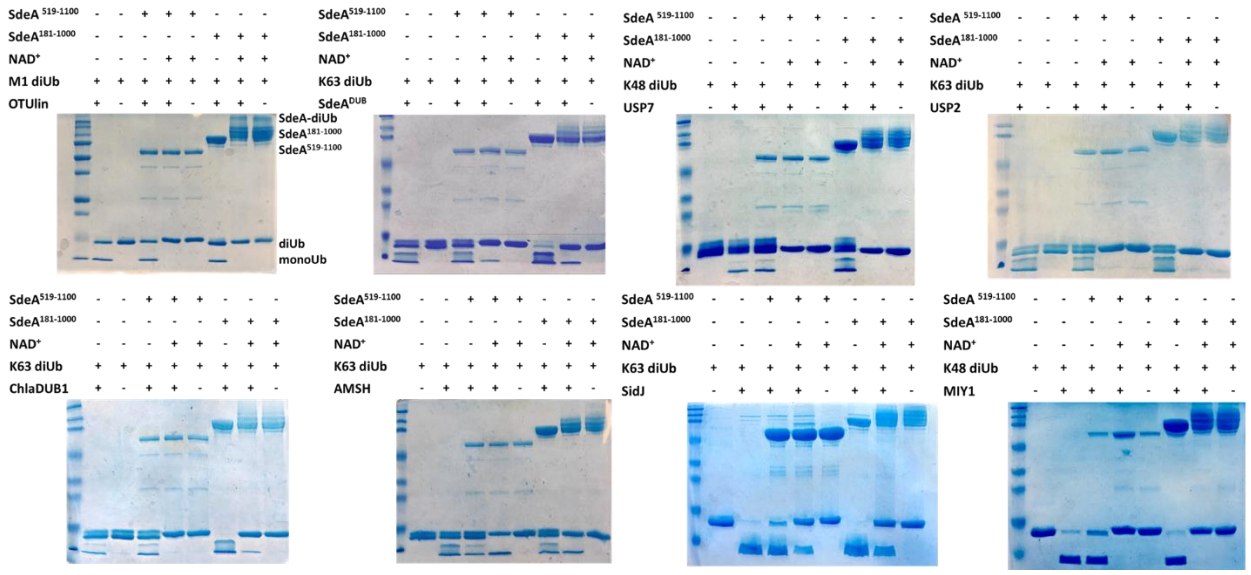


Figure 2-6: Two-step assay showing DUB activity against modified diUb linkages. A) Overview of assay procedure with list of DUBs tested. B) SDS-PAGE gel analysis of inhibition assay using the DUBs listed in A)

We speculated that this relative inability of most DUBs to process modified diUb was due to steric clashes between the introduced -ADPR or -PR groups and the DUB's Ub binding interface. To that end, we examined an available crystal structure of the human DUB OTUlin in complex with M1-linked diUb, its preferred substrate (Figure 2-7A).[84] We noticed that Arg42 (the targeted Arg on the distal Ub moiety) made key interactions with OTUlin in the crystal structure, whereas the Arg42 equivalent on the proximal Ub (Arg118) did not make any contacts with OTUlin and in fact faced away from the enzyme. We hypothesized that Arg118 modification would not affect hydrolysis, but Arg42 modification would. To test this hypothesis, we used the R42K and R118K mutants generated earlier in our two-step assay (Figure 2-7B). Indeed, the R42K mutant, while slower to react with OTUlin overall due to the mutation of a binding residue, showed no difference in hydrolysis when modified or unmodified by SdeA. In contrast to this, the R118K mutant was markedly unable to be cleaved by OTUlin when modified (Figure 2-7B). This supported our hypothesis that steric disruption of ubiquitin binding to DUBs may explain the inhibition of activity after modification.

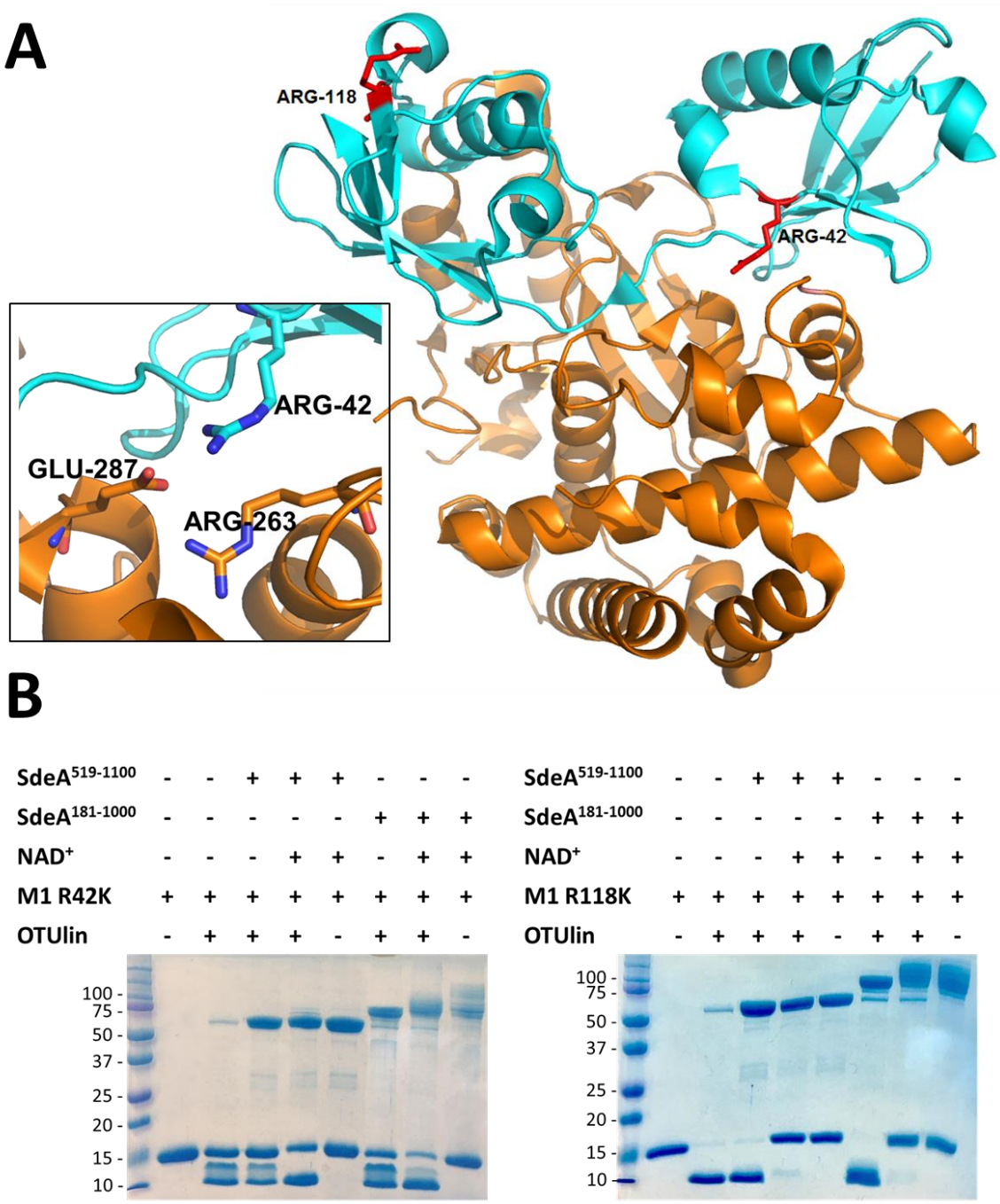


Figure 2-7: Examining structural basis of reduction in DUB hydrolysis activity. A) Crystal structure of M1 diUb in complex with OTUlin. Inset shows binding site for Arg42. B) Two-step assay as described in previous figure using diUb mutants as substrates.

2.3.3 Monitoring of SdeA ADP-ribosyltransferase activity

In order to continuously measure the ADP-ribosylating activity of SdeA, the fluorescent NAD^+ analogues εNAD^+ and $N^{tz}AD^+$ were utilized in an intensiometric assay (Figure 2-8A).[85] As expected, an increase in fluorescence was observed upon ADP-ribosylation, suggesting that these analogues were accommodated as substrates by the mART domain of SdeA (Figure 2-8B-C). The $SdeA^{E/A}$ mutant exhibited no fluorescence increase in agreement with its lack of catalytic activity. Interestingly, $SdeA^{FL}$ was noticeably faster in reaction compared to $SdeA^{181\text{-}1000}$, which in turn was faster than $SdeA^{519\text{-}1100}$. These observations suggested that interdomain interactions play a role in optimal enzymatic activity. The product of this reaction, Ub-εADPR, was also purified and isolated (Figure 2-3).

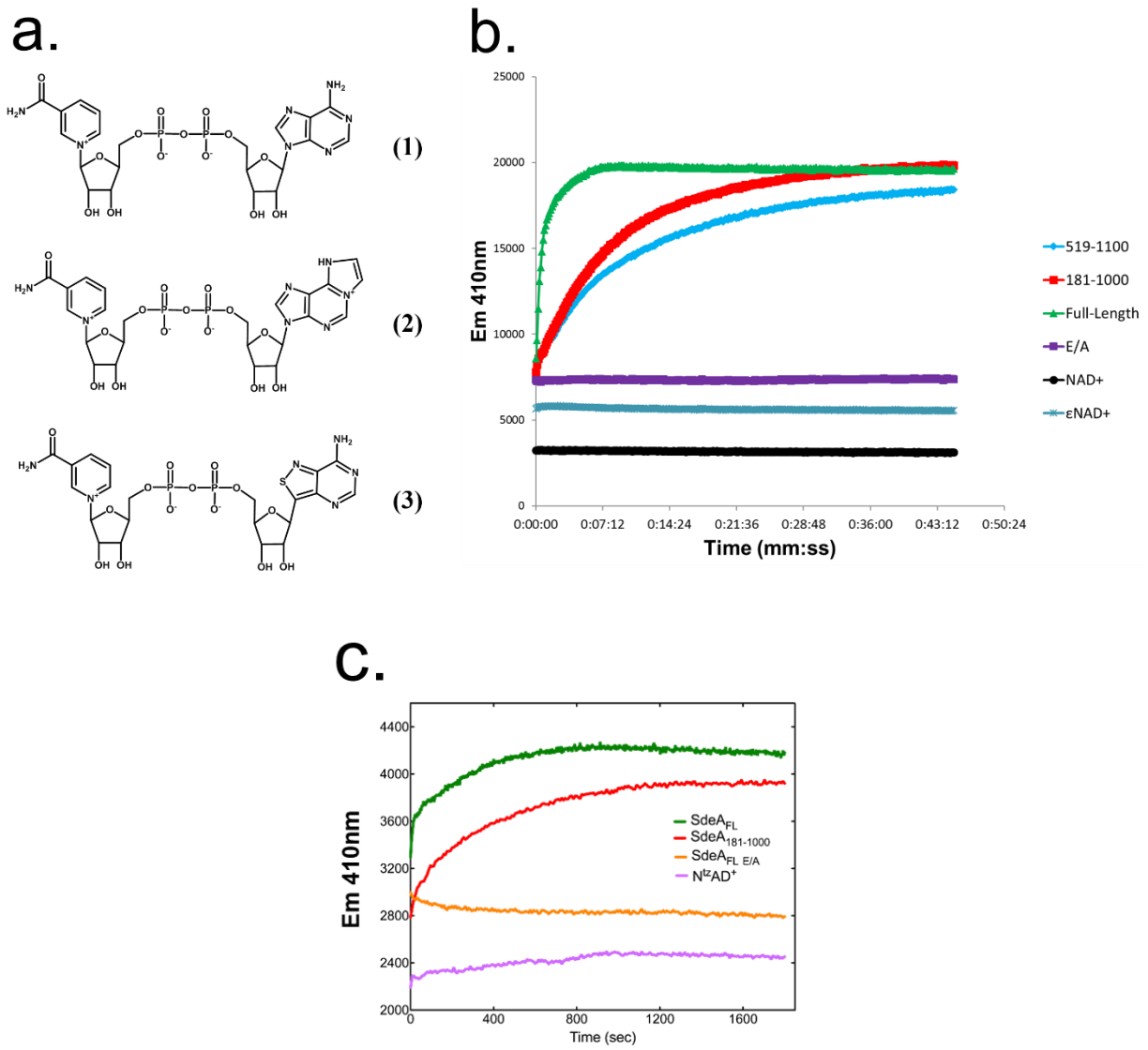


Figure 2-8: Assays of SdeA mART domain using fluorescent analogues. A) Structures of fluorescent analogues utilized. B) The fluorescent intensity increases significantly and can be used to track ADP-ribosylation. C) Fluorescent assay using $N^{tz}AD^{+}$

2.3.4 A Continuous, Fluorogenic Assay for Noncanonical Serine Ubiquitination Catalyzed by SdeA

To measure the ubiquitinating activity of SidE enzymes, a synthetic, peptide substrate was generated (see Materials and Methods) that contained a fluorescein group. This peptide, bearing the sequence MSSMNPEYD, represents the first 9 amino acids of Rab1, a known protein substrate of SdeA (Figure 2-9). The peptide is likely unstructured and flexible; in the crystal structure of Rab1 these residues are unresolved. Previous studies also showed using mass spectrometry that

the N-terminal Ser residues served as the major site of ubiquitination and suggested that flexible regions of proteins containing Ser residues may be targeted.

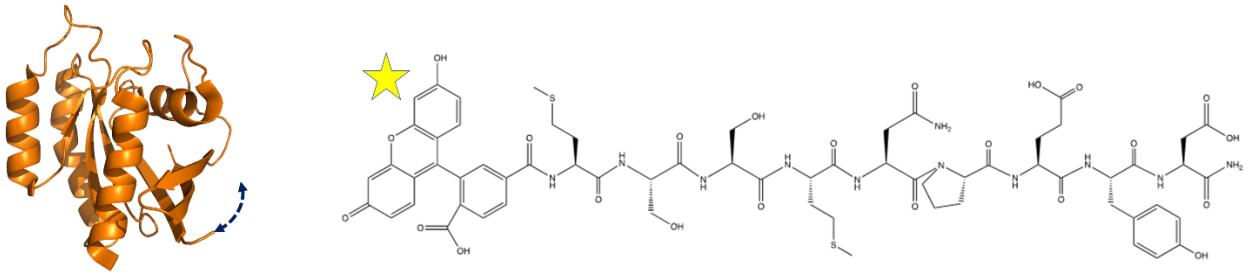


Figure 2-9: Crystal structure of Rab1 (PDB) highlighting the N-terminal region with a dotted line. Structure of the fluorogenic peptide substrate is also reported here.

First, to confirm that ubiquitination could take place, this peptide was allowed to react with SdeA, Ub, and NAD^+. This reaction was separated by SDS-PAGE, and a higher molecular weight fluorescent band appeared in the reacted sample, where a control reaction not containing NAD^+ remained without this new band (Figure 2-10). This result indicated the formation of a ubiquitinated peptide. Indeed, mass spectrometric analysis of the protein extracted from this new band revealed the presence of a peak containing the predicted Ub-PR-Peptide.

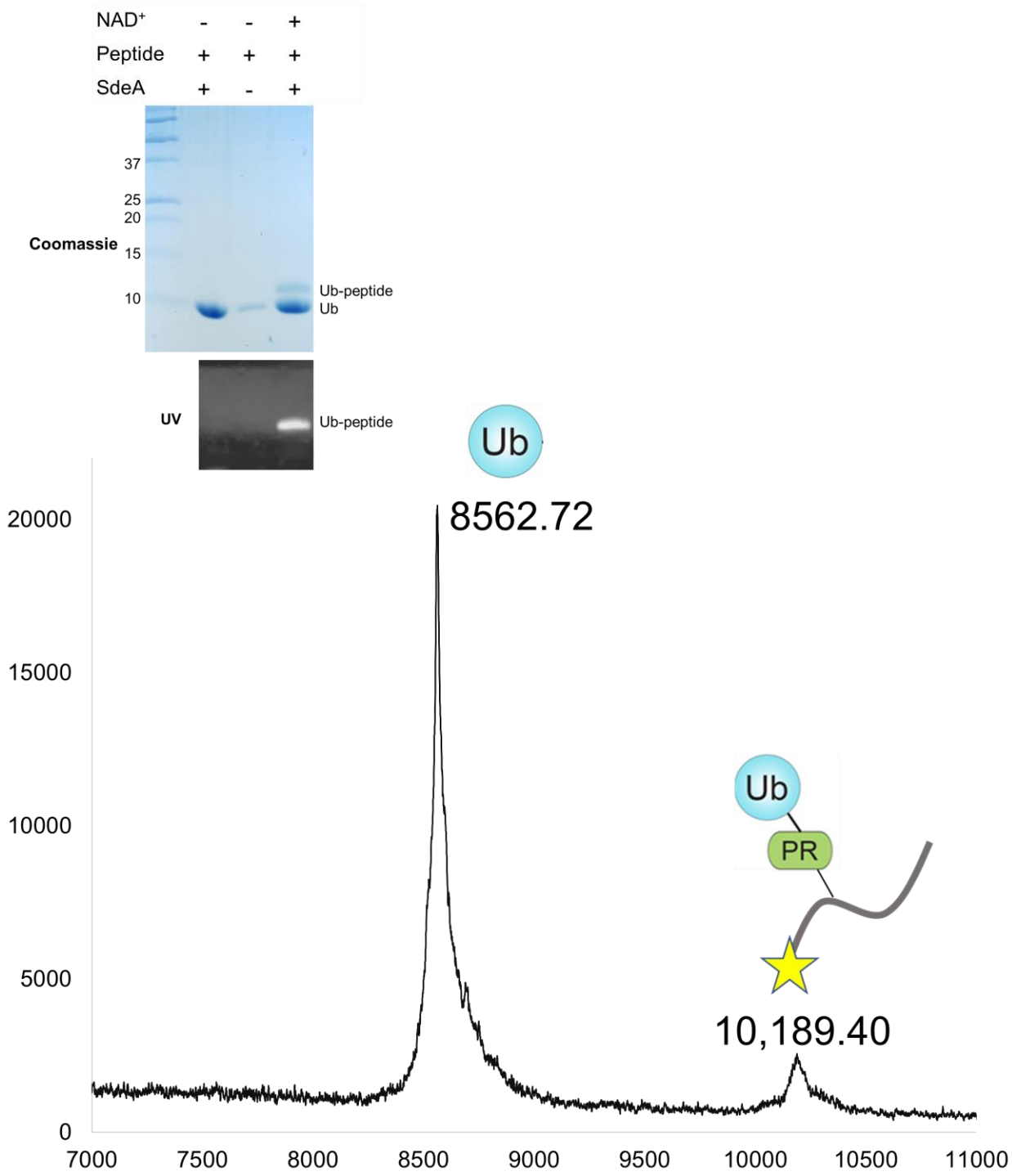


Figure 2-10: Gel-based assay confirming the ability of the synthetic peptide to become ubiquitinated by SdeA. Coomassie and UV-exposed gels shown along with MALDI-MS of Ub-peptide band.

Due to the approximately ten-fold increase in mass of the fluorescent peptide after becoming ubiquitinated, we predicted that fluorescence polarization could be used to track ubiquitination. Fluorescence polarization (FP) is a property that varies directly with the size of a fluorophore in solution; as a larger molecule, tumbling more slowly, will exhibit higher FP and a smaller, more rapidly tumbling molecule, will exhibit lower FP. As expected, when the SdeA-catalyzed ubiquitination reaction was conducted, a significant increase in FP was observed that eventually reached a plateau (Figure 2-11A). In the absence of NAD^+, FP remained at steady levels. Without the need to use gel-based methods, an assay was developed to monitor SdeA-catalyzed ubiquitination, applicable for further studies of this process.

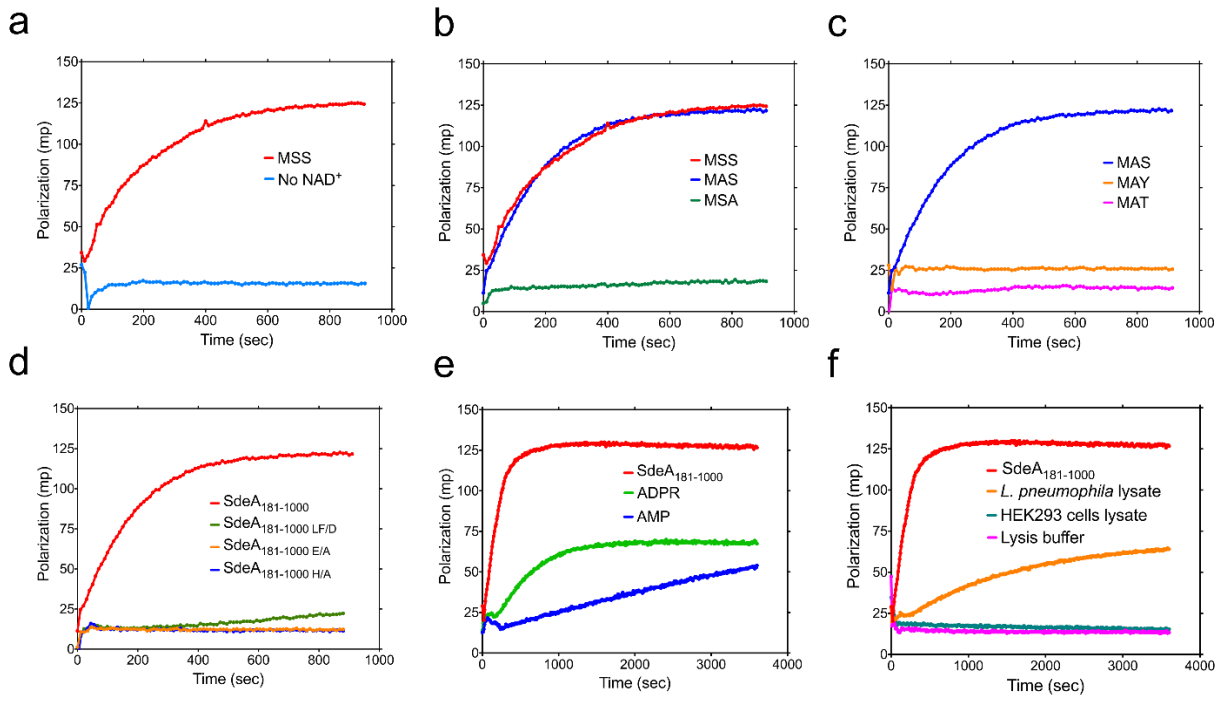


Figure 2-11: Fluorescence polarization assays of SdeA-catalyzed ubiquitination of the synthetic peptide. A) MSS peptide compared to a control reaction lacking NAD^+. B) Comparison of MSS, MAS, and MSA peptides as substrates. C) Probing of non-serine hydroxyl residues as substrates. D) Analysis of the activity of SdeA mutants. E) The effect of small-molecule inhibitors on SdeA activity. F) Probing of cell lysates using FP assay.

2.3.5 Probing Substrate Specificity of SdeA

In order to further explore the types of protein targets recognized by SidE enzymes, we generated new peptides that varied from the original MSSMNPEYD sequence (hereafter referred

to as MSS). Specifically, due two the presence of two Ser residues at positions 2 and 3, to determine which Ser was the major target, two variant peptides were synthesized where either Ser was replaced with Ala (MAS and MSA). When utilized in our FP assay, we found that while MAS retained activity similar to that of the original MSS peptide, MSA did not react (Figure 2-11B). This result suggested that while SdeA appeared to be able to ubiquitinate unstructured peptides, it prefers certain positions of the target Ser residue over others.

To test whether other hydroxyl-containing residues could be targeted, two more variant peptides were synthesized which possessed either Thr or Tyr at position 3 (MAT and MAY, respectively). These peptides were found to not react with SdeA in our FP assay, suggesting a strict specificity for Ser residues (Figure 2-11C).

2.3.6 Mutational Analysis

To determine whether this assay could be utilized in probing different mutants of SdeA, we generated three mutant constructs: $SdeA^{E/A}$, $SdeA^{H277A}$, and a mutant of SdeA where interdomain contacts between mART and PDE are disrupted by replacing two key hydrophobic residues at the interface with charged Asp residues ($SdeA^{LF/D}$). In our assay system, the catalytic mutants tested displayed no activity, and the LF/D mutant was impaired significantly (Figure 2-11D).

2.3.7 Inhibition Assays and Cell Lysate Analysis

Two important questions pertaining to this newly discovered enzyme family are (1) whether or not similar enzymes are found in other organisms and (2) whether they may be able to serve as drug targets. To study the first question, we determined the applicability of our assay to testing cell lysates. Indeed, we found that incubation with *L. pneumophila* lysate in place of enzyme caused an increase in FP, while incubation with HEK293 cell lysate did not cause an increase (Figure 2-11F).

To address the second question, we assessed whether our assay could be used to test inhibitors. Previously, AMP had been determined to inhibit the ubiquitination reaction.[69] Therefore, we incubated SdeA with AMP as well as ADPR with the expectation that these molecules would compete for NAD^+ binding (Figure 2-11E). Interestingly, we found that both of

these molecules caused inhibition of activity, with ADPR causing FP to plateau and AMP causing the rate of FP increase to significantly slow down.

2.3.8 Kinetic Properties of SdeA Towards Peptide Substrate

To gain further insight into the ability of SdeA to recognize the synthetic peptide as a substrate, a Michaelis-Menten kinetic analysis was performed with respect to the peptide by conducting reactions where [peptide] was varied, but not NAD^+ or Ub (Figure 2-12A). From these experiments it was determined that the Michaelis constant K_M was 79.37 μM and the turnover number k_{cat} was 1.63 1/s. While the data fit the Michaelis-Menten equation well, the curve was not fully saturated at the highest peptide concentration used (40 μM), indicating that peptide binding may be relatively weak (Figure 2-12C).

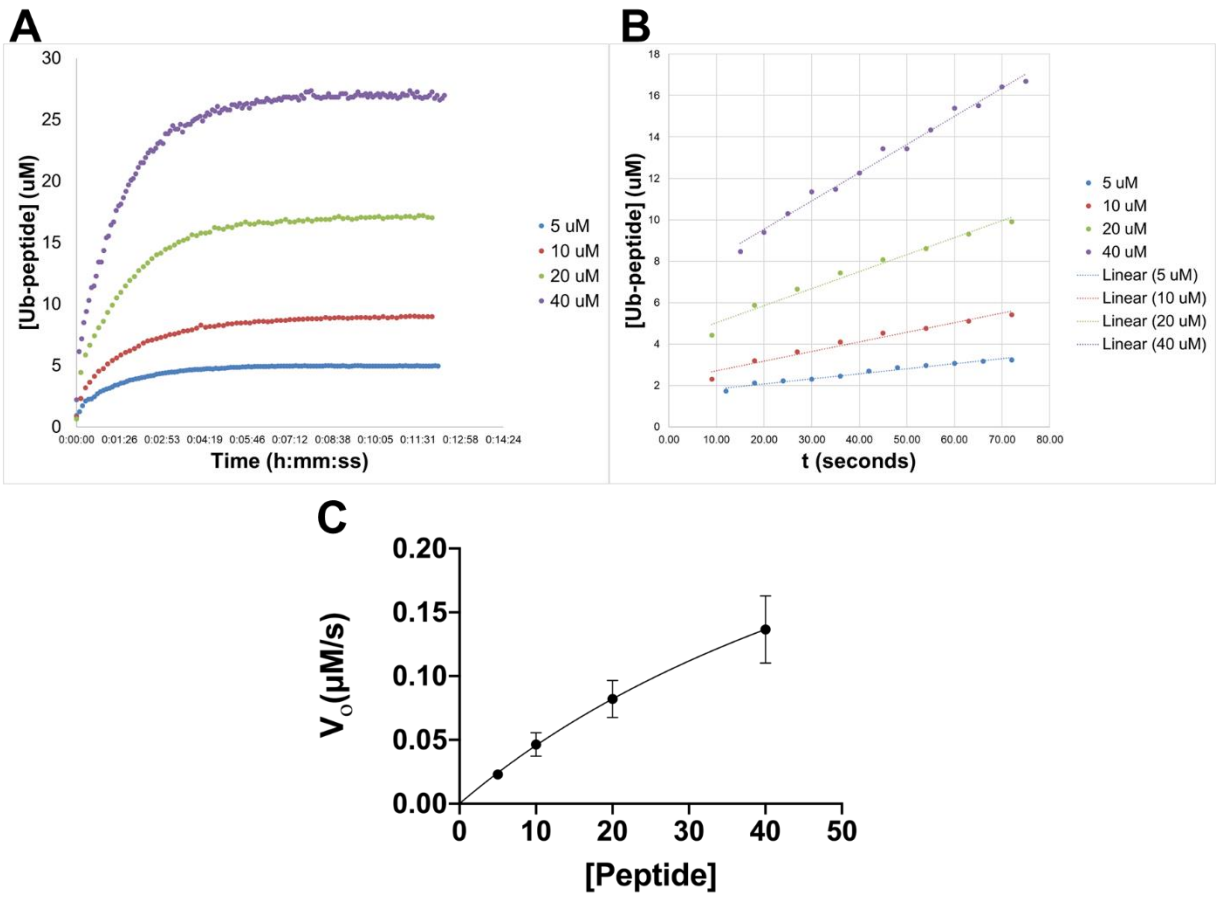


Figure 2-12: Michaelis-Menten analysis of SdeA-catalyzed ubiquitination with respect to concentration of peptide. A) FP progress curves of reactions conducted with varying amounts of peptide, normalized to reflect peptide concentration. B) Initial segments of curves fit to a linear model. The slopes of these curves were used to determine initial velocity. C) Michaelis-Menten curve with initial velocity plotted against peptide concentration.

2.4 Discussion

2.4.1 Overall Significance

Two major insights were gained from the results of these described experiments: first, that SdeA is able to attack Ub chains, both in terms of modification and transfer to protein substrates. This enzymatic modification renders the targeted Ub chains to be resistant to most host DUBs. These observations have increased the known breadth of the SidE family's enzyme function.

Stabilization of Ub chains as well as crosslinking them to protein targets brings up intriguing biological questions. As of now, the extent to which this occurs in a biological situation is not yet clear. More puzzling still is the fact that the full length SdeA possesses a DUB domain itself, which is also impaired in cleaving modified Ub chains. Further studies will be required to work out the extent and the effect of these modifications in a biological setting.

The other key advance was the development of an assay able to continuously monitor this noncanonical ubiquitination process. This assay represents a significant improvement over the gel-based methods used previously and has allowed us to better understand substrate recognition of SidE enzymes. An interesting application of this assay in future studies would be to compare the four members of the SidE family: SdeA, SdeB, SdeC, and SidE. Broadly speaking, while all four are capable of performing ubiquitination of substrate proteins, a more careful analysis comparing their kinetics as well as any variations or similarities in substrate preference has not yet been reported. Furthermore, this assay is performed in a 96-well plate format that opens the door to more high-throughput methods of inhibitor screening and testing for other enzymes in this family.

2.4.2 Insights into SdeA Substrate Recognition

The fact that SdeA recognizes ubiquitin chains is perhaps to be expected, due to the abundance of Ub chains present in the host cell, which have a myriad of signaling functions. Intriguingly, these results may yield insights into Ub binding by SdeA. We found that M1-linked diUb appeared to be not fully converted to the doubly phosphoribosylated species by $SdeA^{181\text{-}1000}$, in contrast to the other linkages tested. When mutant versions of this M1 diUb were tested, it was found that the R42K mutant, which left only the Arg of the proximal Ub as a target residue, underwent partial phosphoribosylation similarly to the WT M1 diUb. In all cases, no difference was observed in ADP-ribosylation levels, as they were fully converted within the parameters of our assay. Taken together, this data suggests that the proximal Ub moiety of M1-linked diUb may not be as well recognized by the PDE domain of SdeA, and that the area near the N-terminus of Ub may be important for PDE binding.

The fluorescent assay also yielded important insights on the nature of substrate recognition from the perspective of the protein target. We have learned that other hydroxyl-containing residues such as Thr and Tyr are not targeted for ubiquitination. Also, while unstructured regions bearing Ser residues may be ubiquitinated, the positioning of the Ser residue is nevertheless important.

Other studies also reached similar conclusions. It remains to be determined whether SidE proteins recognize a specific protein fold, or whether they are more nonspecific in their activity. A more unbiased approach to identifying PR-linked ubiquitinated proteins in an infection setting will be necessary to address this question.

2.4.3 Applications to Inhibitor and Enzyme Discovery Efforts

The SidE family is required for full virulence of *Legionella pneumophila* towards model hosts. Furthermore, it is toxic to eukaryotic cells when ectopically expressed. Due to this biological importance, this novel class of enzymes may represent new drug targets for inhibition and antibiotic development. The assays developed in this work may serve as important platforms for future inhibitor discovery efforts. Two potential inhibitors were thus evaluated: ADPR and AMP due to their general resemblance to the NAD^+ co-substrate. AMP appeared to generally slow the rate of reaction, where ADPR caused the FP value to plateau. We speculate that while AMP might be acting as a traditional competitive inhibitor, the plateau effect observed when ADPR is included in the reaction may be due to ADPR acting as a substrate in a transfer reaction to the fluorescent peptide. A peptide modified in such as way would no longer be able to become ubiquitinated and as such, the maximum FP reached would be lowered.

The fluorescent probes generated in this work have further applicability as tools for enzyme characterization. We have purified fluorescent Ub-εADPR, generated from SdeA mART domain reacting with Ub and εNAD^+. This may be used in future studies as a reagent to measure the activity of the PDE domain of SdeA independently, as well as for other enzymes such as phosphatases and ADP-ribosylhydrolases. The FP-based assay may also be utilized for probing lysates of species besides *L. pneumophila*. This would allow us to experimentally determine the presence of SidE-like enzymes elsewhere in nature that may not appear in bioinformatic searches, perhaps due to weak sequence similarity.

CHAPTER 3. BIOCHEMICAL STUDIES OF TRANGLUTAMINASE-LIKE UBIQUITIN LIGASE MAVC

A portion of this chapter is being considered for publication in *Nature Communications*.

3.1 Introduction

The discovery of a third noncanonical ubiquitination system, also catalyzed by a bacterial effector, further redefined our understanding of the manipulation of ubiquitin signaling by pathogens.[72] In contrast to both the E1/E2/E3 and the SidE mechanisms, MavC requires no cofactor and covalently links Gln40 of Ub with Lys92 (or to a minor extent Lys94) of target protein Ube2N in what appears to be a transglutaminase-like activity. This is an all-in-one ubiquitinating enzyme, but unlike the SidE family, which can ubiquitinate a variety of protein substrates, MavC appears to have a stark preference for Ube2N. Gan et al. tested other structurally similar E2 enzymes, but no activity was detected. This raised an important question of how MavC binds to and recognizes Ube2N. Also, structural analysis of the apo form of MavC revealed a core fold similar to that of the Cif family of bacterial effectors, which are known to catalyze the deamidation of both Ub and the structurally related NEDD8 at Gln40, the same residue attacked by MavC.[76] However, Cif and its related proteins distinctly lack the ubiquitinating activity. A key structural difference between Cif enzymes and MavC is the presence of an insertion between residues 128 and 226 that folds into a separate domain, hereafter referred to as the insertion domain (Figure 3-1). Valleau and colleagues observed that this insertion domain was necessary for Ube2N binding. Here, we further characterize the importance of the insertion domain on overall ubiquitinating and ubiquitin deamidating activity, and conduct kinetic studies of Ube2N and Ub binding to MavC, revealing that Ub binding is very weak. We further explore the possible manner by which MavC can overcome this weak Ub binding and observe that a mimic of Ube2N charged with Ub by E1 is a significantly better substrate than free Ub and free Ube2N. These experiments taken together suggest that MavC attacks the charged Ube2N~Ub thioester-linked conjugate, which is widely available in host cells, and catalyzes an intramolecular transglutamination reaction. The insertion domain likely drives interactions by binding to the Ube2N moiety, with Ub binding playing a more minor or transient role. This ubiquitinated Ube2N species (Ub-Ube2N) is rendered inactive, unable

to be charged by E1, likely due to a steric clash considering the proximity of Lys92 of Ube2N with the catalytic Cys87. Furthermore, Ube2N binds very tightly to its heterodimeric binding partner Uev1A, allowing for the formation of K63-linked Ub chains. Here we also show that Uev1A binding does not interfere with MavC-catalyzed ubiquitination, suggesting the physiological substrate of MavC is the Uev1A/Ube2N~Ub complex.

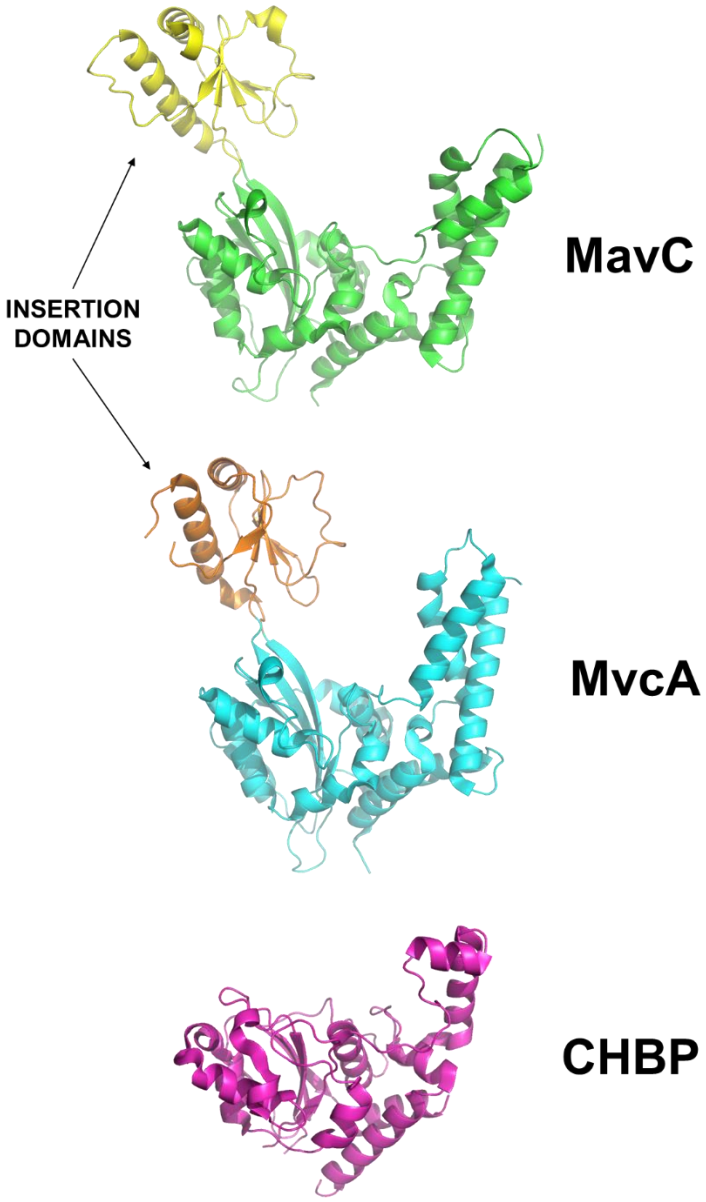


Figure 3-1: Structural comparison of MavC, MvcA, and Cif family member CHBP (PDB 5TSC,5SUJ, 3EIT respectively). Insertion domains of MavC and MvcA are highlighted.

3.2 Materials and Methods

3.2.1 Design of Protein Constructs

The stable 1-462 construct of MavC ($MavC^{1-462}$) was utilized for ubiquitination assays and is equally as active as the full-length construct (see Chapter 4). A construct of MavC lacking the insertion domain ($MavC^{\Delta INS}$) was designed with residues 129-225 deleted, and Cys226 changed to Gly to minimize undesired effects.

3.2.2 Cloning, Expression, and Purification of Recombinant Proteins

$MavC^{1-462}$ was cloned into the pGEX-6P1 vector as an N-terminal GST fusion protein. Briefly, the gene fragment for $MacC^{1-462}$ was amplified from full length MavC in pQE30 (received from Luo Lab) using standard, PCR-based techniques. A BamHI and XhoI restriction site was included on the respective 5' and 3' ends of this gene fragment. After restriction digest, T4 ligase was used to insert the gene into the pGEX 6P-1 vector.

$MavC^{\Delta INS}$ in the pGEX 6P-1 vector was obtained from GenScript (Piscataway, NJ). Ube2N in the pET-SUMO vector, as an N-terminal fusion to His-tagged SUMO protein was obtained from Addgene (#51131). Ubiquitin (Ub) was cloned into the pRSET-A vector. Ub G76C and Q40E mutations were made utilizing standard site-directed mutagenesis techniques. Mutants were confirmed via DNA sequencing. All protein constructs described were transformed into *E. coli* BL-21 DE3 strain for expression except for wild-type Ub which was transformed into the Rosetta expression strain of *E. coli*.

GST-tagged protein constructs and untagged Ub and Ub mutants were purified as described in Chapter 2.

For purification of Ube2N, *E. coli* cells containing the recombinant construct were used to inoculate 100 mL of LB media supplemented with ampicillin and grown overnight at 37º C under shaking conditions. This small-scale culture was used to inoculate up to 6 L of LB media supplemented with ampicillin. Cultures were grown at 37 ºC under shaking conditions until they reached an OD_{600} of ~0.4-0.6. At this point, induction of protein expression was performed by adding IPTG to a final concentration of 300 µM. Induced cultures were allowed to continue shaking at a decreased temperature of 18 ºC for 16-18 hours.

Cells were then harvested by centrifugation and resuspended in buffer (1x PBS, 400 mM KCl). This cellular suspension was then lysed by at least two passes through a French Press and then clarified by ultracentrifugation at 100,000 xg for 1 hour at 4 °C. The supernatant was decanted from the pellet and applied to Ni-NTA resin (Qiagen). The resin was washed with at least 5 column volumes of resuspension buffer, followed by a wash with resuspension buffer containing 50 mM imidazole to remove nonspecifically bound proteins from the resin. Finally, proteins were eluted by the addition of elution buffer (1x PBS, 400 mM KCl, 300 mM imidazole) and collected in four 10 mL fractions. His-SUMO was removed by addition of SENP2 and imidazole was removed by dialysis, conducted overnight into fresh resuspension buffer. Free SUMO was removed by adding the dialyzed and cleaved sample through the Ni-NTA resin once again. Flow-through contained purified Ube2N which was pooled and concentrated.

3.2.3 Generation and Purification of Ubiquitinated Ube2N Product

To generate ubiquitinated Ube2N (Ub-Ube2N), MavC 1-462 at a final concentration of 1 µM was incubated with purified Ube2N at a final concentration of 25 µM and Ub at a concentration of 100 µM. Reaction was allowed to proceed at a final volume of 30 mL for 3 hr at 37 ºC, in a buffer of 50 mM Tris pH 7.4, 100 mM NaCl and 1 mM DTT. Reaction mixture was concentrated to 1 mL and passed through two Superdex 200 Increase 10/300 GL columns (GE Healthcare), connected in tandem, in the reaction buffer described above. Fractions corresponding to pure Ub-Ube2N were pooled and concentrated. Side fractions containing a mixture of Ube2N and Ube2N were also collected and subjected to a second round of size-exclusion chromatography for further purification.

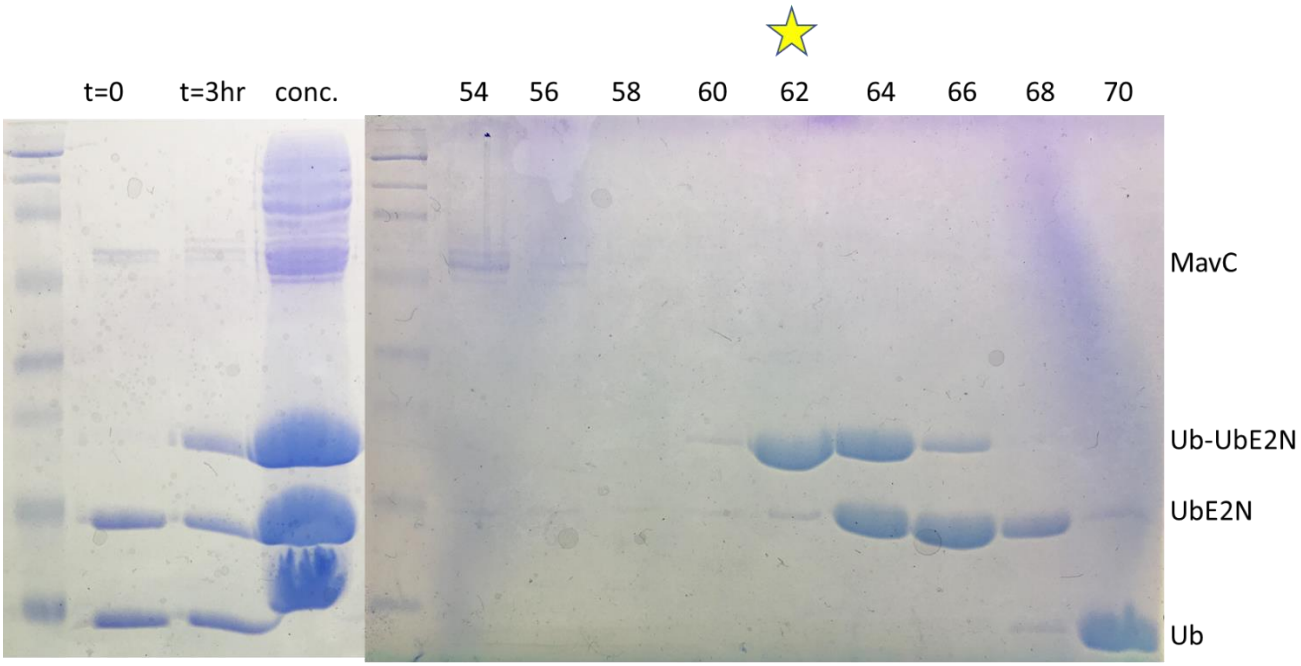


Figure 3-2: Synthesis and purification of Ub-Ube2N. Left panel shows reaction of Ub and Ube2N with MavC. Right panel shows size-exclusion fractions. Collected fraction indicated with a star.

3.2.4 Generation and Purification of Disulfide-Linked Mimic of Ubiquitin-Charged Ube2N

To generate a disulfide-linked mimic of charged Ube2N (Ub-SS-Ube2N), activated Ub as first prepared as previously described, beginning by reaction of the G76C mutant of Ub with 5,5´-dithio-bis-nitrobenzoic acid (DTNB).[86] Reaction mixture contained 250 µM of Ub G76C and 2 mM DTNB in 0.1 M sodium phosphate pH 8 as the buffer. This reaction was allowed to incubate overnight at 4 ºC to generate activated Ub G76C-TNB. A buffer exchange into non-reducing reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl) was then performed to remove excess TNB. To generate Ub-SS-Ube2N, an equimolar ratio of Ube2N and Ub G76-TNB was combined in non-reducing reaction buffer and allowed to incubate for 3 hours at 25 º under gentle rocking. Ub-SS-Ube2N. The reaction was then concentrated to 4 mL and subjected to size-exclusion chromatography using a Superdex S75 column (GE Healthcare) with non-reducing reaction buffer. Fractions containing pure Ub-SS-Ube2N were pooled and concentrated. Generation and purification of Ub-SS-Ube2N was performed with Kristos Negrón Terón and Dr. Shalini Iyer.

3.2.5 Ubiquitination Assays

To compare the ubiquitinating activity of MavC with MavC$^{\Delta INS}$, purified MavC constructs (wild type or mutants) were combined with Ube2N and Ub at a final concentration of 1 µM MavC, 25 µM Ube2N, and 100 µM Ub in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT). Reactions were conducted for 1 hour at 37 °C before quenching with 5X SDS-PAGE loading buffer. Reactions were analyzed by SDS-PAGE and visualized with Coomassie Blue.

To compare the ability of the disulfide-linked Ube2N-SS-Ub and the free Ube2N and Ub to become ubiquitinated by MavC, reactions were conducted at a final concentration of 0.005 µM MavC and either 25 µM Ube2N-SS-Ub or 25 µM Ube2N and 25 µM Ub. Reactions were incubated at 37 °C for up to 30 minutes in non-reducing reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl). The reaction products were analyzed by reducing SDS-PAGE and visualized with Coomassie Blue.

To determine the activity of MavC in the presence of the Uev1a-bound Ube2N-SS-Ub complex, varying amounts of Uev1a were incubated with Ube2N-SS-Ub for 10 minutes on ice in order to form the complex prior to addition of 0.005 µM MavC. Ubiquitination reactions were performed at 37 ºC for 30 minutes in non-reducing reaction buffer. Reactions were quenched with SDS-PAGE loading buffer, separated by SDS-PAGE and visualized with Coomassie Blue.

3.2.6 Deamidation Assays

Ubiquitin deamidating assays were performed by combining purified MavC$_{1\text{-}462}$ constructs (wild type or mutants) with Ub at a final concentration of 1 µM enzyme and 100 µM Ub. The reactions were incubated at 37 °C for 30 minutes in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT). The deamidation reaction products were analyzed by Native-PAGE and visualized with Coomassie Blue.

MavC's deamidating activity on the disulfide conjugate was tested by combining purified MavC$_{1\text{-}462}$ with Ube2N-SS-Ub at a final concentration of 0.005 µM MavC, 25 µM Ube2N-SS-Ub and incubated at 37 °C for 30 minutes in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl). The reaction products were analyzed by Native-PAGE and visualized with Coomassie Blue. As a control to observe the migration band of deamidated Ube2N-SS-Ub, a reaction was run utilizing the known deamidase Cif (obtained from the Luo Lab) at a concentration of 0.5 µM.

3.2.7 Kinetic Analysis of Ubiquitination and Deamidation

To determine the Michaelis-Menten kinetic parameters of the Ube2N ubiquitinating activity of MavC, reactions were conducted with MavC (0.5 μM), and varying concentrations of Ube2N at 37 ºC for 30 minutes. Reactions were quenched with SDS-PAGE loading buffer and separated by SDS-PAGE along with Ube2N standards of known concentrations and visualized with Coomassie Blue. Gels were analyzed with ImageJ, and a standard curve was generated using the band intensities of the Ube2N standards. This standard curve was used to quantify Ube2N remaining after each reaction. Data was fit to the Michaelis-Menten equation. Linear regression and plotting were performed using SigmaPlot.

To determine the Michaelis-Menten kinetic parameters of the Ub deamidating activity of MavC, reactions were conducted with MavC (0.5 μM), and varying concentrations of Ub at 37 ºC for 30 minutes. Reactions along with Ub Q40E standards of known concentrations were separated by Native-PAGE. Gels were analyzed by ImageJ as described above. All reactions were performed in triplicate for kinetic analysis.

3.2.8 E1 Charging Assay

To compare the ability of E1 to charge Ube2N versus Ub-Ube2N, a reaction mixture of 0.5 uM E1 enzyme, 200 uM Ube2N or Ub-Ube2N, 400 uM Ub was conducted in a reaction buffer consisting of 50 mM Tris pH 7.4, 100 mM NaCl, 2.5 mM ATP, 5 mM MgCl2. Reactions were allowed to proceed for 30 minutes at 37 ºC before quenching with either reducing or non-reducing SDS-PAGE buffer, separated by SDS-PAGE and visualized with Coomassie Blue.

3.3 Results

3.3.1 Effect of Insertion Domain on MavC Activity

Previous studies suggested that the insertion domain was important for Ube2N binding by MavC.[76] To further analyze the role played by the INS domain, especially on the ubiquitination activity, the MavC$^{\Delta INS}$ construct was used to attempt a ubiquitination reaction (Figure 3-3). As a negative control, the catalytically inactive MavC C74A mutant was used, and wild type MavC was

used as a positive control. Under these assay conditions, MavC$^{\Delta INS}$ was found to be totally inactive in ubiquitinating Ube2N.

This construct MavC$^{\Delta INS}$ was also tested in a deamidation reaction, once again compared alongside wild type and C74A mutant of MavC. Interestingly, while deamidating activity of MavC$^{\Delta INS}$ was reduced in comparison to the wild type, it was not totally abolished, as was the case with the ubiquitination experiment described above. This suggested that the insertion domain played a critical role in Ube2N recognition and a relatively minor role in Ub recognition.

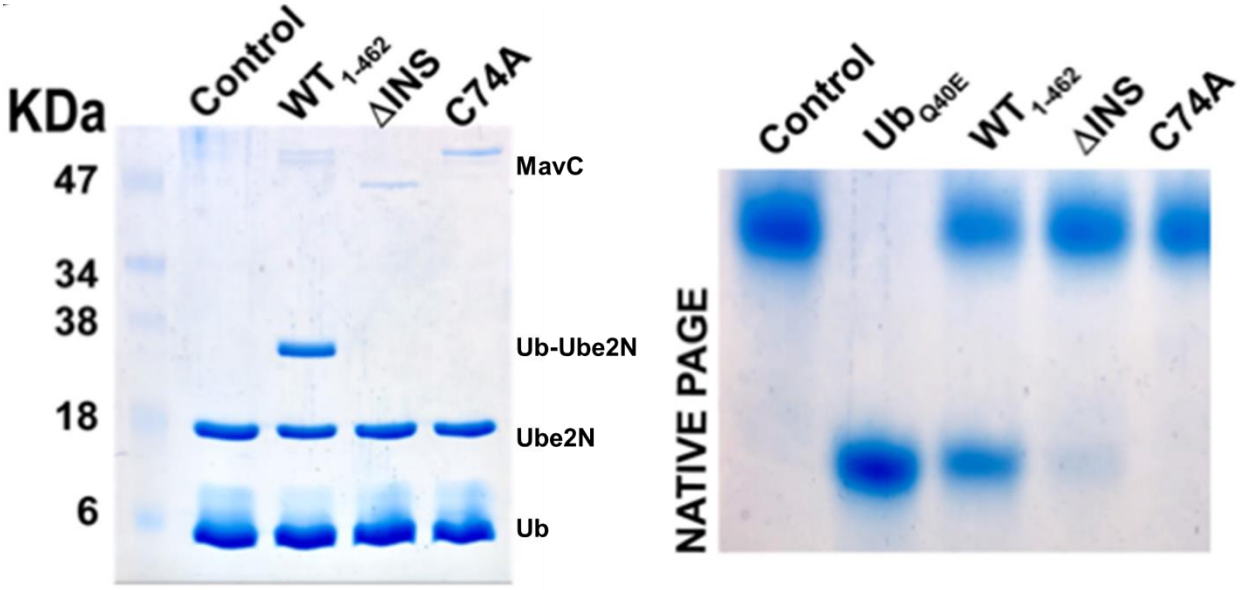


Figure 3-3: Analysis of insertion domain on ubiquitination and deamidation. Left panel: SDS-PAGE of ubiquitination reaction. Right panel: Native-PAGE of deamidation reaction.

3.3.2 Kinetic Analysis of MavC-Catalyzed Ubiquitination and Deamidation

To better understand the substrate recognition and catalytic abilities of MavC, a Michaelis-Menten-type kinetic analysis was performed of the MavC-catalyzed ubiquitination reaction with respect to Ube2N, and of the deamidation reaction with respect to Ub (Figure 3-4). Ube2N binding appeared to be fairly robust, with a K_M of ubiquitination approximately 10.4 µM (Table 3). However, in stark contrast, Ub binding was found to be very weak, with an estimated K_M determined to be over 500 µM. Indeed, the deamidation reaction failed to be saturated even at the highest [Ub] of 375 µM was used in the reaction (Figure 3-4B). Taken together, it was evident that

Ube2N binding was significantly stronger than Ub binding. Considering this weak Ub binding and the fact that only 100 µM Ub was utilized in the ubiquitination reactions in this analysis, it is probable that saturating Michaelis-Menten conditions were not met, so the calculated K_M of Ube2N is likely an overestimation with respect to the actual Michaelis constant.

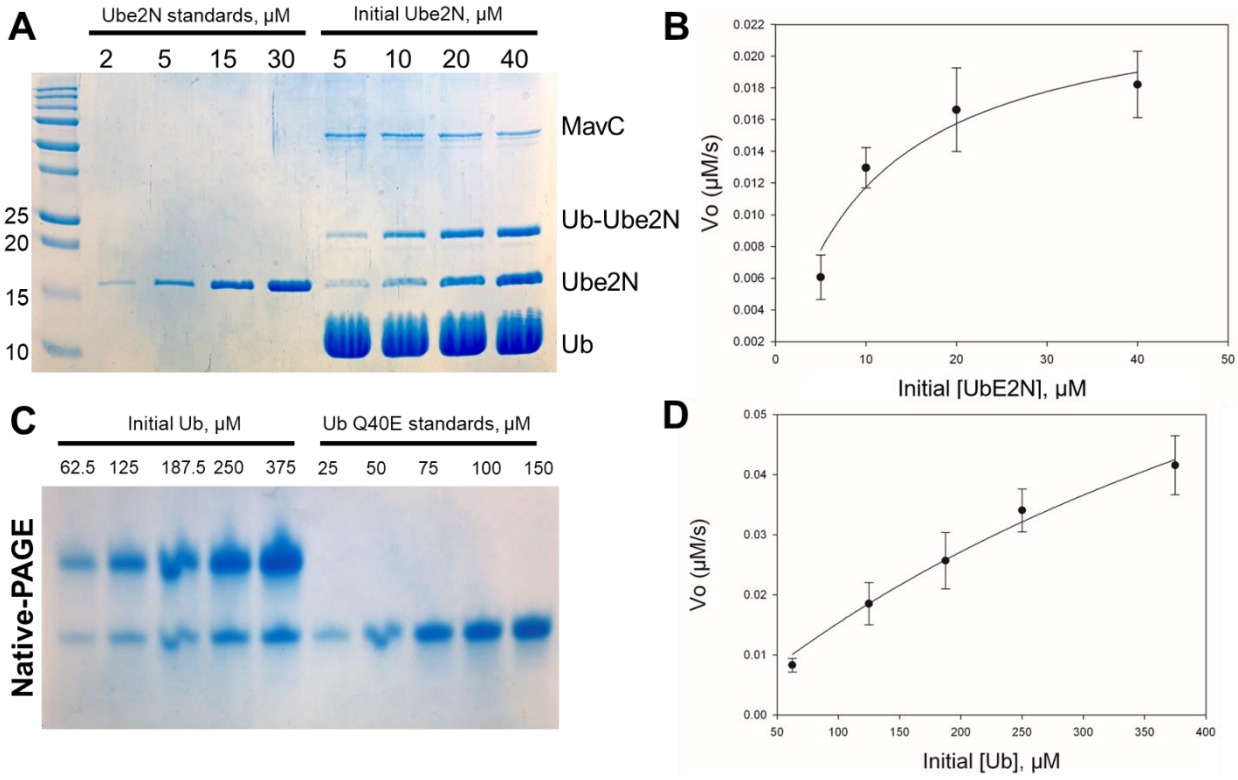


Figure 3-4: Kinetic analysis. A) SDS-PAGE gel of ubiquitination reactions with standards. B) Plots of initial velocity versus [Ube2N] fit to the Michaelis-Menten equation. C) Native PAGE gel of deamidation reactions with standards. B) Plots of initial velocity versus [Ub] fit to the Michaelis-Menten equation.

Table 3-1: Calculated Michaelis-Menten parameters for ubiquitination and deamidation reactions.

	Ubiquitination	Deamidation
K_M (µM)	10.40 ± 4.29	686.27 ± 235.49
k_{cat} (1/s)	0.0239 ± 0.0036	0.1203 ± 0.030

3.3.3 Comparison of Ub-SS-Ube2N and free Ub/Ube2N as MavC Substrates

This weak affinity between MavC and Ub was somewhat counterintuitive, as it suggested that MavC would be unlikely to catalyze transglutamination in any biologically relevant capacity under intracellular conditions, as the concentration of free Ub in human cells has been determined in previous quantitative studies to be around 20 μM.[79] However, it had been readily observed that Ub-Ube2N product does form in cells appreciable amounts under *Legionella* infection conditions. Considering that most E2s in cells exist in a ubiquitin-charged state, where Gly76 of Ub is bound via a thioester linkage to the catalytic Cys of the E2 (in the case of Ube2N, Cys87),[87,88] we hypothesized that MavC binds Ub poorly because it instead recognizes the charged state of Ube2N (Ube2N~Ub) and catalyzes an intramolecular transglutamination reaction to crosslink Ub and Ube2N with a much more stable isopeptide bond. To test this hypothesis, due to the labile nature of Ube2N~Ub, a more stable analog was generated by engineering a disulfide bond between residue 76 of Ub (by mutating Gly76 to Cys) and residue 87 of Ube2N (Figure 3-5).

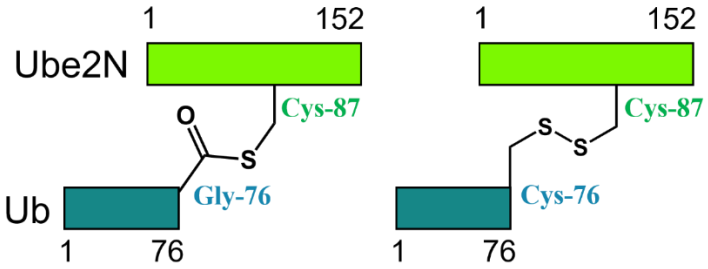


Figure 3-5: Comparison of Ube2N charged with Ub (left), and the disulfide conjugate Ub-SS-Ube2N utilized in this study (right).

This disulfide conjugate, or Ube2N-SS-Ub, was purified and its reactivity toward MavC was compared with reactivity of equimolar amounts of free Ube2N and free Ub toward MavC. Strikingly, Ube2N-SS-Ub proved to be a tremendously more reactive substrate than the free species. Even when a catalytic amount of MavC was utilized (5 nM as compared to 1 μM used in previous assays), nearly complete conversion of Ube2N-SS-Ub into the isopeptide-linked product was observed within 30 minutes, whereas at that time point no reactivity at all was observed with free Ube2N and Ub (Figure 3-6). This result strongly suggests that under cellular conditions, MavC targets Ube2N~Ub, crosslinking them via an isoopeptide bond. This result is also in agreement with an observation from a previous study, where a mutant of Ub lacking the two C-terminal Gly

residues and therefore unable to become charged by Ube2N was modified significantly less than wild-type Ub.

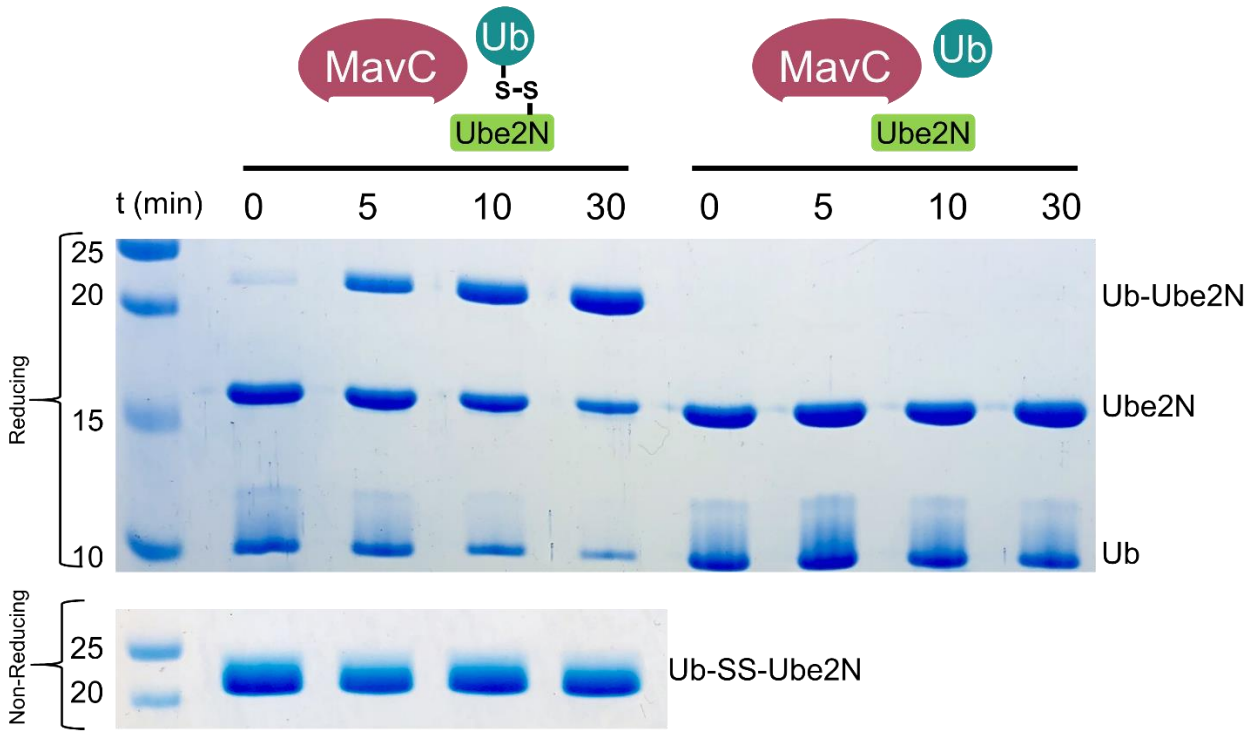


Figure 3-6: Ubiquitination assay to compare Ub-SS-Ube2N vs free Ube2N and Ub as MavC substrates. A non-reducing SDS-PAGE control to show the stability of Ub-SS-Ube2N over time course of this assay is also included.

To additionally determine whether deamidation of the Ub moeity from Ube2N~Ub was a relevant process, the disulfide conjugate was incubated with MavC under the same assay conditions where nearly complete conversion to crosslinked product was observed. Within this assay timescale and conditions, no deamidated Ube2N-SS-Ub was detected (Figure 3-7), suggesting that deamidation is unlikely to play a significant role in MavC's action as an effector and that ubiquitination is instead the dominant reaction.

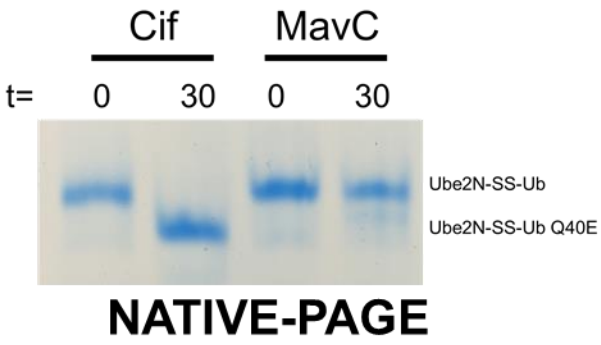


Figure 3-7: Native PAGE assay of Ub-SS-Ube2N to detect deamidation of the Ub moiety. As a control to determine the migration pattern of deamidated Ube2N-SS-Ub, a reaction with the known deamidase Cif was included.

3.3.4 Effect of Uev1A on MavC Activity

Since Ube2N exists in cells as part of a tight heterodimeric complex with Uev1A,[89] we determined whether this complex was also able to be ubiquitinated by MavC. As such, Ube2N-SS-Ub was subjected to pre-incubation with increasing concentrations of Uev1A and then reacted with MavC. The formation of Ub-Ube2N product in the presence of the highest concentration of Uev1A was comparable to when the reaction was conducted without Uev1A, suggesting that Uev1A binding did not affect MavC activity (Figure 3-8).

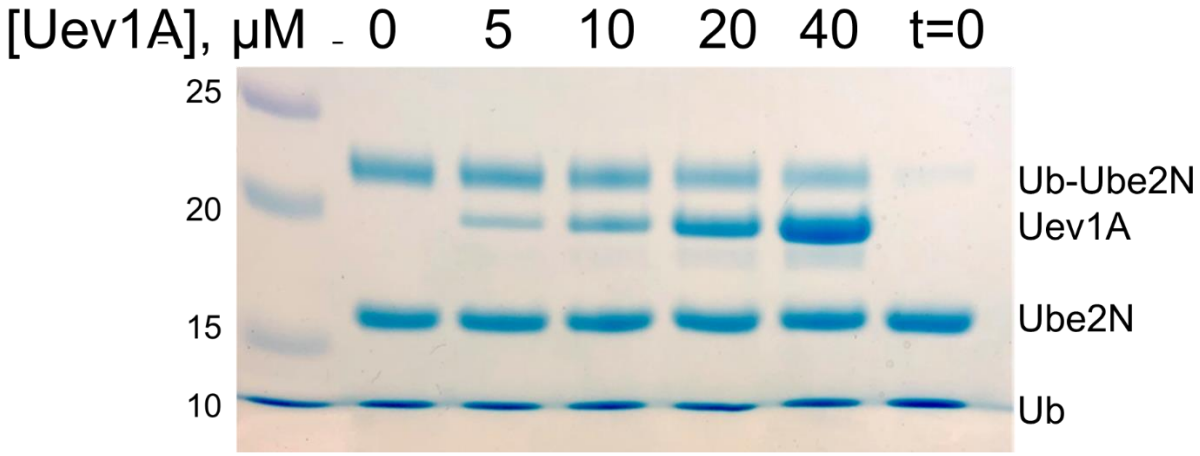


Figure 3-8: Analysis of Uev1A binding on Ube2N activity. Increasing concentrations of Uev1A were added to Ub-SS-Ube2N before conducting ubiquitination reaction.

3.3.5 E1 Charging of Ube2N Versus Ub-Ube2N

Previously, it had been shown that Ub-Ube2N, the product of MavC-catalyzed ubiquitination, was unable to form K63-linked ubiquitin chains. To further explore the basis of this loss in activity, we attempted to charge the Ub-Ube2N by incubation with E1, Ub, and ATP. While the control unmodified Ube2N was robustly charged, no charging of Ub-Ube2N was observed to any extent (Figure 3-9). This suggests that the ubiquitination of Ube2N by MavC halts the crucial charging step of canonical ubiquitination, thereby blocking Ube2N activity.

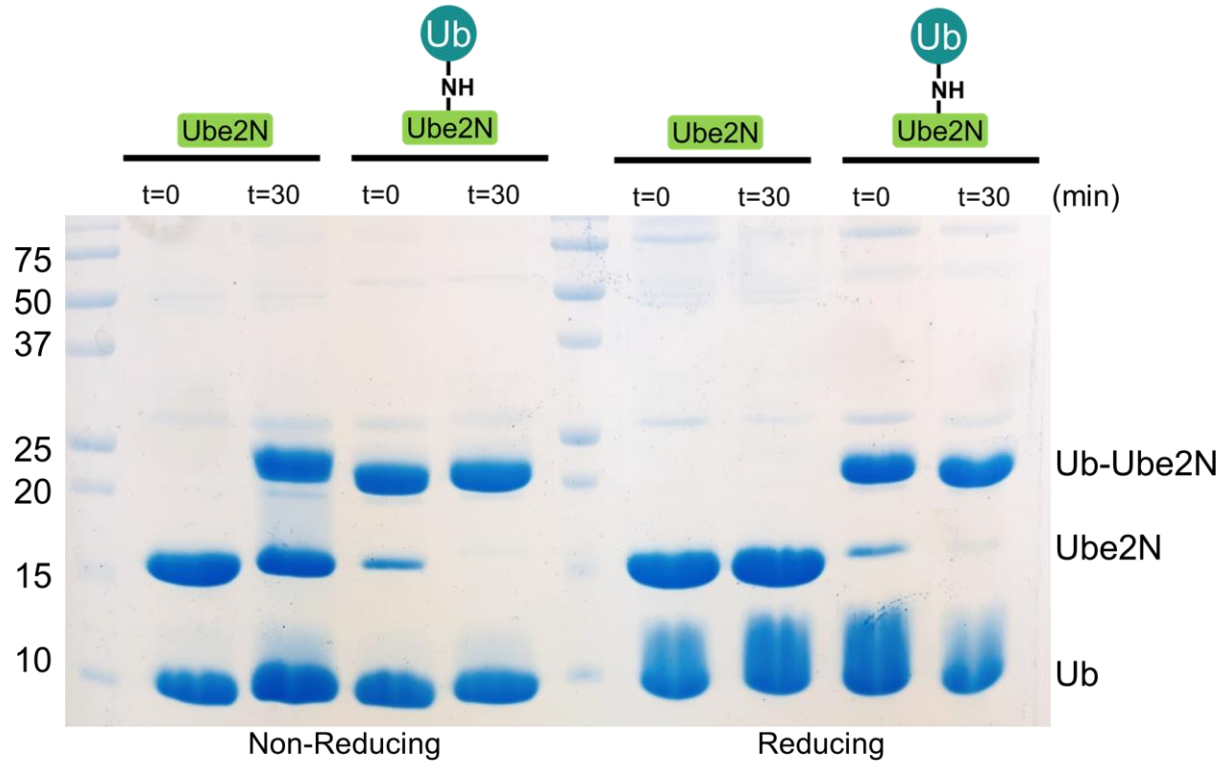


Figure 3-9: E1 charging assay. Formation of the charged conjugate is detected via non-reducing SDS-PAGE (left lanes).

3.4 Discussion

3.4.1 Overall Significance

The biochemical studies described here showcase the importance of the insertion domain to the ubiquitinating activity of MavC. Furthermore, kinetic analyses comparing the ubiquitinating and the deamidating reactions catalyzed by MavC revealed that free Ub binding was weak, likely too weak for adequate reaction in a cellular context. It was then shown that the Ube2N~Ub charged

conjugate is the likely target of MavC by comparing the action of this effector against a mimic of the conjugate with free Ube2N and Ub. The accommodation of Uev1A by MavC, an important binding partner of Ube2N was also shown. Finally, the effect of ubiquitination of Ube2N at the Lys92 site is explored, with Ub-Ube2N found to specifically be defective in the E1 charging step despite possessing a free catalytic Cys.

3.4.2 Role of Insertion Domain

The insertion domain defines a structurally new and unique Ube2N-binding motif and is the distinguishing feature of MavC and MvcA from the Cif family (Figure 3-1). Indeed, it is required for full MavC ubiquitinating, but not deamidating activity. The latter observation interestingly suggests that this domain may also play a role in recognizing ubiquitin as well. A DALI search of a model of the insertion domain alone revealed no clear structural homologs. Further adding to the unique nature of this domain is the fact that of the entire *Legionella* genus, only the *Legionella pneumophila* species possesses MavC, possibly showing that the acquisition of the insertion domain is a relatively recent evolutionary development.

3.4.3 The Physiologically Relevant Substrate of MavC

Instead of joining together free Ub and Ube2N, as had been previously demonstrated *in vitro*, MavC recognizes the thioester linked Ube2N~Ub conjugate. This is borne out by the fact that E2 enzymes typically exist as these charged conjugates in cells, and also due to the observation that Ub binding was very low (Figure 3-4D, Table 3). MavC was also initially reported to be a Ub-specific deamidase by Valleau et al,[76] but deamidated Ub was undetectable in cells infected with wild-type *L. pneumophila*.[72] The result from the kinetic analysis here provides a possible explanation for this. During an infection setting, a relatively small copy number of effectors will be deployed into the host cell. Therefore, in order to have a robust effect, it is especially critical for MavC to be able to efficiently recognize its protein substrate. By targeting Ube2N~Ub, MavC capitalizes on its relatively stronger interactions with the E2 subunit to bring in the adjacent Ub, thereby overcoming its low free Ub affinity. Additionally, having both the acyl acceptor ($Lys92^{Ube2N}$) and acyl donor ($Gln40^{Ub}$) units in one tethered molecule permits efficient capture of the Ub-thioester intermediate through a transamidation reaction rather than allowing a futile

reaction via hydrolysis (deamidation). Even using the disulfide substrate as a means to provide Ub to MavC at higher affinity, we found no detectable deamidation of Ub, whereas a Cif family enzyme can efficiently deamidate Ub in the same substrate (Figure 3-7). Altogether, the results indicate that the acquisition of the insertion domain by MavC and its evolution into a Ube2N-binding motif has shifted the balance in favor of the transglutamination reaction at the cost of Ub deamidation.

3.4.4 Conclusions and Further Directions

While MavC shares the core fold of the Cif family, our results show that it has diverged both structurally and functionally, having effectively lost the original function of Ub and NEDD8 deamidation through a lower Ub affinity (Figure 3-7). Instead, it preferentially attacks the Ube2N~Ub conjugate to turn off Ube2N's ability to generate K63 poly-Ub chains (Figure 3-6, 3-10). Despite both free Ub/NEDD8 deamidation and Ube2N ubiquitination ultimately being shown to lead to decrease in NF-κB activation, we speculate that the divergence of MavC occurred to accommodate other *L. pneumophila* effectors that utilize the host's free Ub such as the E3 ligases LegU1 and SidC and the noncanonical ligases of the SidE family.[37,90] Poisoning of the host cell's supply of Ub by deamidation could antagonize these other effectors' activity. Therefore, MavC may satisfy a need for an alternative method of attenuating host immune signaling without perturbing the free Ub pool.

The fact that Ub-Ube2N is catalytically switched off has importantly been shown to attenuate immune response of a human host due to the importance of Ube2N in the NF-κB signaling pathway (Figure 3-10). How hampering the host's ability to make K63 poly-Ub chains may lead to other cellular effects, however, remains to be elucidated. This is an especially relevant point, as effects besides immune signaling are attributed to K63 linked Ub chains, such as vesicular trafficking and autophagy. Moreover, it has been shown that much of the effector arsenal of *Legionella* has evolved from pathogenic interactions against host organisms other than humans. This is due to the fact that the *Legionella*'s natural hosts are amoebae and ciliated protozoa.[91–93] A careful analysis of the effect of MavC against these organisms may lead to an understanding of the pressures that drove its evolution, as well as broader insights into the role of Ube2N in non-immunological roles.

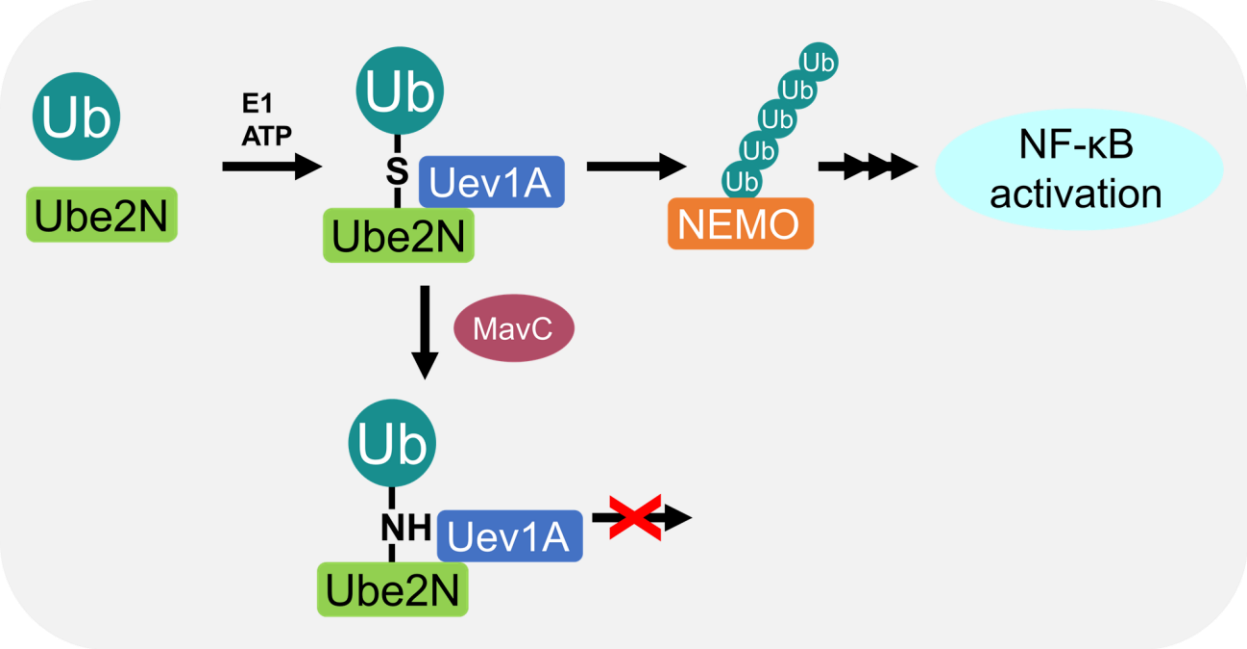


Figure 3-10: Proposed scheme of MavC's effect in host cells. The biochemical studies presented in this Chapter determine the likely target of MavC as the Uev1A:Ube2N~Ub complex, catalyzing an intramolecular crosslinking of Ub and Ube2N, preventing the ubiquitination of substrates such as NEMO that lead to downstream NF-κB activation.

CHAPTER 4. THE CRYSTAL STRUCTURE OF MAVC IN COMPLEX WITH SUBSTRATE AND PRODUCT

A portion of this chapter is being considered for publication in *Nature Communications*.

4.1 Introduction

The noncanonical ubiquitinating enzyme MavC, found within the *Legionella pneumophila* genome and injected as an effector into the host cell during infection, shares a core structural resemblance to the Cif family of enzymes. [72,76]This family consists of enzymes are also known to be bacterial effectors that attack host ubiquitin and/or NEDD8 via deamidating Gln40. The structural difference lies in the presence of an insertion domain found near the middle of MavC, and, as described in the previous chapter, is indispensable for ubiquitinating activity. This activity is also the key functional distinction between MavC and the Cif enzymes.

Notably, the mechanism of MavC-catalyzed ubiquitination is unique from both the E1/E2/E3 and the SidE pathways. This mechanism is reminiscent of transglutaminases, which also form isopeptide crosslinks between glutamine and lysine residues of their protein co-substrates. Also, similarly to the transglutaminase family, MavC uses a catalytic triad of residues for its activity, with a cysteine, Cys74, serving as the nucleophile. The first step consists of a nucleophilic attack of the catalytic Cys to the target Gln of one of the co-substrates, causing a release of ammonia and forming an obligate thioester. This thioester is then attacked by the ε-amino group of a Lys residue of the other protein substrate, ultimately forming the stable thioester linkage. In contrast to MavC, however, transglutaminases as they are currently understood are much less specific in their substrate recognition in comparison to MavC. A notable example of this broad substrate recognition is Factor XIII, an enzyme important in formation of blood clots.[94] Previous studies have shown that this transglutaminase has 147 different protein substrates.

An important question raised then is the mode of recognition of Ube2N~Ub by MavC. The biochemical evidence suggests that binding of Ube2N~Ub is likely driven by interactions between Ube2N and the insertion domain. The crystal structure of MavC had been resolved, revealing the presence and unique fold of this insertion domain, but it leaves open the question of how either Ub or Ube2N interact in order for catalysis to occur. Further, previously solved crystal structures of

Cif enzymes in complex with NEDD8 provide only limited information on substrate recognition due to MavC's selectivity for Ub over NEDD8.

MavC also strongly prefers to ubiquitinate Lys92 of Ube2N over the nearby Lys94. Indeed, mutation of Lys92 to Ala resulted in a drastic decrease in ability of Ube2N to become ubiquitinated *in vitro*. Conversely, the Lys94 to Ala mutant of Ube2N continued to become ubiquitinated at a comparable level to that of wild type Ube2N. This raises further questions as to how MavC is able to discern between these two nearby Lys residues.

To address the outlined questions, characterization of the structure of MavC in complex with its substrate as well as its product became the objective of these studies. In this Chapter, the cocrystal structures of three disparate crystal forms of MavC in complex with a disulfide-linked mimic of Ube2N~Ub representing the substrate-bound form along with the cocrystal structure of MavC in complex with the ubiquitinated, isopeptide-linked product Ub-Ube2N are presented. Analysis of these structures revealed the molecular basis of Ub and Ube2N recognition as well as a fascinating dynamic quality of the insertion domain which had not been previously appreciated. Further, it was observed that a significant remodeling of the Ube2N active site by MavC is required for ubiquitination.

4.2 Materials and Methods

4.2.1 Design of Protein Constructs

A construct of MavC spanning residues 1-384 ($MavC^{1-384}$) was utilized for crystallization experiments. Examination of the previously solved apo structure of MavC in the PDB (5TSC) showed that while the 1-462 construct was crystallized, only residues 1-384 were resolved in the structure.[76] This suggested that the C-terminal residues were likely disordered, therefore truncations in the C-terminus were made to minimize flexible regions that may hinder crystal packing (Figure 4-1 A). A ubiquitination assay showed that $MavC^{WT}$, $MavC^{1-462}$, $MavC^{8-400}$, and $MavC^{1-384}$ were equally active, indicating that the C-terminus likely plays a non-essential role in MavC's function, and may be used a translocation signal for the Legionella type IV secretion system (Figure 4-1B).

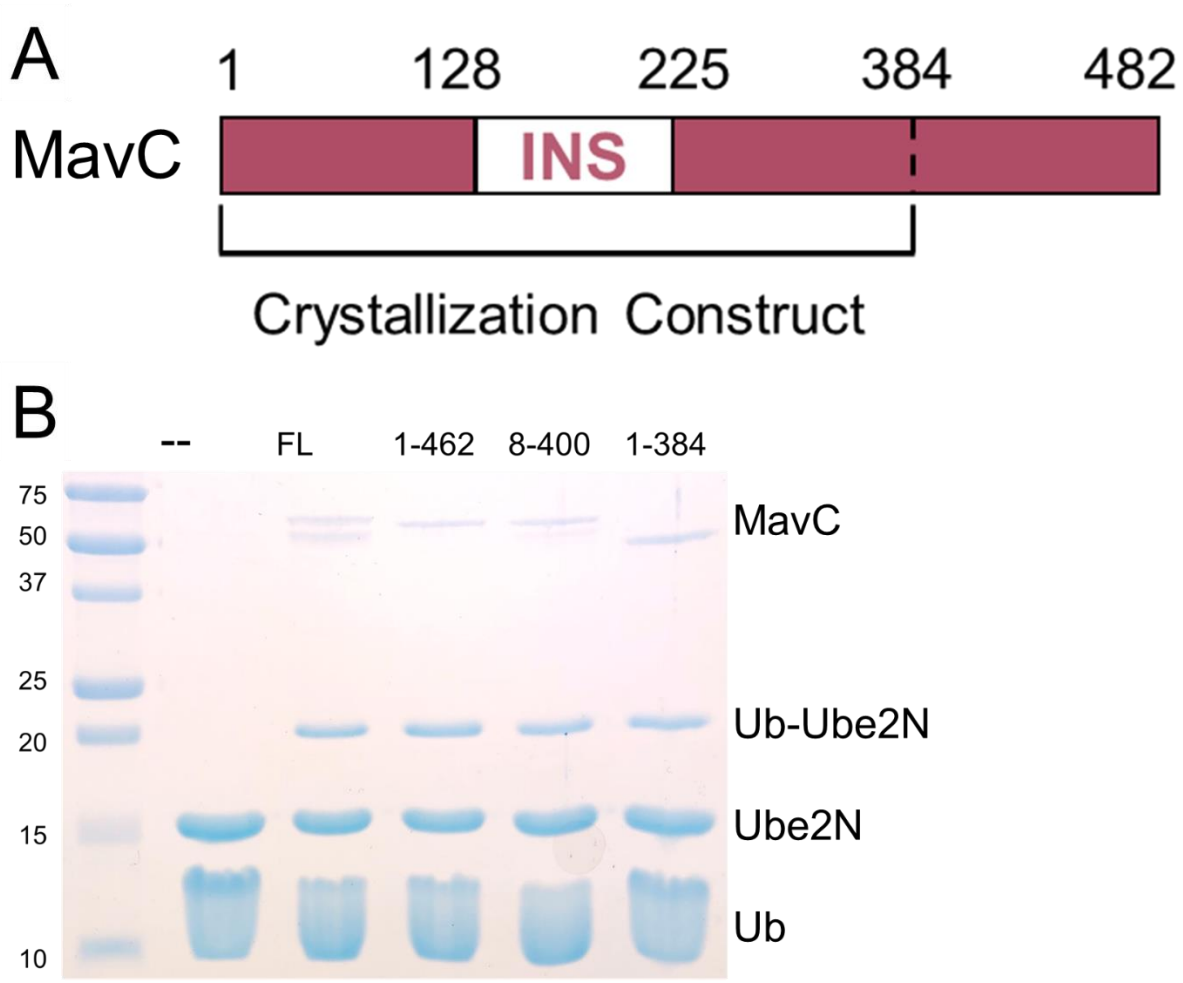


Figure 4-1: Constructs of MavC utilized. A) Domain diagram highlighting the 1-384 region used for crystallography. B) Ubiquitination assay comparing different MavC constructs.

In addition to the truncation, the catalytically inactive C74A mutant was introduced as to avoid reaction with Ub-SS-Ube2N as well as a precaution to avoid the possibility of nucleophilic attack toward Ub-Ube2N.

4.2.2 Cloning, Expression, and Purification of Recombinant Proteins

$MavC^{1-384}$ and $MavC^{8-400}$ were cloned into the pGEX 6P-1 vector as an N-terminal GST fusion construct. Briefly, the desired gene fragment was amplified from full length MavC in pQE30 (received from Luo Lab) using standard, PCR-based techniques. A BamHI and XhoI

restriction site was included on the respective 5' and 3' ends of this gene fragment. After restriction digest, T4 ligase was used to insert the gene into the pGEX 6P-1 vector.

Untagged wild-type ubiquitin was cloned into the pRSET-A vector. Ube2N in the pET-SUMO vector, as an N-terminal fusion to His-tagged SUMO protein was obtained from Addgene (#51131). All mutants described in this study were created using site-directed mutagenesis, with each mutant sequenced to confirm the presence of the mutation.

Recombinant protein constructs were transformed into *E. coli* BL-21 DE3 cells for protein expression with the exception of wild-type Ub, which was transformed into the *E. coli* Rosetta strain.

MavC and its mutants were expressed and purified via GST-affinity chromatography as has been described in Chapter 2. The GST tag was cleaved using PreScission™ Protease and subsequently removed from the solution by incubation with the Glutathione-Sepharose resin. MavC mutants were purified in conjunction with Dr. Shalini Iyer, Kristos Negrón-Terón, and Sebastian Kenny (Das Lab, Purdue University).

Ubiquitin was expressed and purified via cation-exchange chromatography as has been described in Chapter 2.

Ube2N was expressed and purified via nickel affinity chromatography as has been described in Chapter 3. The His-SUMO tag was cleaved using SENP2 and subsequently removed from the solution by incubation with the Ni-NTA resin.

4.2.3 Generation and Purification of Ubiquitinated Ube2N Product

Ubiquitinated Ube2N (Ub-Ube2N) was synthesized and purified as described in Chapter 3. Briefly, $MavC^{1-462}$ at a final concentration was incubated with purified Ube2N and Ub for 3 hr at 37 ºC. Components of the ubiquitination reaction were separated using size-exclusion chromatography. Fractions corresponding to Ub-Ube2N were pooled and concentrated.

4.2.4 Generation and Purification of Disulfide-Linked Mimic of Ubiquitin-Charged Ube2N

To generate a disulfide-linked mimic of charged Ube2N (Ub-SS-Ube2N), activated Ub was prepared as previously described, beginning by reaction of the G76C mutant of Ub with 5,5´-dithio-bis-nitrobenzoic acid (DTNB). Reaction mixture contained 250 µM of Ub G76C and 2 mM

DTNB in 0.1 M sodium phosphate pH 8 as the buffer. This reaction was allowed to incubate overnight at 4 ºC to generate activated Ub G76C-TNB. A buffer exchange into non-reducing reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl) was then performed to remove excess TNB. To generate Ub-SS-Ube2N, an equimolar ratio of Ube2N and Ub G76-TNB was combined in non-reducing reaction buffer and allowed to incubate for 3 hours at 25 º under gentle rocking. Ub-SS-Ube2N. The reaction was then concentrated to 4 mL and subjected to size-exclusion chromatography using a Superdex S75 column (GE Healthcare) with non-reducing reaction buffer. Fractions containing pure Ub-SS-Ube2N were pooled and concentrated.

4.2.5 Crystallization of MavC in Complex with Ub-Ube2N

The complex was formed by mixing purified $MavC^{1-384}$ C74A and Ub-Ube2N together in an equimolar ratio. This mixture was incubated on ice for 1 hour followed by concentration of the sample to give a final protein concentration of 28 mg/mL. Conditions were screened using the hanging drop vapor diffusion method at 21 ºC. Crystals appeared in numerous conditions across many screens as needles and thin rods. To confirm the presence of the desired complex in the crystals, SDS-PAGE analysis was performed after washing crystals with crystallization buffer to remove protein in solution and dissolving in water (Figure 4-2). However, crystals from initially identified hit conditions diffracted poorly. Therefore, optimization of the crystallization conditions was carried out. The condition 0.2 M sodium malonate and 25% w/v PEG 3350 was further optimized by sampling different pH levels and different concentrations of PEG 3350, with more robust, 3-dimensional hexagonal rod-like crystals growing in a condition containing 25% PEG 3350 and 0.2 M sodium malonate pH 9.0. Further optimization was performed via additive screen, where the best-diffracting crystals grew in a condition of 25% PEG 3350 and 0.2M sodium malonate at pH 9.0 with 10 mM $NiCl_2$.

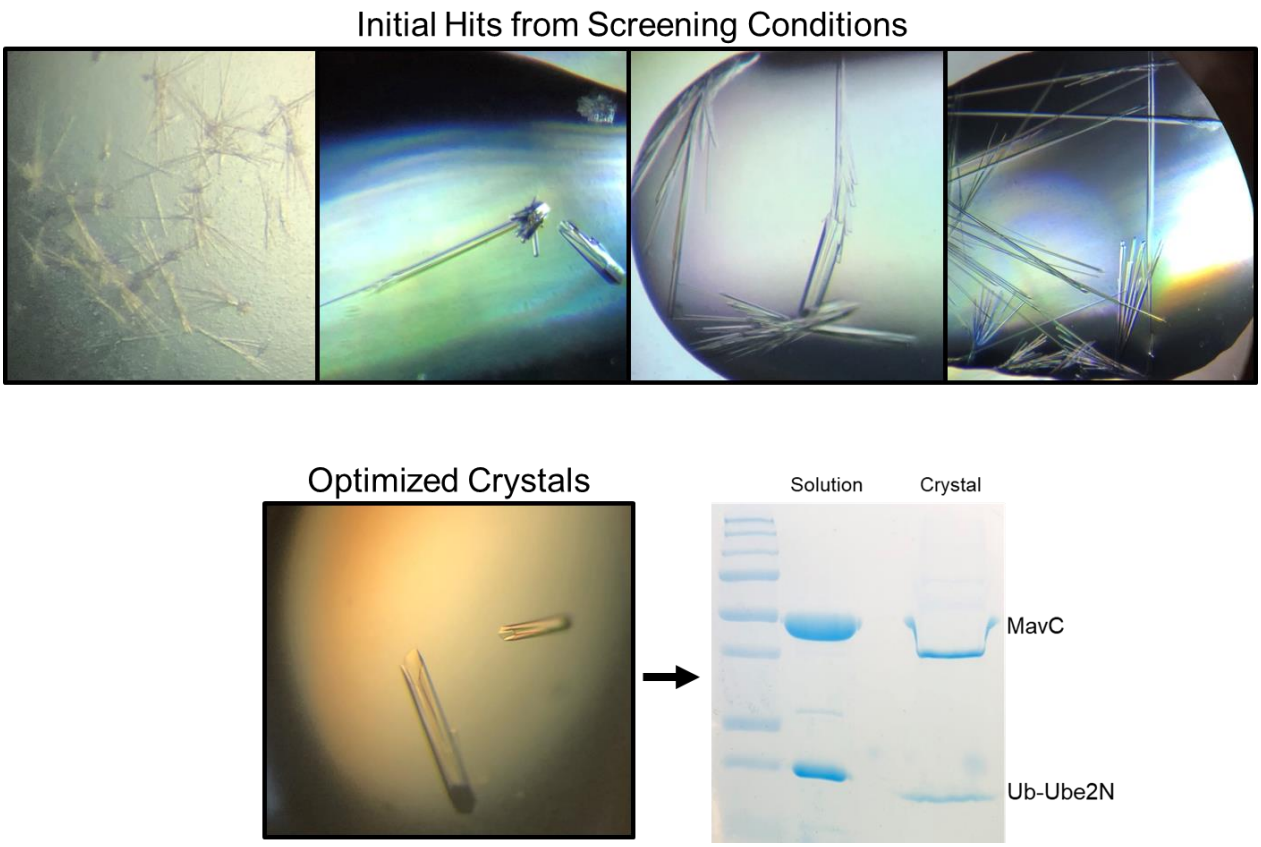


Figure 4-2: Crystallization of MavC$^{1-384}$-Ub-Ube2N complex. Top panels show crystals grown in initial screening conditions. Bottom panels show crystals used for diffraction and SDS-PAGE analysis.

4.2.6 Crystallization of MavC in Complex with Ube2N-SS-Ub

The complex was formed by mixing purified MavC$^{1-384}$ C74A and Ube2N-SS-Ub together in an equimolar ratio to give a final protein concentration of 32 mg/mL. This mixture was incubated on ice for 1 hour. Conditions were screened using the hanging drop vapor diffusion method at 21 ºC. Crystals appeared in several conditions of the PEG-Ion Screen (Hampton Research) (Figure 4-3). To optimize crystal formation, the protein concentration was reduced to 24 mg/mL. Crystals used for data collection were grown in three different conditions: Crystal 1) 0.1 M sodium acetate pH 4.6 and 3.5 M sodium formate. Crystal 2) 0.2 M potassium bromide and 30% PEG 2000 MME, Crystal 3) 0.2 M sodium formate and 20% PEG 3350. Crystallization of

MavC$^{1-384}$ C74A mutant in complex with Ube2N-SS-Ub was performed by Dr. Shalini Iyer (Das Lab, Purdue University).

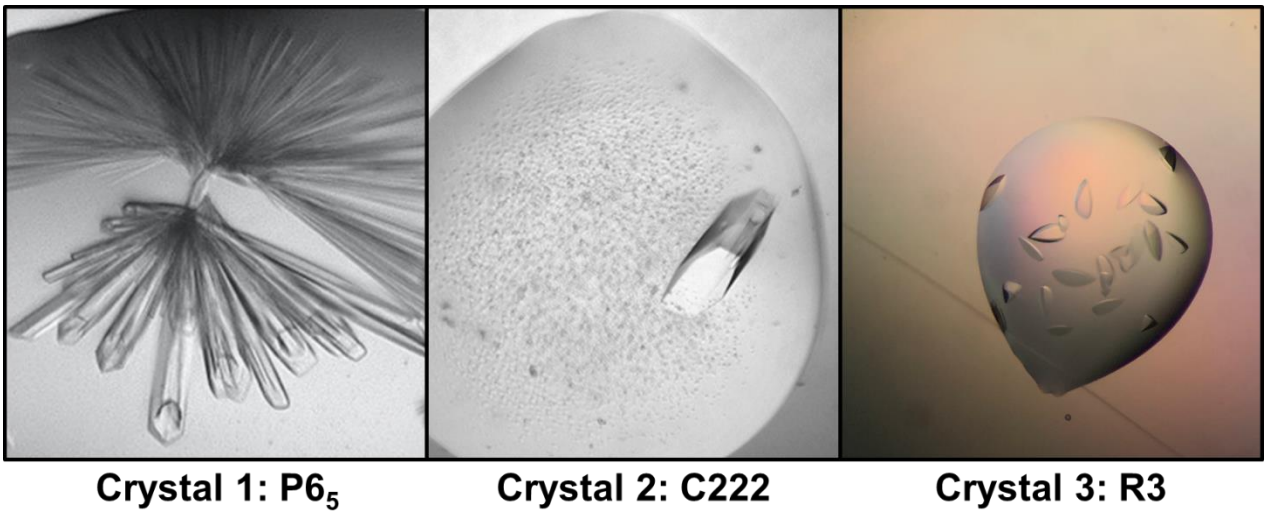


Figure 4-3: Crystallization of Ub-SS-Ube2N-MavC complexes. Three crystal forms are shown along with space groups and nomenclature.

4.2.7 Data Collection and Structure Determination

For the complex of MavC$^{1-384}$ C74A with Ub-Ube2N (product-bound complex), a complete dataset to 2.07 Å was collected from a single crystal at the Advanced Photon Source (APS) at Argonne National Laboratory on the LS-CAT 21-ID-G (λ = 0.97857 A) beam line. Data were processed and scaled using HKL2000 in the hexagonal space group, $P6_5$ (Table 1).[95]

For the complex of MavC$^{1-384}$ C74A and Ube2N-SS-Ub (substrate-bound complexes), complete datasets of Crystal 1 to 1.97 Å, Crystal 2 to 2.34 Å, and Crystal 3 to 2.80 Å were collected at the Advanced Photon Source (APS) at Argonne National Laboratory on the LS-CAT 21-ID-G (λ = 0.97 Å) beam line. Data were processed and scaled using HKL2000 in the rhombohedral R3:H space group (Crystal 1), the centered orthorhombic $C222_1$ space group (Crystal 2), and the hexagonal $P6_5$ space group (Crystal 3).

The structure of the product-bound complex was determined by molecular replacement using Phaser from the CCP4 suite.[96] The apo crystal structure of MavC (PDB: 5TSC), Ube2N (PDB: 2C2V), and Ub (PDB: 1UBQ) were all utilized together as the search models. The asymmetric unit contained one complex of MavC and Ub-Ube2N. The structure was refined by

the Phenix Refine program in conjunction with the model-building program Coot.[97] The final structure was validated using MolProbity and deposited in the Protein Data Bank under the code 6P5B.

The structures of the substrate-bound complexes were determined using similar methods to that of the product-bound complex. All substrate-bound structures' asymmetric units contained one complex of MavC and Ub-SS-Ube2N.

In the complex of Crystal 2 ($C222_1$), no clear electron density was observed between residues 89-94 of Ube2N, and therefore this region was left unmodeled in the structure. Structure determination of Crystal 1 and 2 was performed by Shalini Iyer (Das Lab, Purdue University). Detailed statistics are given in Table 4.

Table 4-1: Data collection and refinement statistics.

	$MavC_{1-384}$_Ub-SS-UbE2N Crystal 1	$MavC_{1-384}$_Ub-SS-UbE2N Crystal 2	$MavC_{1-384}$_Ub-SS-UbE2N Crystal 3	$MavC_{1-384}$_Ub-UbE2N product complex
PDB ID	**6ULH**	**6UMS**	**6UMP**	**6P5B**
Resolution range	25.34 - 1.968 (2.039 - 1.968)	38.63 - 2.344 (2.428 - 2.344)	30.75 - 2.8 (2.85 - 2.80)	30.359 - 2.099 (2.174 - 2.099)
Space group	R 3:H	$C222_1$	P 6_5	P 6_5
Unit cell	a=b=171.86 Å c=58.36 Å γ=120°	a=98.55 Å b=104.37 Å c=124.44 Å α=β=γ=90°	a=b=150.478 Å c=53.357 Å γ=120°	a=b=147.84 Å c=53.23 Å γ=120°
Total reflections	1107581	694463	1888913	1383292
Unique reflections	45560	27260	19287	39130
Multiplicity	5.8 (5.5)	7.1 (5.3)	10.2 (9.7)	10.4 (10.0)
Completeness (%)	99.91 (99.74)	99.83 (97.83)	99.75 (99.88)	99.16 (97.89)
Mean I/sigma(I)	10.5 (1.7)	12.24 (1.45)	15.13 (1.17)	22.0 (2.5)
Wilson B-factor	31.44	40.03	66.93	32.82
R-merge	0.098	0.187	0.201	0.136
CC1/2 (last shell)	0.475	0.433	0.514	0.777
Reflections (refinement)	45553 (4570)	27196 (2620)	17235 (1696)	38834 (3802)
Reflections (R-free)	2007 (194)	1286 (127)	1738 (174)	2020 (177)
R-work	0.1767 (0.2483)	0.2065 (0.3013)	0.2426 (0.3433)	0.1782 (0.2235)
R-free	0.2297 (0.3211)	0.2437 (0.3631)	0.3013 (0.3830)	0.2316 (0.3093)
Non-hydrogen atoms	4953	4688	4608	5084
macromolecules	4655	4520	4608	4784
solvent	298	168	0	300
RMS bonds (Å)	0.007	0.002	0.003	0.008
RMS angles (°)	0.86	0.49	0.59	0.89
Ramachandran - favored	98.81	97.79	95.31	97.87
Ramachandran - allowed	1.19	2.21	4.69	2.13
Average B-factor (Å²)	40.96	59.17	67.39	39.7

4.2.8 Ubiquitination and Deamidation Assays

The ubiquitinating activity of MavC mutants was determined by combining with Ube2N-SS-Ub at a concentration of 0.005 µM MavC, 25 µM Ube2N-SS-Ub and incubated at 37 °C for 30 minutes in non-reducing reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl). The reaction products were analyzed by SDS-PAGE and visualized with Coomassie Blue.

Ubiquitin deamidating activity of MavC mutants was determined by combining with Ub at a final concentration of 1 µM enzyme and 100 µM Ub. The reactions were incubated at 37 °C for

30 minutes in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT). The deamidation reaction products were analyzed by Native-PAGE and visualized with Coomassie Blue.

4.3 Results

4.3.1 Overall Mode of Binding and Active Site

The general architecture and domain organization of MavC in the complexes solved is similar to the structure of apo MavC (PDB: 5TSC). The protein can be divided up into three distinct regions: first, a central core globular domain (CG) that contains the enzyme's active site, second, a tail-like α-helical extension (HE) making up a region next to the CG domain and which has been defined in previous structural studies of Cif family proteins, and lastly the described insertion domain (INS), spanning residues 128-225 located opposite the HE. Together, these domains make up a C-shape, with the CG domain and catalytic triad residues (Cys74, mutated to Ala in the structure, His231, and Gln252) located in the middle of this C-shape (Figure 4-4).

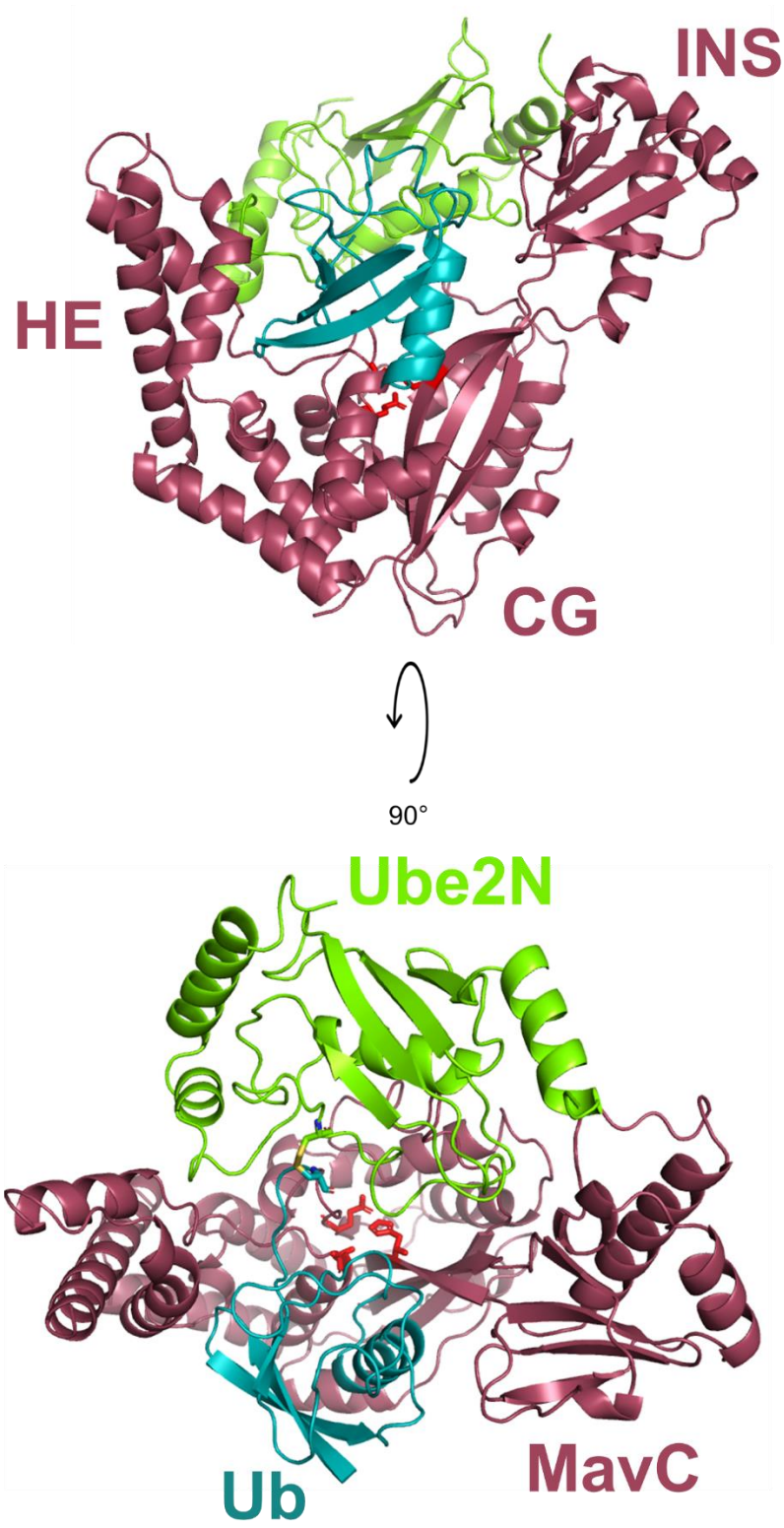


Figure 4-4: Overall binding mode of MavC in complex with Ub and Ube2N. Presented here is a model of the Ub-Ube2N bound structure. Active-site residues of MavC are highlighted in red.

Ube2N and Ub bind in an overall similar fashion in all complex structures (with some key variations discussed below), and are present in an extended conformation, bridging the active site with their respective linkages. In the Ub-Ube2N-bound MavC, clear electron density was observed between Gln40 of Ub and Lys92 of Ube2N (Figure 4-5).

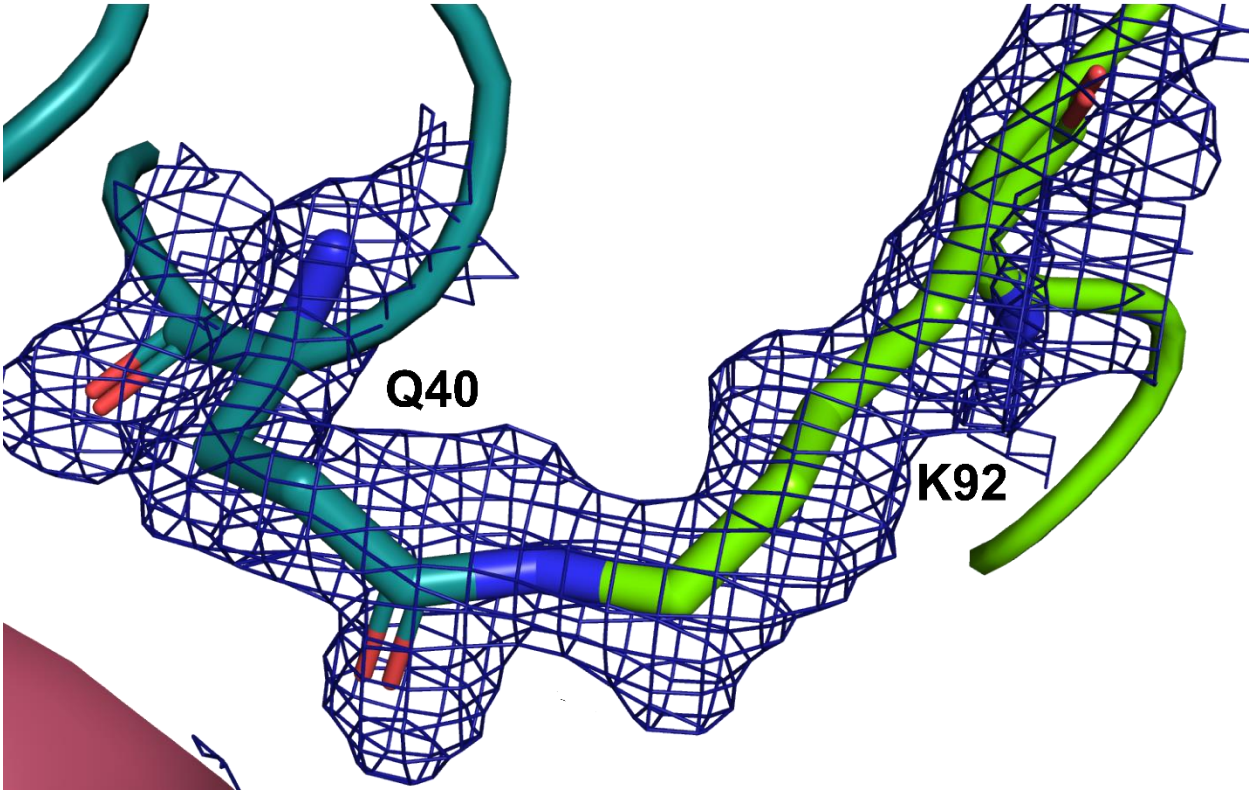


Figure 4-5: Zoomed-in view of the Ub-Ube2N complex with MavC, with electron density map showing clear evidence of isopeptide linkage between Gln40 of Ub and Lys92 of Ube2N.

The overall Ub binding mode resembles that of CHBP, a Cif family member found in *Burkholderia pseudomallei*, whose structure in complex with Ub has been determined.[98] A structural alignment between these two shows a general similarity, with an RMSD of 0.782 between CHBP-Ub and MavC-Ub. This suggests that despite weak Ub affinity, the canonical Cif substrate binding site has not diverged in MavC.

Consistent with our biochemical understanding of MavC's activity, in all of our complex structures, Gln40 of Ub is positioned near residue 74 of MavC in the CG domain. While the construct crystallized had Cys74 mutated to Ala in order to present unwanted Ub modification, modeling a Cys in that position reveals that the γ-S atom is located within 3.0 Å of the target

carboxamide group of Gln40 of Ub and is therefore in the correct position to make a nucleophilic attack. The active site residues Cys74, His231, and Gln232 are in the same position between all complex structures as well as the apo structure of MavC, indicating that it is a stable region of the enzyme (Figure 4-6).

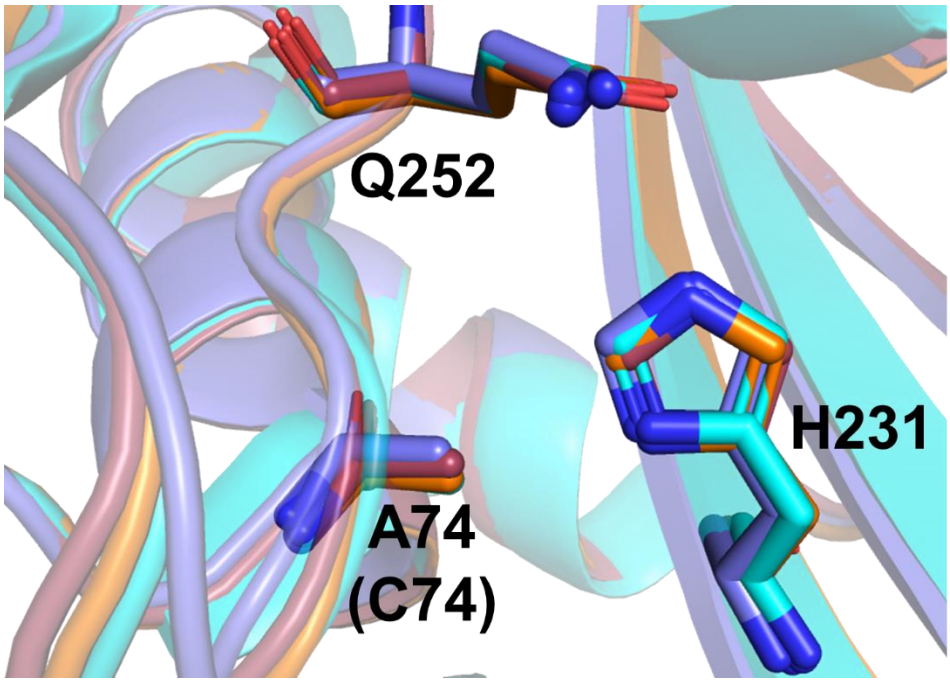


Figure 4-6: Alignment of all MavC structures, with active site triad residues depicted in sticks.

Numerous interactions with other residues of the MavC active site serve to hold Gln40 of Ub in place (Figure 4-7). First, it is located with its side-chain C=O group within hydrogen-bonding distance of the backbone amide groups of the catalytic Cys74 (2.9 Å) as well as Trp255 (3.0 Å) of MavC, whereas the NH_2 group of Gln40 is within hydrogen-bonding range of the backbone carbonyl group of Thr230 of MavC (2.8 Å). These backbone amide interactions would be important for stabilization of the oxyanion transition state which is an important feature of catalytic triad catalytic mechanisms. Also, His231 is likely able to donate a proton to the leaving ammonia from Gln during the formation of the thioester. The indole side chain of Trp255 of MavC is stacked against the peptide linkage between Gln40 and Gln41 of Ub, an arrangement that permits the backbone carbonyl of Gln40 of Ub to come within hydrogen-bonding distance from the hydroxyl group of Ser73 of MavC (Figure 4-7). Combined, these interactions appear to fix the Gln40 side chain of Ub in a reactive arrangement for attack by the nucleophilic Cys to facilitate

formation of the thioester intermediate. In this arrangement the NH_2 group of Gln40 points toward a solvent-filled area, which would allow the ammonia released during catalysis to diffuse away from the active site.

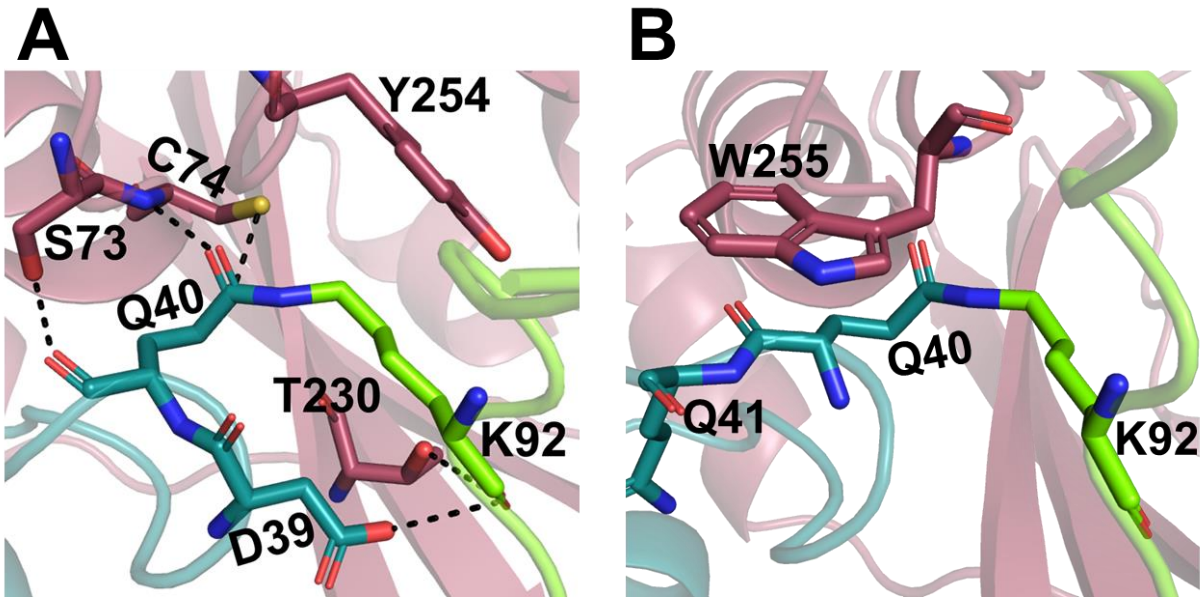


Figure 4-7: Close up view of active site, showing interactions between MavC, Ub, and Ube2N. A) and B) show different orientations highlighting different interactions. Key residues are shown in sticks. MavC is colored burgundy, Ub teal, and Ube2N green.

Similarly to Gln40 of Ub, the key interactions positioning Lys92 of Ube2N into the active site are also discerned in the Ub-Ube2N product bound complex (Figure 4-7). The carbon chain of Lys92 makes likely van der Waals contacts with Tyr254 of MavC. Thr230 of MavC, whose backbone interacts with Gln40 of Ub, may play a double role in recognition as it also appears to exist within hydrogen-bonding range of the amino group of Lys92 of Ube2N.

4.3.2 The Ube2N Binding Interface

There are broadly three main patches of interaction between Ube2N and MavC observed in our structures (Figure 4-8A). The common features as well as the variations will be described here. Region 1 consists of a series of hydrophobic and hydrogen-bonding interactions between the insertion domain of MavC and Helix 1 and Loops 4 and 7 of Ube2N (Figure 4-8B). This interface resembles the interface utilized by some E3 ligases, such as TRAF6.[55] As observed in Chapter 3,

the insertion domain is essential for recognition of Ube2N and ultimately ubiquitination by MavC. Of the three interacting regions, Region 1 is the most consistent between all four complex structures. Region 2 involves interactions between Loop 7 of Ube2N and the CG domain of MavC. The product-bound complex structure shows more extensive interactions, especially with Met317 of MavC, that are not observed in the substrate-bound complexes (Figure 4-8C). These variations may represent a remodeling of Ube2N to promote attack of the Ub-MavC thioester by Lys92 of Ube2N. A further analysis of this are described in **4.3.6**. The third region of interaction between Ube2N and MavC is located between Helix 3 of Ube2N and the HE domain of MavC, where extensive interactions between acidic and basic residues are observed (Figure 4-8D). Here there is considerable variation between the complexes solved. For example, interactions involving Arg63 and Lys64 of MavC with Helix 3 of Ube2N that are observed in Crystal 3 and the product-bound complex are not observed in Crystals 1 and 2 of the substrate-bound complex, where instead Helix 3 of Ube2N appears to bind to residues of the HE domain located nearer to the CG. This change in the position of Ube2N in these structures can likely be explained by a movement of the insertion domain, further discussed in **4.3.5**.

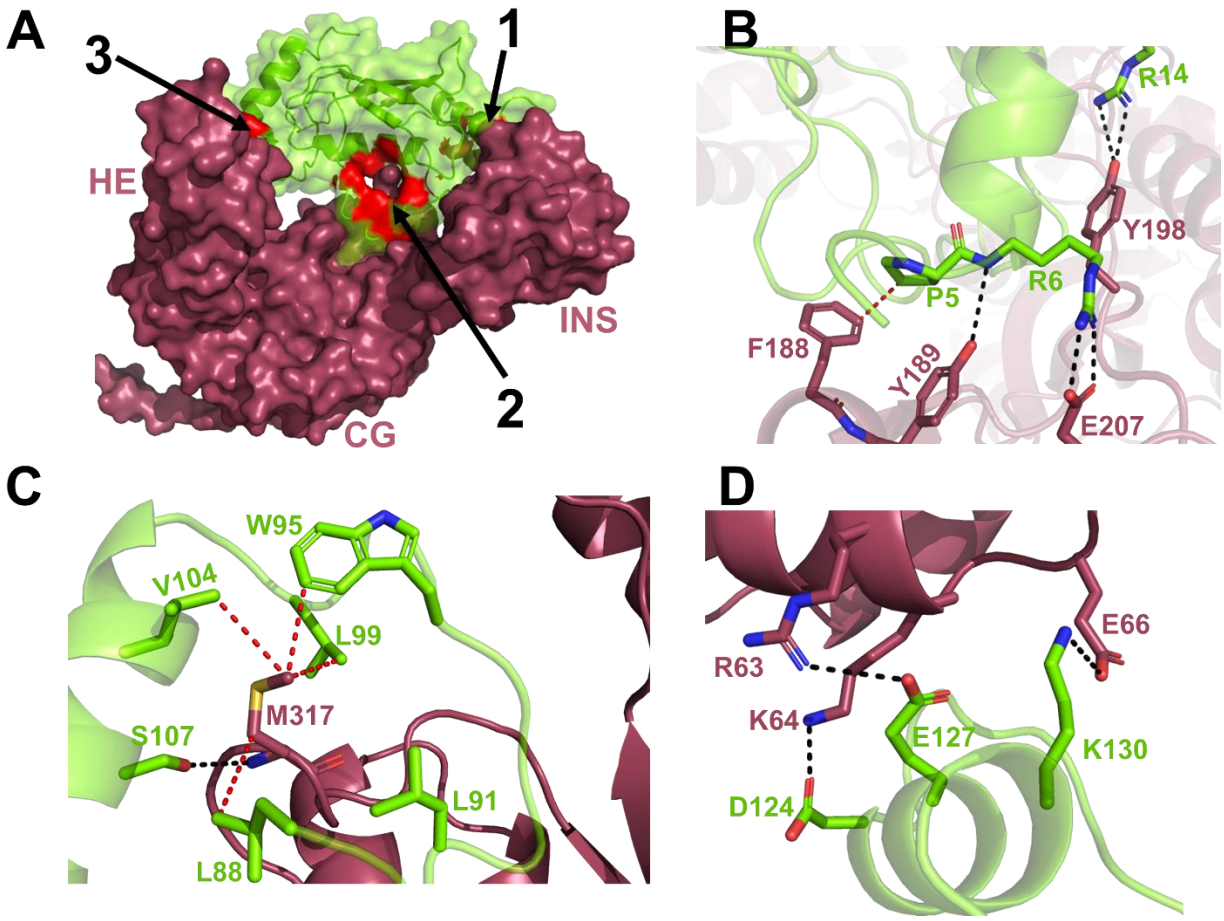


Figure 4-8: Interactions between Ube2N and MavC. A) Product-bound MavC structure with Ub removed and regions of Ube2N interaction highlighted in red. B) Close up view of interactions between MavC and Ube2N in Region 1. C) Interactions in Region 2. D) Interactions in Region 3. MavC is depicted in burgundy and Ube2N in green. Hydrophobic interactions depicted as red dashed lines and electrostatic interactions depicted as black dashed lines.

4.3.3 The Ubiquitin Binding Interface

The Ub binding site was well-conserved across all four complex structures. An examination of the Ub-MavC interface reveals three distinct patches of recognition that interact with each of the three domains of MavC (Figure 4-9A). Region 1 involves Ser21 and Asn25 of Ub forming hydrogen bonds with the insertion domain of MavC. (Figure 4-9B) As described in Chapter 3, when this insert was deleted ($MavC^{\Delta INS}$), Ub deamidating activity was retained but found to be weaker than that of the full construct (Figure 3-3). This observation, along with the contacts in the structure, suggests that the insert domain makes a meaningful contribution to Ub

binding. Region 2 of the interface is situated around the active site of the enzyme, involving a series of interactions around Gln40, the target residue as described in **4.3.1**. (Figure 4-9C). Region 3 involves contacts between the N-terminal β-hairpin turn residues Lys6, Leu8, and Thr9 of Ub, as well as nearby Arg72, forming interactions with the HE of MavC, specifically residues Leu36, Asn39, Glu40, Ile43, and Glu66 (Figure 4-9D). The engagement of Arg72 may serve to distinguish Ub from the related NEDD8, which has an Ala residue at that position (Figure 4-9F) and is consequently not recognized by MavC. When the Arg72 of Ub was mutated to Ala, MavC proved unable to ubiquitinate Ube2N. This method of discrimination between Ub and NEDD8 via the Arg72 residue has been observed before in other enzymes, such as the NEDD8-specific E1 and NEDD8-specific proteases.[99,100]

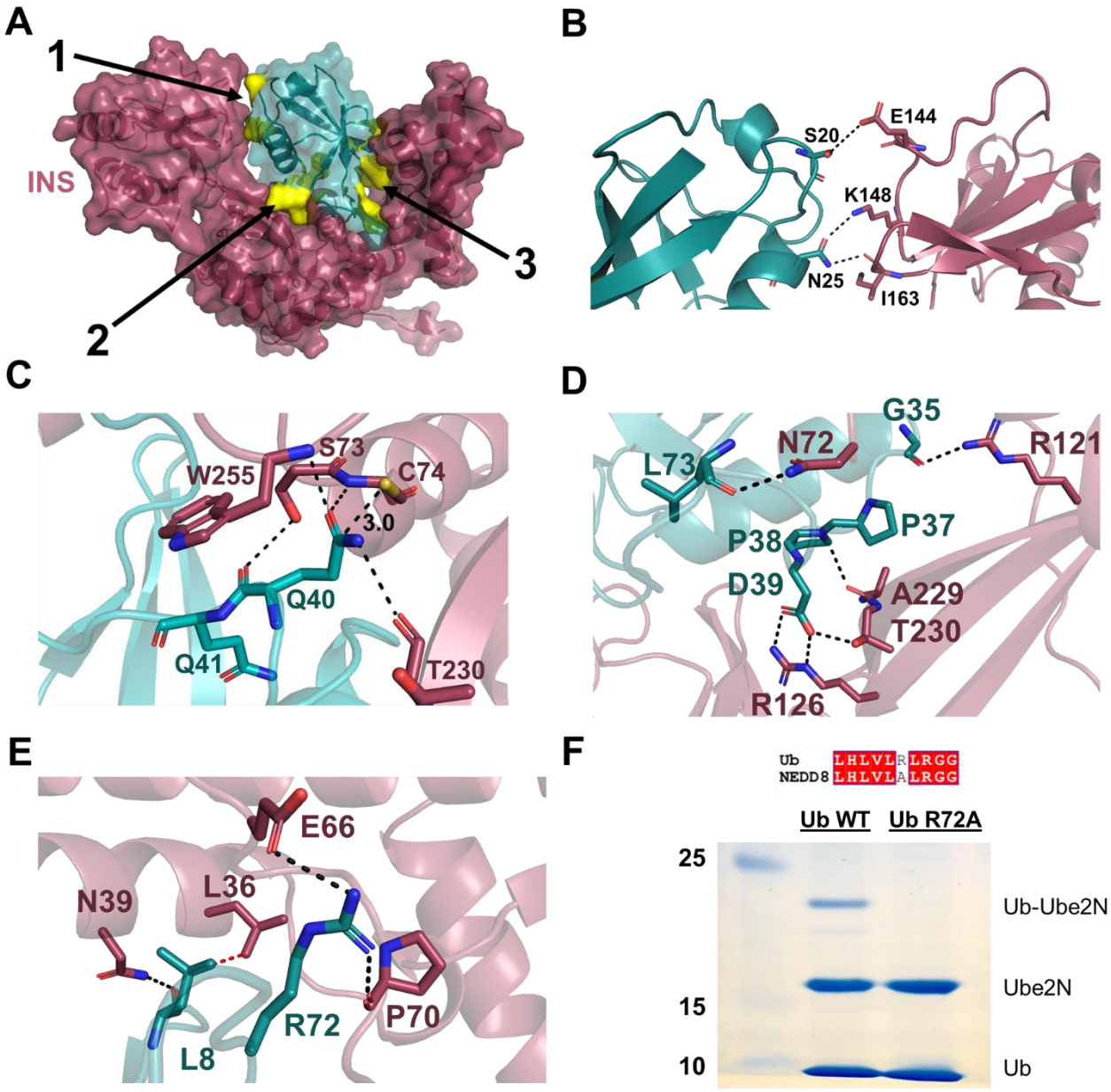


Figure 4-9: Interactions between Ub and MavC. A) Product-bound MavC structure with Ube2N removed and regions of Ub interaction highlighted in yellow. B) Close up view of interactions between MavC insertion domain and Ub in Region 1. C) Interactions in Region 2. D) and E) Interactions in Region 3. MavC is depicted in burgundy and Ub in teal. Hydrophobic interactions depicted as red dashed lines and electrostatic interactions depicted as black dashed lines. F) SDS-PAGE of ubiquitination reaction when Arg72 of Ub is converted to Ala. Sequence alignment of Ub and NEDD8 is also given.

4.3.4 Mutational Analysis of Key Binding Residues

To experimentally verify the importance of the interactions between MavC and both Ube2N and Ub observed in the crystal structures, selected residues of MavC were mutated and the activity of these mutants was compared to that of the wild type enzyme (Figure 4-10A). As a control, the catalytic mutant C74A was also included in these assays. In addition to comparing transglutamination of the disulfide conjugate, a Ub deamidation assay was also employed to further assess the activity of mutations at the Ub-binding interface of MavC (Figure 4-10B).

Within the parameters of the ubiquitination assay, all Ube2N binding mutants tested failed to produce any detectable levels of Ub-Ube2N, indicating their importance in Ube2N recognition.

Interestingly, a similar abolishment of transglutamination activity was also observed when Ub binding mutants of MavC were utilized in the assay. However, when deamidation activity was compared, while all MavC mutants exhibited a defect in comparison to the wild type, deamidation nevertheless was able to be observed in the case of the N39A, E66A, and N72A mutants, indicating that these mutations likely had a smaller effect on MavC-Ub interactions.

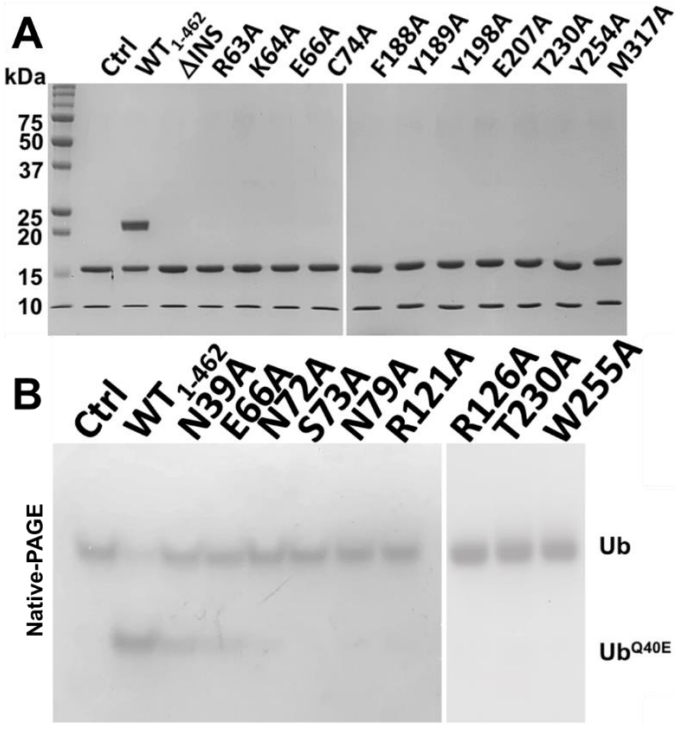


Figure 4-10: Analysis of MavC mutants. A) SDS-PAGE of ubiquitination reactions with Ube2N binding mutants of MavC. B) Native-PAGE of deamidation reactions conducted using Ub binding mutants of MavC.

4.3.5 Conformational Differences of Insertion Domain

While the overall layout and secondary structure of MavC in our newly characterized complexes was similar to the apo form of MavC, the conformation of the insertion domain exhibits significant differences between the apo MavC as well as between different complexes. A comparison of MavC from our crystal structures with the unbound structure reveals that conformational differences lie mainly in the insertion domain of the protein. The mean r.m.s.d. improves from 3 Å to 1.00 Å when the insertion domain is excluded.

Two key movements were observed. First, a common difference that set apart the apo structure specifically from the complex structures was a pronounced 30º rigid body rotation that the entire insertion domain undergoes to enable MavC to bind both Ub and Ube2N (Figure 4-11).

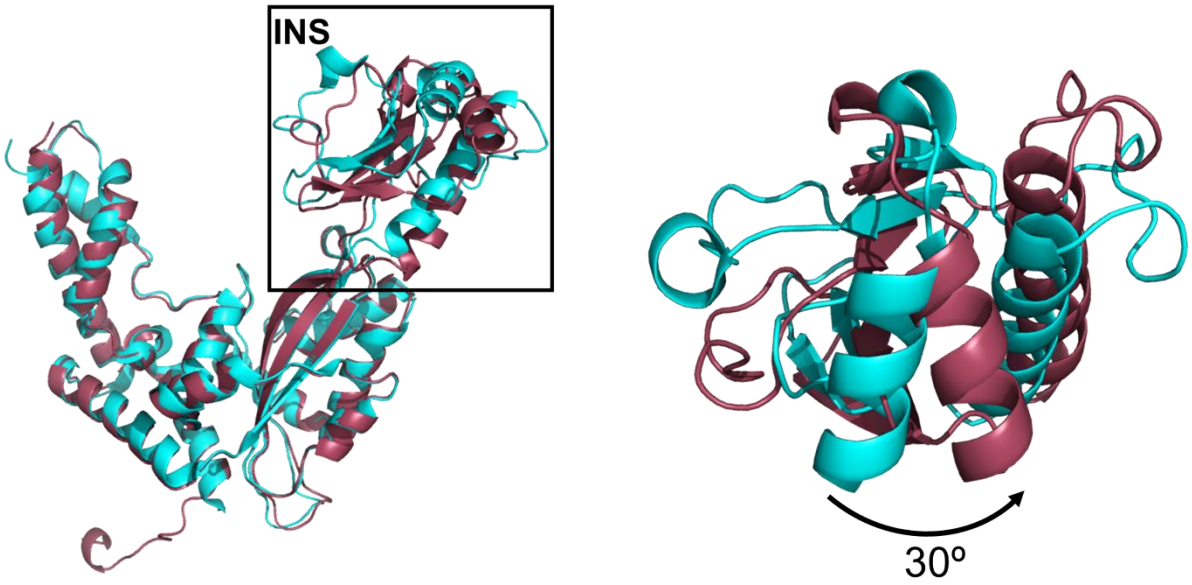


Figure 4-11: Alignment of MavC in the product-bound complex (burgundy) with apo MavC (PDB: 5TSC, blue) and a depiction of rigid-body rotation of insertion domain.

We observed that this rigid body rotation would be necessary for proper Ub binding, as superimposing Ub-Ube2N with the apo structure reveals that the “unrotated” form of the insertion domain encroaches onto the Ub binding site. An alignment of Ube2N with the apo form of MavC

reveals that the overall acidic patch of the insertion domain is facing away from the complementary basic N-terminal region of Ube2N (Figure 4-12).

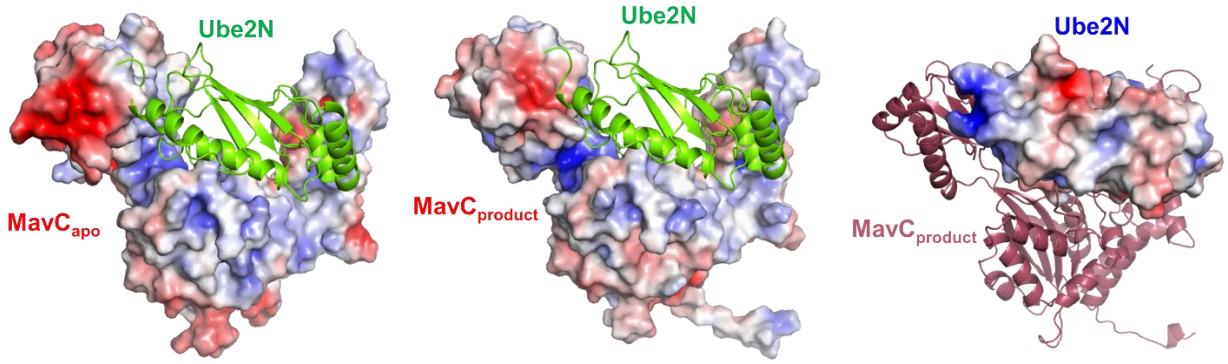


Figure 4-12: Depiction of charge surface of MavC and Ube2N, clearly showing an overall negative (red) region of MavC insertion domain complementing the positive (blue) N-terminal lobe of Ube2N.

This is in clear contrast to the bound forms of MavC, where the acidic patch has realigned, brought on by the rotation, to form a binding interface, facilitating binding of both Ube2N and Ub. The second key movement was observed by examining our crystal structures of the substrate-bound complex and noticing Crystal 3's difference from Crystals 1 and 2. This second movement is a rigid-body pendulum movement where the insertion domain further moves while keeping Ube2N bound (Figure 4-13).

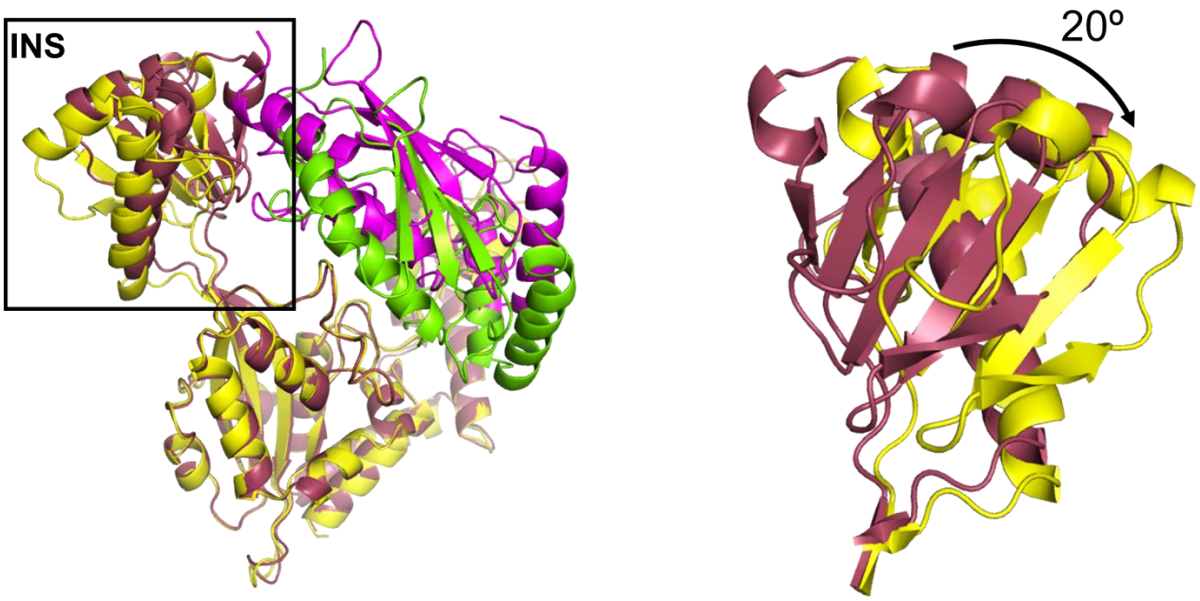


Figure 4-13: Alignment of MavC and Ube2N of the substrate-bound complex Crystal 3 (MavC: burgundy, Ube2N, magenta) with Crystal 1 (MavC: yellow, Ube2N, green) and a depiction of pendulum movement of insertion domain. The Ub moieties are hidden for clarity.

We find that in the product-bound structure, despite the absence of a covalent linkage, $Gly76^{Ub}$ and the catalytic $Cys87^{UbE2N}$ are positioned adjacent to each other, further suggesting that the Ub-charged Ube2N species is the relevant substrate (Figure 4-14).

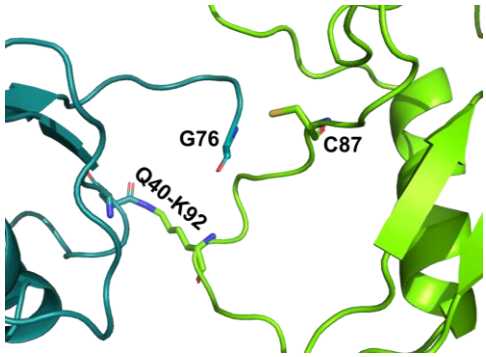


Figure 4-14: Positioning of Gly76 of Ub and Cys87 of Ube2N in the product-bound complex. These residues are covalently linked in the Ube2N~Ub charged complex.

Taken together, the insertion domain may freely interact with Ube2N as an independent binding motif, but as a part of the MavC enzymatic machinery these conformational movements are necessary to productively bind Ube2N.

4.3.6 Conformational Differences of Ube2N

Fascinatingly, while the structure of Ub in all of our complex structures remained largely unchanged with the exception of the flexible C-terminal tail, when we compared Ube2N across the complexes, we observed significant conformational differences which appear to be required in order for catalysis to occur (Figure 4-15). The most striking observation was that the Ube2N active site containing catalytic Cys87 as well as the target residues Lys92 and Lys94 underwent a significant remodeling across our solved structures.

The Ube2N moiety resolved in Crystal 3 of the substrate-bound structures resembled unbound Ube2N the most. Examination of the active site fascinatingly showed that Lys92, the major residue that is linked to Gln40 of Ub, was positioned over 16 Å away from the γ-carbon of Gln40 of Ub (The entirety of the Lys92 side chain was not resolved in this structure, so the distance given was measured from the β-carbon). Moreover, the Lys92-bearing region of Ube2N was wound into a 3-10 helix that is invariable in all known structures of Ube2N in the PDB. This placement is incompatible attacking the thioester intermediate and forming the isopeptide bond, indicating that a substantial change in this region would be required to bring Lys92 into the MavC active site.

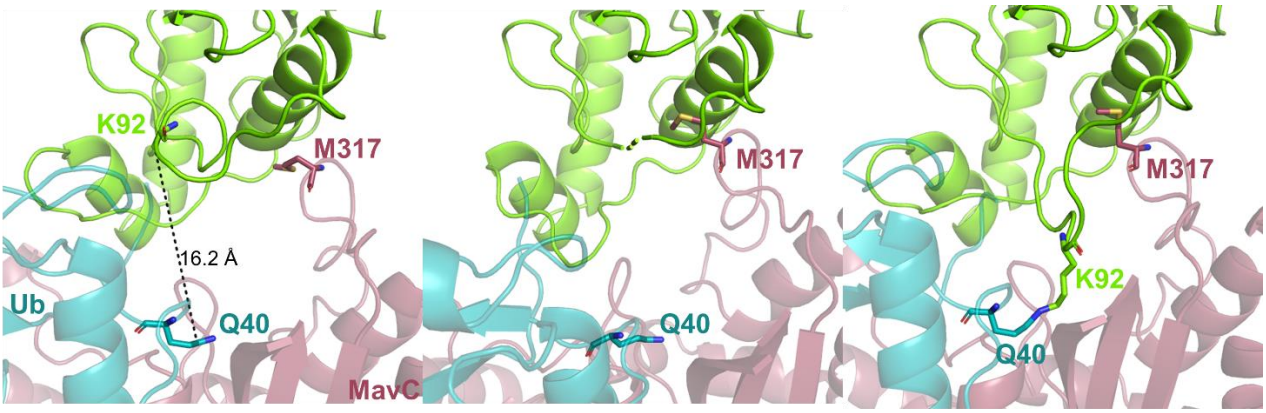


Figure 4-15: Scheme of Ube2N loop remodeling. Left panel: Substrate-bound MavC Crystal 3. Center panel: Substrate-bound MavC Crystal 1. Right panel: product-bound MavC. Ube2N is colored green, Ub in team, and MavC in burgundy.

In Crystals 1 and 2 of the substrate-bound structures, despite a good overall resolution, there was no discernible electron density for the 3-10 helix in our structure. This suggested a conformational heterogeneity (Figure 4-15, center panel).

Finally, in the product-bound complex, as described earlier in **4.3.1**, Lys92 of Ube2N is covalently tethered to Gln40 of Ub and the 3-10 helical region is totally unfolded, forming an extended loop which bridges the gap originally observed in the substrate-bound structure. Met317 of MavC is the cornerstone of this interface, stabilizing this loop in the complex via hydrophobic interactions as well as with Helix 2 of Ube2N, giving Lys92 the opportunity to be presented to the active site of MavC (Figure 4-8C). In the complex, this unfolded loop from Ube2N appears to be dynamic considering the B-factor of ~70 $Å^2$ for the region spanning residues 86 to 95 compared to the average B-factor of ~40 $Å^2$ for the whole complex. Further, the electron density of the loop residues around Lys92 was also relatively poor.

Besides the remodeling of the active site of Ube2N, the loop region of Ube2N spanning residues 118-122 is also shifted, making contacts with the HE domain of MavC in both Crystal 3 of the substrate-bound structure and the product-bound structure.

4.3.7 Structural Analysis of Ub-Ube2N Product

The Ub-Ube2N product resolved in complex with MavC appeared to be in an “open” conformation, with Gly76 of Ub and C87 of Ube2N appearing next to each other. We compared

this structure with available crystal structures containing UbE2N~Ub or its mimics (Figure 4-16). Three predominant conformations are adopted by Ub~Ube2N: position 1 is considered a "closed" conformation, with Ub making some additional contacts with Ube2N, and positions 2 and 3 more "open", where no significant interactions between Ub and Ube2N besides for the covalent linkage exist. The ubiquitinated MavC product was found to observe a conformation resembling position 2 in our structure. However, it cannot be ruled out that Ub-Ube2N may adopt multiple conformations after it is released from the MavC active site.

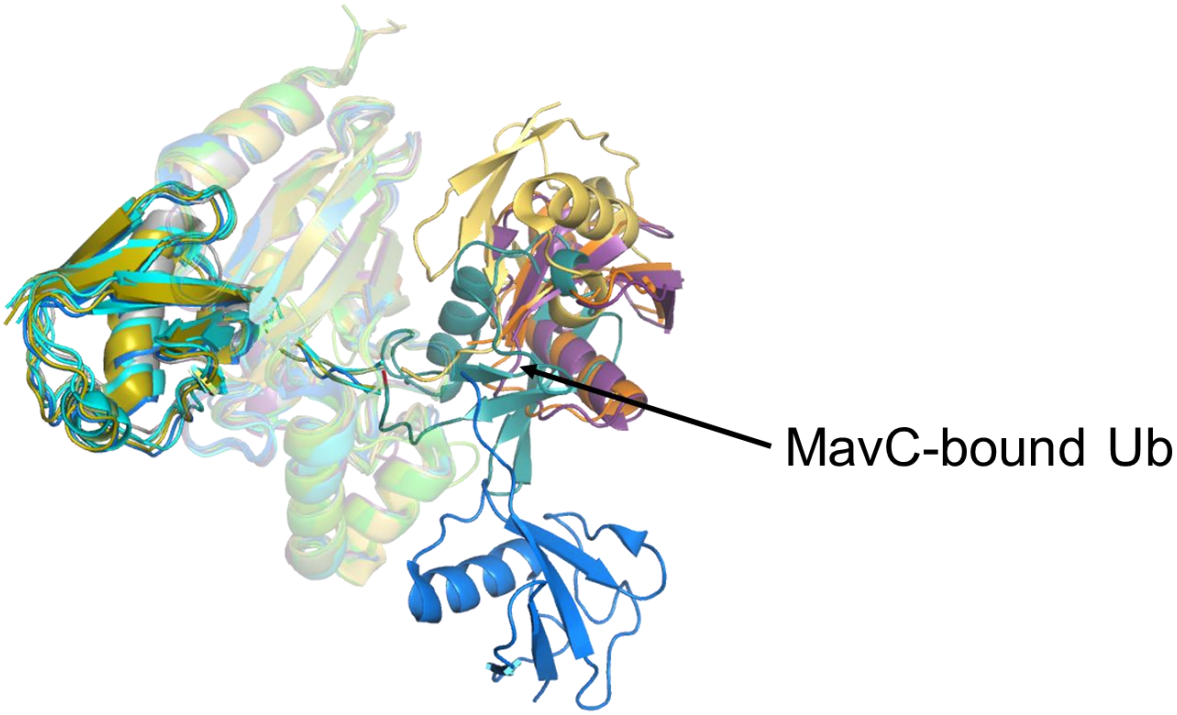


Figure 4-16: Alignment of Ube2N or Ubc13~Ub structures in the PDB with MavC product, Ub-Ube2N.

4.4 Discussion

4.4.1 Overall Significance

The structural data presented here represent important atomic-level snapshots of noncanonical ubiquitination catalyzed by the *Legionella pneumophila* effector MavC and provide a clear structural explanation for the previous biochemical results. Furthermore, the basis of

MavC's interactions with Ub and Ube2N have also been determined. This is also to our knowledge the first known structure of a transglutaminase in complex with both donor and acceptor substrates. This may also yield more general insights into the principles driving transglutamination in general. Interesting conformational changes were observed between the substrate and product bound structures, showing that MavC is a more dynamic enzyme than originally thought. Also, a remodeling of the active site region of Ube2N in order to allow for correct positioning of the target Lys92 residue was found to be important, and the structures indicate how MavC likely performs this process.

4.4.2 Binding and Substrate Recognition

From these structures, we have determined a molecular basis for the poor Ub binding by MavC observed in Chapter 3. By superimposing Ub bound to MavC in our complex structure onto the structure of the apo form of MavC, we clearly observe that Ub is occluded by the insertion domain (Figure 4-17). This likely proves the biggest challenge to Ub binding, despite the relative abundance of MavC-Ub interactions observed and surface area buried in the complex structures.

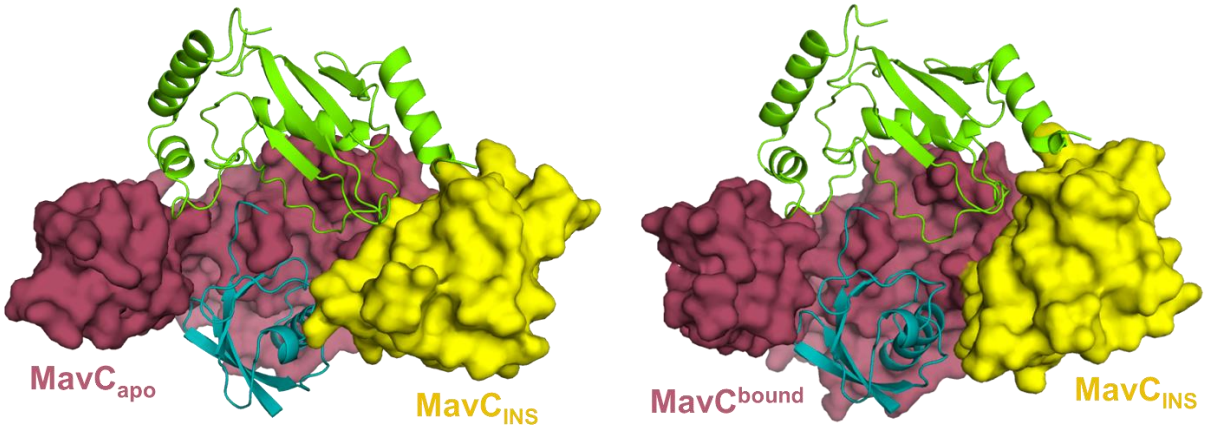


Figure 4-17: Structural comparison of apo vs product-bound MavC with Ub. Apo MavC structure was aligned with Ub-Ube2N from product-bound structure. Note the occluded binding site in the apo form.

Other *Legionella* effectors such as the DUB domain of SdeA are known to instead bury less surface area of Ub but nonetheless bind efficiently and without any potential steric obstacles.[45]

It is likely that this insertion domain can adopt multiple conformations in the apo form, where the Ub binding site may occasionally be open, since deamidation of free Ub by MavC is clearly observed *in vitro*. However, by binding to the charged Ube2N~Ub conjugate, the effective local concentration is increased, greatly increasing the likelihood of the tethered Ub moiety to bind. The Ub binding site is consistent across all four of our complex structures and resembles the Ub binding site of CHBP, a Cif family protein. It is likely that MavC, while retaining this binding site, lost the ability to robustly bind free Ub over the course of its evolutionary history as it gained the ability to bind Ube2N. The fact that the majority of E2s exist in this charged form may have influenced this process, as there is less of a need for this enzyme to separately bind Ub. Nevertheless, the Ub binding pocket still plays an important role, as mutations there cause a significant disruption in MavC's ubiquitinating ability. A more detailed analysis will be required in future studies to work out the relative importance of Ub binding versus Ube2N binding mutations.

Ube2N binding, as hypothesized, is strongly driven by the insertion domain, which consistently makes contacts with the N-terminal lobe of Ube2N in all crystal structures solved. This is a similar interface used by many E3 ligases to recognize Ube2N and may indicate how MavC is specific for this particular E2. In a more biological context, MavC competing with E3s that also recognize Ube2N may also have signaling effects in the cell. Due to this consistent binding of the insertion domain to Ube2N (in contrast with the HE and CG domain interactions), it is reasonable to conclude that initial recognition of Ube2N by MavC is driven by this interface.

4.4.3 Remodeling of Ube2N

The product-bound complex structure is unique in that it shows for the first time an unfolding of the 3-10 helical segment of an E2 enzyme driven by protein-protein interactions. Indeed, this unfolding is required in order to bring the amino group of Lys92 to the MavC active site (Figure 4-15). The dynamic nature of this loop may also explain why both Lys92 and Lys94 may be ubiquitinated by MavC, as detected in previous mass spectrometric analyses. The importance of Met317 is highlighted in the biochemical assays, as its mutation to alanine causes MavC to lose its ubiquitinating activity (Figure 4-10). Interestingly, in the homolog MvcA which is known to remove Ube2N from Ub-Ube2N (further discussed in Chapter 5), the key Met317 is a Leu instead, which may also provide a hydrophobic interface to interact with this region, which contains residues such as Leu88, Leu91, and Trp95. A structural comparison of MavC with the

determined structures of Cif and CHBP[98] reveals that instead of Met317, these deamidating effectors have charged residues in that position namely Glu and Asp respectively (Figure 4-18A).

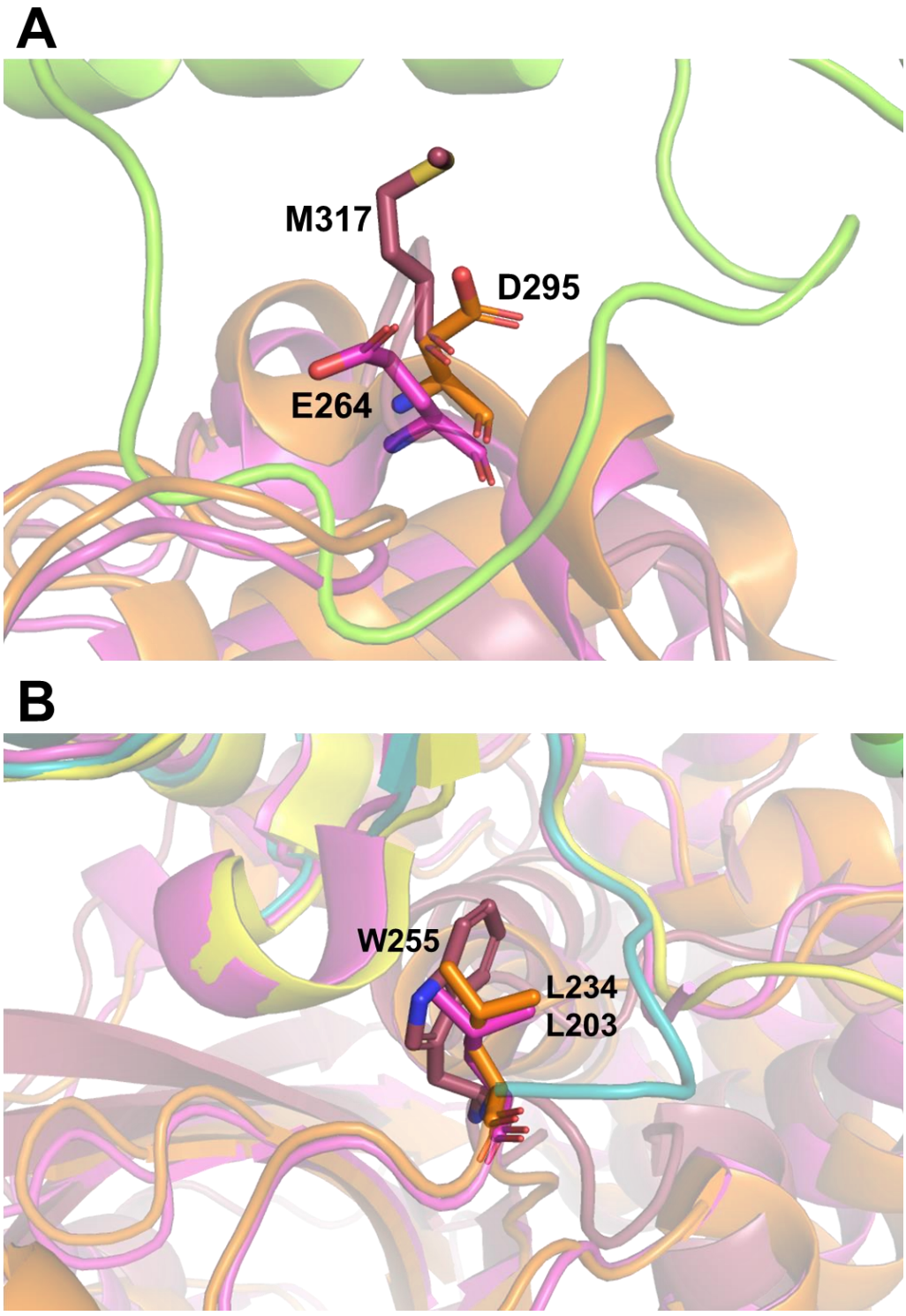


Figure 4-18: Alignment of MavC in complex with product with Cif members. A) Comparison of the M317 site of MavC B) Comparison of the W255 of MavC. The product-bound MavC complex (in burgundy) was aligned with CHBP in complex with Ub (PDB: 4HCN, in orange), and Cif in complex with NEDD8 (PDB 4FBJ, in magenta).

The observation of active site remodeling of Ube2N calls into question the stability of this region in the canonical ubiquitin-conjugating function of this enzyme. Though no other structural study until this one showed any variation in the 3-10 helix, one cannot rule out a dynamic function in other contexts besides that of MavC. Indeed, the catalytic residue Cys87 is adjacent to this helix, therefore the role of this region and its possible remodeling in Ube2N's ubiquitin conjugating and transferring activities may be an intriguing area of further investigation.

4.4.4 Dynamics of MavC and General Reaction Model

It is evident that the insertion domain must undergo a rigid-body rotation in order to properly align both Ube2N and Ub for catalysis (Figure 4-11, 4-17). This rotation is observed in all solved complex structures when compared to the apo structure. However, it is possible that in solution, there exists a state where the Ube2N moiety is bound to the insertion domain, but the Ub has not yet moved into the active site. In fact, considering the poor Ub binding of MavC and the initial steric occlusion of the Ub binding site, this is likely the initial state of the MavC-Ube2N~Ub bound complex.

Since the conformation of MavC in Crystal 3 resembled the product-bound structure, we surmised that Crystals 1 and 2, which are similar to each other, represented an earlier stage in MavC's ubiquitination process. We see that in Crystals 1 and 2, there are no contacts between Helix 3 of Ube2N and the HE domain of MavC, with Region 3 contacts being made instead elsewhere. As such, Ube2N is misaligned, with the Lys92-bearing region located too far from the MavC active site (Figure 4-15, 4-19). Therefore, a pendulum-like movement of the insertion domain is therefore necessary to bring the entire Ube2N moiety into position for catalysis. This is enabled not only by the conformational flexibility of the insertion, which is attached to the remainder of MavC via two unstructured linkers, but also by the fact that in the charged conjugate, Ube2N is attached to the flexible C-terminal tail region of Ub, which can adopt many configurations itself.

This series of structures allow us to map out a possible scheme for substrate recognition and catalysis over the course of MavC-catalyzed noncanonical ubiquitination (Figure 4-21). First, the apo form of MavC, where the insertion domain is in a dynamic state, recognizes and binds the Ube2N portion of Ube2N~Ub, using the Region 1 interface observed in all of our complex structures. Next, this complex, where Ube2N but not Ub is bound, undergoes conformational

adjustments of the insertion domain, the sum of which is a 30º rotation which opens up the sterically occluded Ub binding site and allows Ub to bind. It is important to emphasize that it is not clear that the rotation is the only movement performed by the insertion domain after Ube2N binding, and that there likely exists any number of possible conformations of the insertion domain that are ultimately unproductive for the ubiquitination reaction to occur. Following this process, a further adjustment of the insertion domain is made, which is a pendulum motion bringing Ube2N into the appropriate active site position and poising it for catalysis.

However, these conformational movements on the domain-wide scale are still not enough for the ubiquitination reaction to occur. Along with these movements described, the 3-10 helix of Ube2N containing the acyl-accepting Lys residue must unwind and extend into a loop conformation allowing the Lys residue to attack the thioester. Our structures are especially interesting in that Crystals 1 and 2 do not possess any electron density for this region, whereas Crystal 3 does, and shows an intact 3-10 helix. This may indicate that unfolding of this helical region and the adjustments of the insertion domain after Ub binding (e.g. the pendulum motion) may be decoupled from each other. It is possible for Lys92 to be in an extended conformation but Ube2N is still positioned too far from the thioester to react, and also vice versa. Due to these stipulations as well as the probable transience of Lys92 to be in the correct location for reaction, we speculate that the formation of the thioester between Ub Gln40 and MavC Cys74 may occur relatively soon after Ub binding. The fact that Gln40 of Ub is in a consistent position in all of our complex structures, located near residue 74 of MavC also supports this notion.

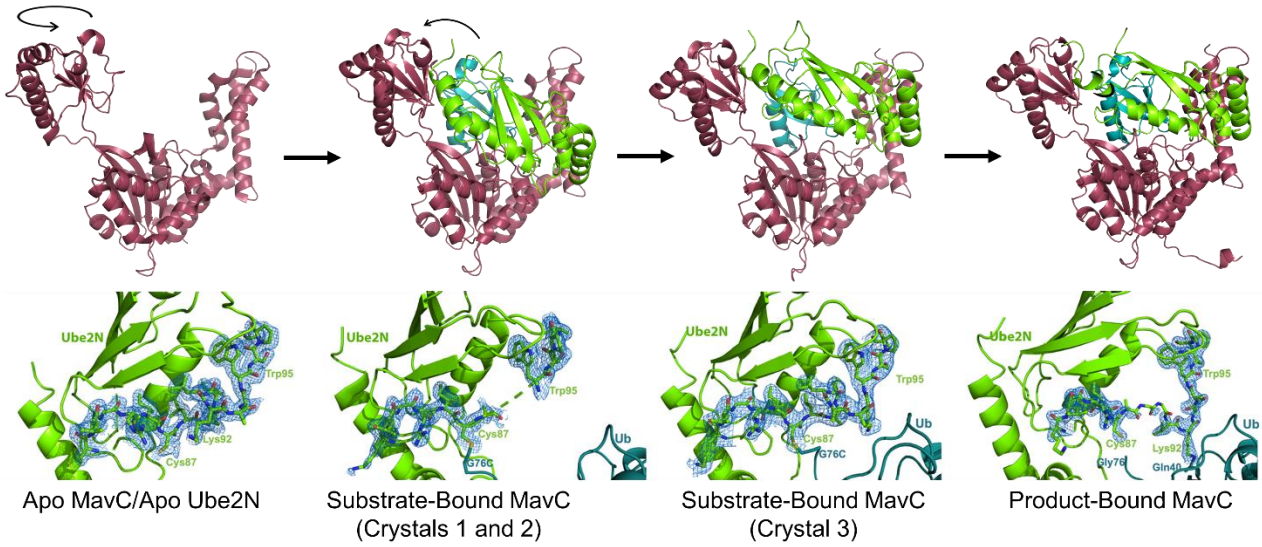


Figure 4-19: Summary of molecular transitions observed in crystal structures of MavC. Top panels show dynamic transitions of insertion domain, carrying Ube2N towards active site. Bottom panels show corresponding differences observed in 3-10 helical region of Ube2N, with electron density map aligned with structural models. Key residues are labeled.

Throughout this process, the biochemical evidence shows that deamidation of Ub is prevented from occurring to any significant degree. To protect this thioester from deamidation, it is possible that Trp255 and Tyr254 are responsible for preventing water from entering the active site, forming a hydrophobic pocket (Figure 4-20). Fascinatingly, transglutaminases classically have been known to possess a conserved Trp and in some cases a nearby Tyr residue in the active site that are essential for activity.[101,102] Previous studies have proposed a similar role for this tryptophan, where it acts as a "gate" to exclude water from entering the active site. Other studies also point to a possible role in stabilizing the oxyanion formed during catalysis,[103] which is less clear from examining our structures, but cannot be ruled out. Interestingly, superimposition of MavC with CHBP reveals that CHBP has a Leu instead of Trp residue in that position and it has a flexible loop region in the place of an adjacent Tyr (Figure 4-18B).[98] This is consistent with the fact that CHBP is a Ub deamidase and therefore requires ample solvent access for optimal activity.

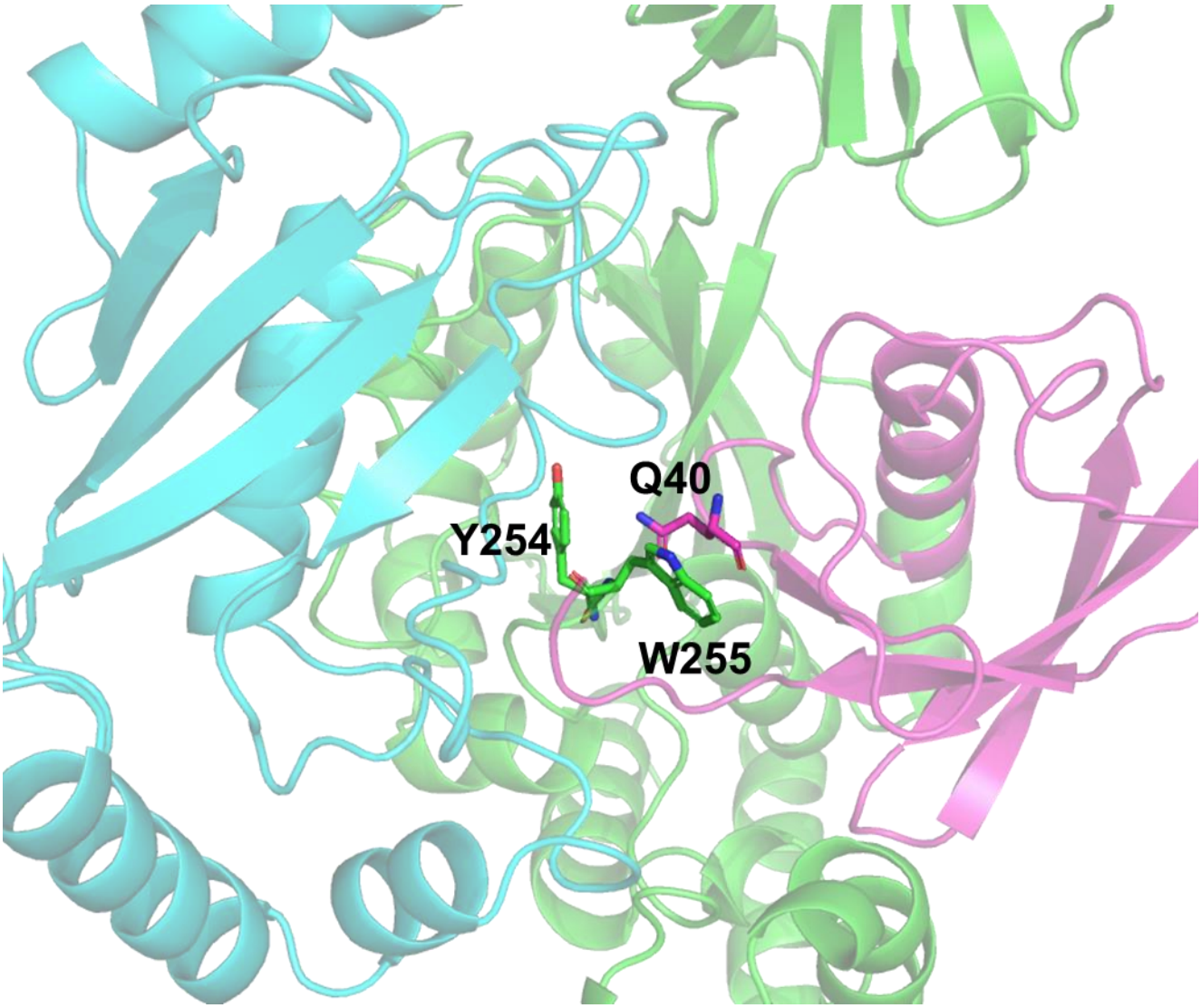


Figure 4-20: Substrate-bound structure of MavC (Crystal 3) with MavC colored green, Ube2N cyan, and Ub magenta. Hydrophobic residues of MavC surrounding target residue Gln40 of Ub are indicated in sticks and labeled.

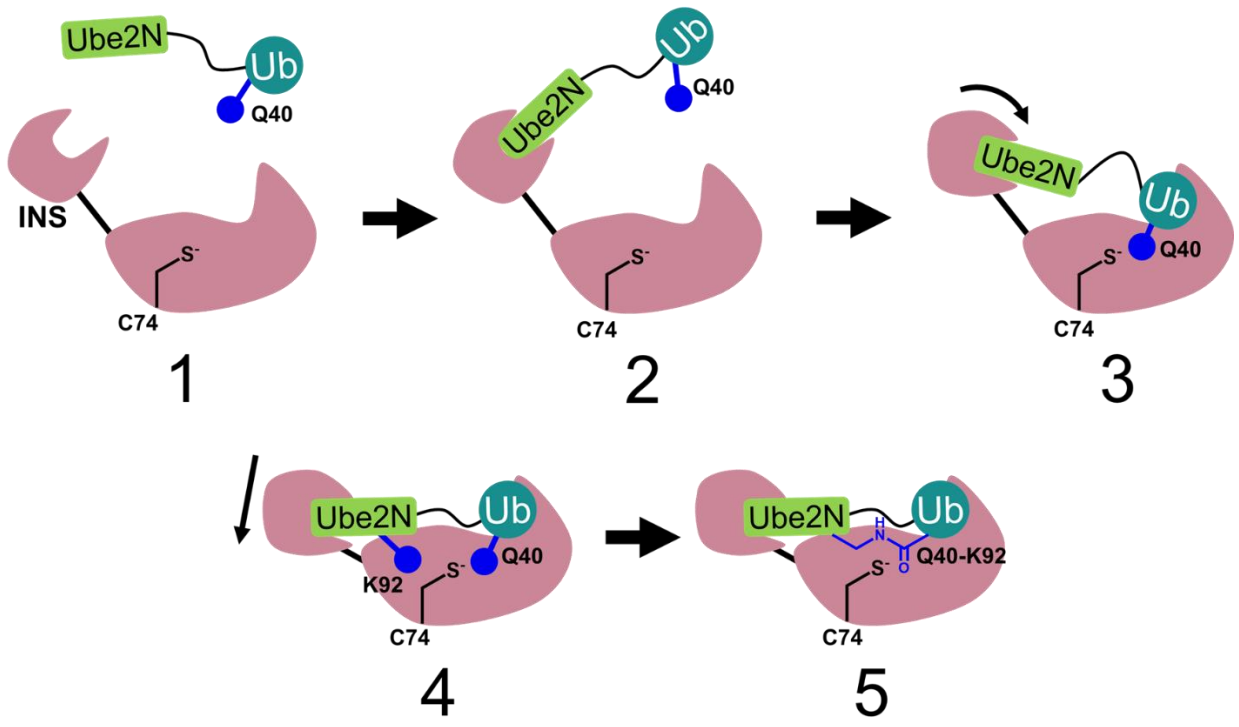


Figure 4-21: Proposed scheme of MavC-catalyzed ubiquitination based on structural and biochemical data. 1) Apo MavC. 2) Recognition of Ube2N moiety of Ube2N~Ub by insertion domain. 3) Adjustment of insertion domain to open up Ub binding site, allowing Ub to bind. 4) Further movement of insertion domain to bring Ube2N near active site, and positioning of Lys92 via unfolding of helix. 5) Catalysis of ubiquitination forming isopeptide linkage.

Conclusions and Future Directions

These open questions that remain in our model of MavC-catalyzed ubiquitination would be better addressed in future studies via careful biochemical and biophysical analysis of the roles of solvent access and transition-state stabilization, as well as the kinetics of thioester bond formation. Also, a substrate-bound crystal structure clearly showing the Lys92 in an attacking position would also further our understanding of the events occurring before reaction.

The origin of the insertion domain also remains an open question. Previous work in the microbiology field suggests that horizontal gene transfer from eukaryotic host organisms may play a role in the acquisition of new effectors by *Legionella* species. However, a DALI structural similarity search using the insertion as a query structure shows no significant hits which may point to this domain having independently evolved in *L. pneumophila*.

CHAPTER 5. REGULATION OF NONCANONICAL UBIQUITIN LIGASES

Portions of this chapter have been published in *Biochemistry* (DOI: 10.1021/acs.biochem.0c00067) and in *The EMBO Journal* (DOI: 10.15252/embj.2019102806)

5.1 Introduction

The myriad of *Legionella* effectors that are injected into the host cell are able to cause a number of different effects and serve many functions, both catalytic and non-catalytic. It is important to note that the deployment of effectors by pathogenic bacteria is more akin to a scalpel than a shotgun. The ultimate objective of these effectors is to carefully manipulate host cell signaling to generate a hospitable niche allowing replication of the pathogen.[104] Therefore, precise mechanisms of spatial and temporal regulation of effectors are necessary for *Legionella* to avoid undesirable effects on its host cell, as well as coordinate different stages of infection. Many such regulatory mechanisms have been well established, for example the presence of conserved phosphatidylinositol-4-phosphate (PI4P) binding domains in several effectors that provide spatial targeting and are required for LCV biogenesis and ultimately optimal *Legionella* replication.[105] Temporal regulation has also been observed via different means, with one example being control of expression. The ankyrin-like effectors (Ank) have been found to be regulated at the level of expression, enabling *Legionella* to only deploy them at the appropriate times.[106]

Perhaps more dramatic is the presence of effectors whose singular role is to regulate other effectors. These have been referred to as metaeffectors.[107] Many fascinating examples have been observed, such as the E3 ligase LubX which causes ubiquitination and proteasomal degradation of SidH.[31] Though the precise function of SidH remains unknown, *Legionella* strains lacking LubX have been found to become hyper-lethal when introduced to *Drosophila.* Further, SidH is toxic when ectopically expressed in yeast, with co-expression of LubX able to rescue that toxicity. A high degree of toxicity is not always desirable for the intracellular pathogen, as the host cell must be kept alive at least for long enough for intracellular replication to occur. Another example of metaeffectors include SidD, which serves to selectively de-AMPylate Rab1 modified by the effector SidM.[108–110] A rescuing of SidM's toxic effect in yeast was also observed when SidD was coexpressed. Intriguingly this was the first identified enzyme able to remove AMP from another

protein in a mechanism dependent on metal ions, reminiscent of some phosphatases. A more recently identified metaeffector has been Lpg2505, named SusF, which binds with nanomolar affinity to the cytotoxic effector SidI, whose translation-inhibiting and glycosyltransferase activities are both significantly attenuated upon binding.[111]

Both the SidE family and MavC, the two noncanonical Ub ligase systems described in the previous chapters, are not exempt from regulatory processes from metaeffectors. Even prior to the SidE proteins' unmasking as ubiquitinating enzymes, SidJ had been known to negate their toxic effect when coexpressed in yeast.[60] Further, it was observed that deletion of SidJ from *L. pneumophila* led to a defect in intracellular growth.[59] SdeA had been identified to localize to the LCV in early stages of infection but dissociated at later time points in a SidJ-dependent manner. Indeed, after the biochemical function of SidE enzymes was determined to be that of ubiquitinating Ser residues of protein targets via a PR-linker, SidJ activity was found to reduce the measured levels of SidE-ubiquitinated proteins in infected cells after a few hours, suggesting that SidJ's modulatory effect was related to SidE's enzymatic activity.[112] The initial study reporting this function concluded that SidJ acted as a deubiquitinase, serving to remove PR-linked ubiquitination from protein substrates. However, subsequent investigations by this group and 3 other independent groups intriguingly found that in reality, SidJ acted to switch off the activity of the SidE proteins altogether.[113–116] Mass spectrometric analysis revealed that this was accomplished via ATP-dependent addition of one or more glutamates to the catalytic Glu860 of SdeA which led to a loss in mART activity (Figure 5-1). This was a striking discovery of the first known instance of a bacterial glutamylase. Glutamylation had been understood before this discovery as a eukaryotic phenomenon, utilized to modify tubulin in cells.[117] SidJ was also found to tightly bind to calmodulin (CaM) and require it for this glutamylating function.

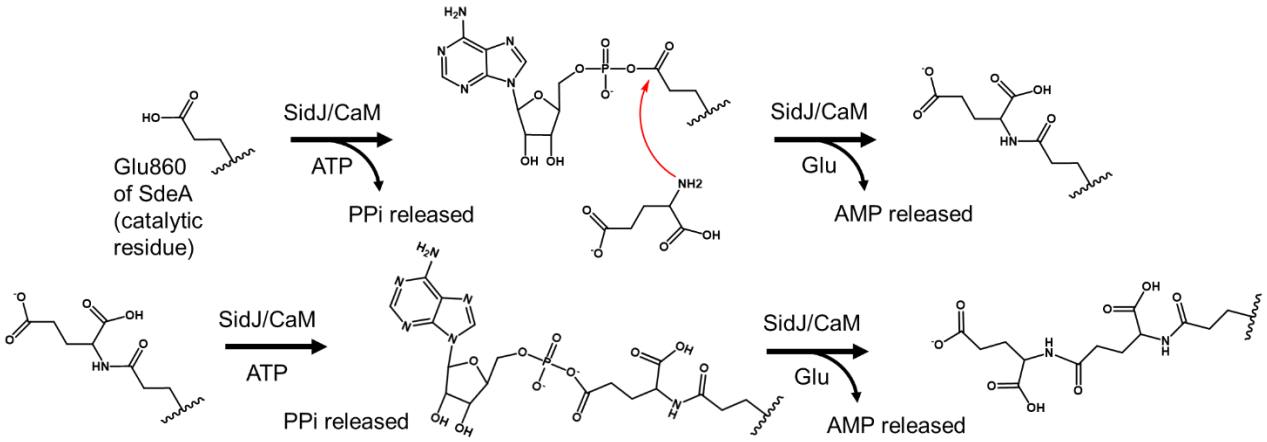


Figure 5-1: Proposed mechanism of SidJ-catalyzed glutamylation of Glu860 of SdeA.

The structure of SidJ in complex with CaM revealed a kinase-like fold of SidJ that contained two putative active sites, both of which were important for glutamylating activity.[113] CaM was bound via an IQ motif (Figure 5-2).[118]

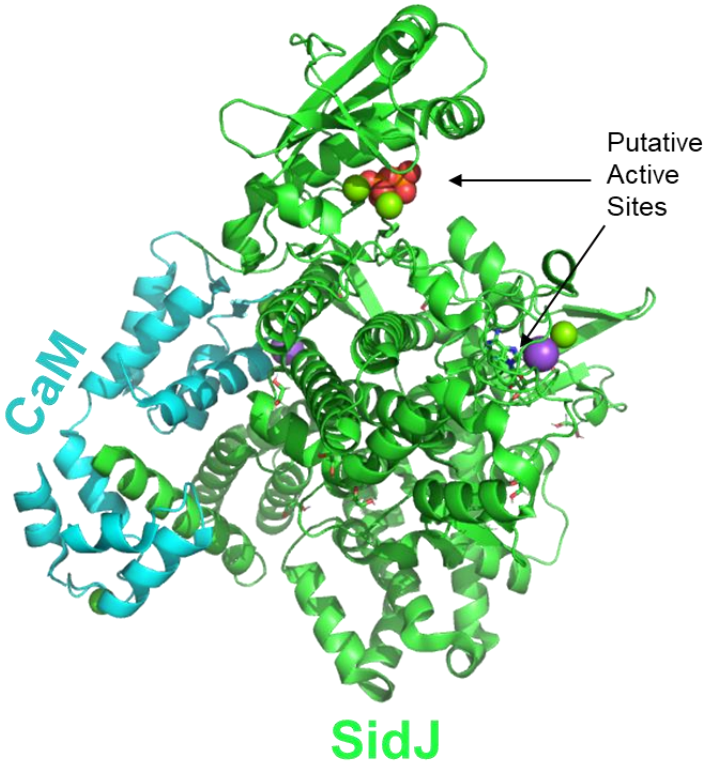


Figure 5-2: Structure of SidJ (green) in complex with CaM (cyan) with putative active sites pointed out (PDB: 6OQQ).

The activation of SidJ by a host factor is an important feature, otherwise SidE enzymes would be inactivated before their deployment by *Legionella* into the host cell. Only after SidJ's release is activation needed. This elegant method of utilizing a host factor to switch on a secreted bacterial effector is not without precedent in the microbiological world, with two noteworthy examples: one being the edema factor (EF) of anthrax toxin, a highly aggressive adenylate cyclase that is also activated by CaM binding.[119,120] Another example is the mARTX toxin family expressed in a variety of gram-negative bacteria, including *Vibrio* species.[121,122] These large proteins contain several effector domains including a cysteine protease that is activated by inositol hexakisphosphate, a molecule found in eukaryotic cells. This protease, upon activation, is necessary for processing the effector domains and allowing their release into the host. The dependence of host factors for effector activity thus may serve many roles such as preventing toxicity to the bacterium, or for appropriate timing of toxin autoprocessing.

The PR-linked ubiquitin deconjugating activity originally attributed to SidJ turned out to be catalyzed by different *Legionella* effectors, thereby providing another interesting layer of regulation of SidE enzymes. This activity was shown by two independent groups to be performed by a sequentially similar pair of effectors named DupA and DupB.[123,124] The structure of DupB (previously referred to as SdeD) showed that these enzymes resembled the PDE domain of the SidE family. However, they contained no other domains, and as such were able to target PR-linked ubiquitinated proteins with high affinity and catalyze the removal of Ub by a PDE activity, generating Ub-PR and unmodified substrate proteins as the products (Figure 5-3). This was shown to reverse both the molecular and biological effects of SidE enzymes. Due to the fact that the initial SidJ study reported a deconjugation of PR-linked ubiquitin, and that the SidJ was purified from *Legionella pneumophila* culture, it is likely that DupA or DupB were pulled down along with SidJ and are responsible for this observed activity.

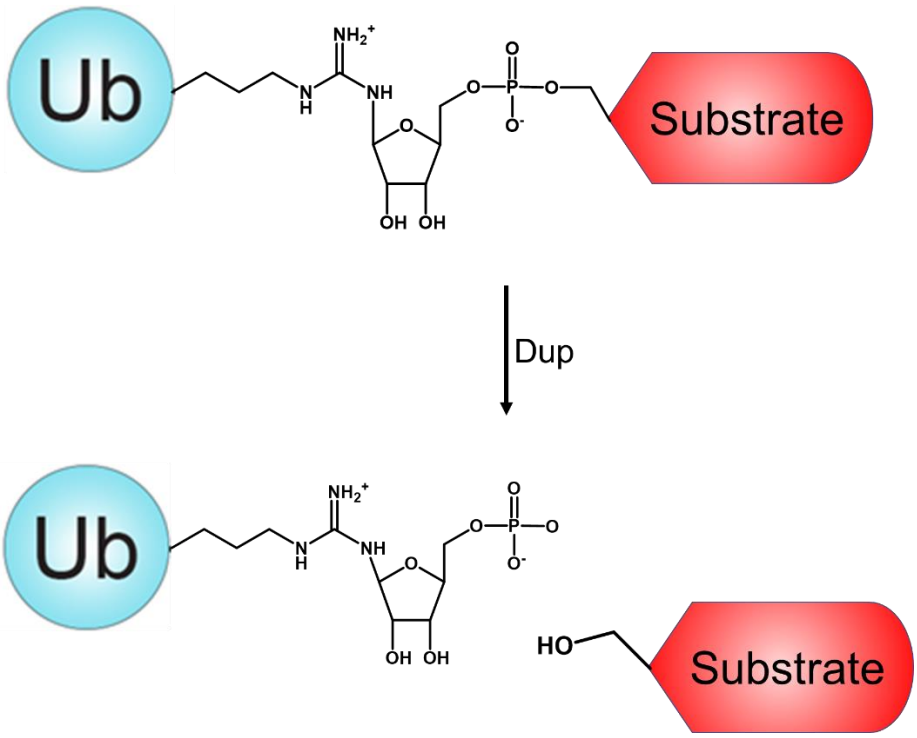


Figure 5-3: Scheme of deubiquitination of PR-linked ubiquitin conjugates by Dup enzymes.

MavC is also a highly regulated effector, with two other effectors found to modulate its activity. First, its paralog MvcA shares strong sequence identity (49%) and was also reported by Valleau et al to be able to deamidate Ub.[76] Subsequent investigations showed that it was unable to catalyze ubiquitination of Ube2N like MavC, and instead it efficiently hydrolyzed the Ub-Ube2N product (Figure 5-4).[125] This represented another interesting example of a structurally similar motif catalyzing the opposite enzymatic function, similar in principle to the Dups.

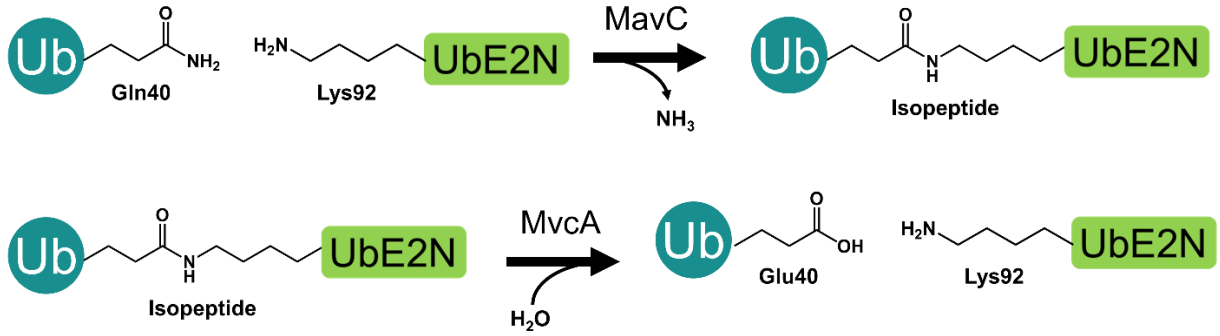


Figure 5-4: Schemes of MavC ubiquitination and MvcA deubiquitination reactions.

The second regulatory effector of MavC was Lpg2149, identified by Valleau et al.[76] This protein was shown to inhibit deamidation of both MavC and MvcA. Further biological experiments showed that this protein was expressed at the early exponential phase of *Legionella* growth, which takes place after infection, when the bacteria begin replication.

In this chapter, the fluorescent assay developed in Chapter 2 is used to study the inhibition of SdeA by SidJ, as well as the deconjugation of Ub-PR by a Dup enzyme. Additionally, biochemical analyses yielding insights into MvcA and Lpg2149-mediated regulation of MavC are also presented.

5.2 Materials and Methods

5.2.1 Cloning, Expression, and Purification of Recombinant Proteins

The 96-873 construct of SidJ was previously cloned by Dr. Michael Sheedlo (Das Lab, Purdue University) into the pGEX-6P1 vector between the BamHI and XhoI sites, using standard PCR-based methods. Full length SidJ served as a template for amplification of the desired construct.

SdeA$^{181\text{-}1000}$ was also cloned into pGEX-6P-1 as described in Chapter 2. Full length DupB, cloned in the pGEX-6P-1 vector, was obtained from the Luo Lab (Purdue University).

MvcA$^{12\text{-}398}$ and MvcA$^{\Delta INS}$ were previously cloned by John Hausman (Das Lab, Purdue University) into the pGEX-6P-1 vector. The C83A mutant was generated using site-directed mutagenesis with the mutation confirmed by DNA sequencing.

LPG2149 in the pQE30 vector was obtained from the Luo Lab (Purdue University).

Untagged wild-type ubiquitin (Ub) was cloned into the pRSET-A vector. SdeA and SidJ constructs used for this study were transformed into the Rosetta strain of *E. coli* for protein expression. MvcA constructs in this study were transformed into the BL-21 DE3 strain of *E. coli*.

All proteins in the pGEX vector were expressed with GST fused to its N-terminus and purified with GST affinity chromatography as described in earlier chapters.

Untagged Ub was purified via cation-exchange chromatography as described in earlier chapters.

His-LPG2149 was purified via the previously described Ni-NTA affinity chromatography method.

His-SUMO-Ube2N was purified via Ni-NTA affinity chromatography and the His-SUMO tag removed and subtracted as described in earlier chapters.

Human calmodulin (CaM) was received as a generous gift from Dr. Mark Wilson (University of Nebraska).

5.2.2 Synthesis of Fluorescent Peptide Substrates

The MAS fluorescent peptide was utilized in this study and was synthesized using standard Fmoc-based solid phase peptide synthesis methods, as described in Chapter 2. NHS-fluorescein (Sigma) was utilized for N-terminal labeling of the peptide. The peptide was then purified by reverse-phase HPLC, as described in Chapter 2.

5.2.3 In-Gel Deconjugation Assay

0.25 μM $SdeA^{181-1000}$, 100 μm ubiquitin and 10 μM fluorescent peptide were incubated in the presence of 100 μm NAD^+ and reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, and 1 mM DTT) for 10 min at room temperature. A negative control was included lacking NAD^+. After this ubiquitination step, DupB was added to a final concentration of 6 μm to one reaction, and an equivalent volume of reaction buffer was added to another reaction sample as a control. Reactions were quenched by 5X SDS/PAGE loading buffer and analyzed by SDS-PAGE as well as Coomassie Blue staining.

5.2.4 Fluorescence Polarization (FP) Assays

To measure the activity of SidJ via inhibition of SdeA activity, two steps were performed. First, 1 μm SidJ, 5 μm $SdeA^{181-1000}$, 5 μm CaM, 500 μm L-glutamate, 5 mM $MgCl_2$ and 1 mM ATP were combined in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mg/ml BSA). Reaction mixtures were incubated at 37°C for 10, 20, 30, 40 and 60 min. SidJ activity was quenched by adding 40 mM EDTA at each time point. The quenched reactions were then mixed with 100 μm ubiquitin, 100 μm NAD^+ and 10 μm fluorescein-labeled MAS peptide, with the final concentration of SdeA in these reactions being 0.25 μm. Reactions were initiated by adding 100

μm NAD^+. Fluorescence polarization was measured using a Cytation Multi-Mode Plate Reader (BioTek) using 485 nm excitation and 528 nm emission filters.

To assess the activity of Dups, FP assays as performed in Chapter 2 were conducted, using 0.25 μm $SdeA^{181-1000}$, 100 μm ubiquitin, 100 μm NAD^+ and 10 μm MAS peptide, with FP being continuously measured. After 20 minutes, DupB was added to a final concentration of 6 μm, and FP continued to be measured. To analyze the activity of Dups in cell lysate, the lysate of a strain of *Legionella pneumophila* lacking the SidE family (Luo Lab) was utilized in the place of purified DupB.

5.2.5 Generation and Purification of Ubiquitinated Ube2N Product

Ubiquitinated Ube2N (Ub-Ube2N) was synthesized and purified as described in Chapter 3. Briefly, $MavC^{1-462}$ at a final concentration was incubated with purified Ube2N and Ub for 3 hr at 37 ºC. Components of the ubiquitination reaction were separated using size-exclusion chromatography. Fractions corresponding to Ub-Ube2N were pooled and concentrated.

5.2.6 Deubiquitination Assays

To compare different constructs of LPG2148, deubiquitination of Ub-Ube2N was assayed by combining 25 μm of Ub-Ube2N with 1 μm of MvcA of varying constructs. Reactions were conducted for 1 hour at 37 º C in a reaction buffer of 50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT, quenched with 5X SDS-PAGE loading buffer, and analyzed by SDS-PAGE and stained with Coomassie Blue.

5.2.7 Deamidation Assays

Deamidation assays were conducted by combining 100 μm of wild type Ub with 1 μm of MvcA of various constructs. Reactions were conducted for 1 hour at 37 º C in a similar reaction buffer as in the deubiquitination assays and subjected to Native-PAGE analysis followed by staining with Coomassie Blue. The deamidated Q40E mutant of Ub was included as a standard for deamidation.

5.2.8 Kinetic Analysis of Deubiquitination

Michaelis-Menten kinetic analysis of the deubiquitination of Ub-Ube2N by MvcA was conducted by combining varying concentrations of Ub-Ube2N (10-65 μm) with 0.1 μm of MvcA in the same reaction buffer as in the previous deubiquitination assays. Reactions were conducted for 5 minutes at 37 º C, quenched with 5X SDS-PAGE loading buffer, and then analyzed by SDS-PAGE with Coomassie staining. Known concentrations of Ube2N were also run on the same gel to serve as standards to quantify Ube2N formation. Reactions were performed in triplicate.

Bands corresponding to Ube2N were analyzed by ImageJ, and a standard curve was generated using the values obtained from the Ube2N standard lanes. This standard curve was used to determine the concentration of Ube2N produced in each reaction lanes. These values allowed us to calculate initial velocity (Vo) of deubiquitination at each concentration of Ub-Ube2N. Vo values were then plotted against corresponding initial [Ub-Ube2N] and the data were fit to the Michaelis-Menten equation.

5.2.9 Lpg2149 Inhibition Assays

To determine inhibition of MavC-catalyzed ubiquitination by Lpg2149, reactions containing final concentrations of 0.25 μm MavC, 25 μm Ube2N and 100 μm ubiquitin were incubated for 10 minutes at 37 º C in the presence of 0-5 μm of Lpg2149. Lpg2149 was added to MavC in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT) and incubated on ice for 10 min before initiating the reaction with Ube2N and Ub. Reactions were quenched with 5X SDS-PAGE loading buffer, subjected to SDS-PAGE and stained with Coomassie Blue.

To determine inhibition of MvcA-catalyzed deubiquitination of Ub-Ube2N by Lpg2149, reactions containing final concentrations of 0.1 μm MvcA and 25 μm Ub were incubated for 10 minutes at 37 º C in the presence of 0-5 μm of Lpg2149. Lpg2149 was added to MvcA in reaction buffer (50 mM Tris pH 7.4, 100 mM NaCl, 1 mM DTT) and incubated on ice for 10 min before initiating the reaction with Ube2N and Ub. Reactions were quenched with 5X SDS-PAGE loading buffer, subjected to SDS-PAGE and stained with Coomassie Blue.

5.3 Results

5.3.1 Analysis of SidE regulators SidJ and DupB

SidJ has been identified as a regulator of SidE enzymes, effectively acting as an "off" switch by glutamylating a key catalytic residue in the mART domain. In the case of SdeA, Glu860 is targeted for glutamylation. One important question about this process pertains to the kinetic properties of this enzyme. Due to the small size of the glutamyl group added, it is difficult to measure activity of this enzyme by direct observation. In order to analyze SidJ activity in a time-dependent manner not requiring radiolabeling, we conducted SidJ reactions at different time points and utilized the FP assay from Chapter 2 to determine the effect of SidJ incubation on the activity of SdeA. We found that SdeA's activity decreased the longer it was reacted with SidJ. Using the initial rates of reaction, were able to generate a progress curve of SdeA inhibition which corresponds to glutamylation by SidJ (Figure 5-5A).

We also believed that our fluorogenic peptide substrate could also be deubiquitinated by Dup enzymes. By designing a two-step assay to first generate the Ub-PR-peptide before adding DupB, we observed the disappearance of the Ub-PR-peptide band when this reaction was analyzed by SDS-PAGE (Figure 5-5D). Furthermore, when the FP of this reaction mixture was monitored over time, we observed a clear and continuous decrease in FP upon the addition of DupB, returning to baseline levels (Figure 5-5B). This deubiquitinating activity was also found to exist in *L. pneumophila* lysate lacking the SidE family, where a decrease in FP was also observed (Figure 5-5C).

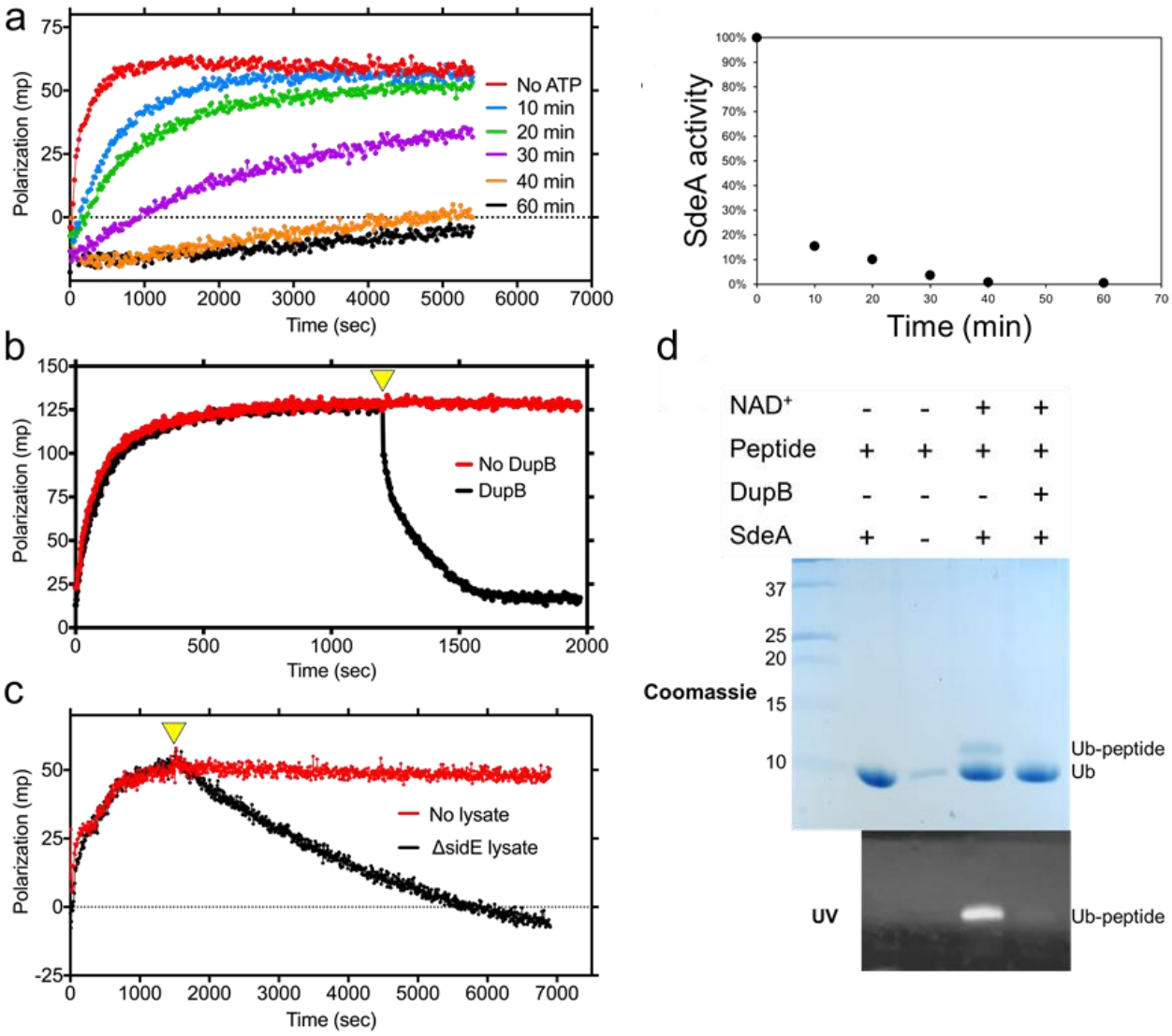


Figure 5-5: Analysis of SidE regulators using newly-developed fluorescence polarization assay. A) SidJ was incubated with SdeA for indicated time points, and ubiquitination was compared. Control SidJ reaction without ATP was tested. Initial rates were plotted to show inhibition of SdeA over time by SidJ. B) Two-step assay where ubiquitinated peptide was first generated, followed by DupB to show deubiquitinating activity. C) Two-step assay as in B), but with *Legionella pneumophila* lysate lacking SidE effectors instead of purified DupB. D) In-gel deconjugation assay, showing removal of Ub from fluorescent peptide by DupB.

5.3.2 MvcA requires the insertion domain for deubiquitination activity

MavC and MvcA both possess an insertion domain that distinguishes them from the Cif family of enzymes. In Chapter 3, it was shown that the insertion domain of MavC was essential for ubiquitinating activity. To elucidate the role of this domain in MvcA's ability to catalyze

deubiquitination of Ub-Ube2N, a construct lacking the insertion was compared to the wild type, as well as a catalytic C83A mutant. The result clearly showed that deubiquitinating activity was lost when the insertion domain was removed (Figure 5-6).

Similarly to MavC, MvcA also possesses ubiquitin deamidating activity *in vitro*.[76] Interestingly, when the MvcA$^{\Delta INS}$ construct was compared to the wild type enzyme in a deamidation assay, similar levels of Ub deamidation were observed between the two (Figure 5-6). This was in contrast to MavC$^{\Delta INS}$ construct which is notably weaker in deamidation than the wild type and suggested that the insertion domain of MvcA plays a more important role in Ube2N rather than Ub recognition.

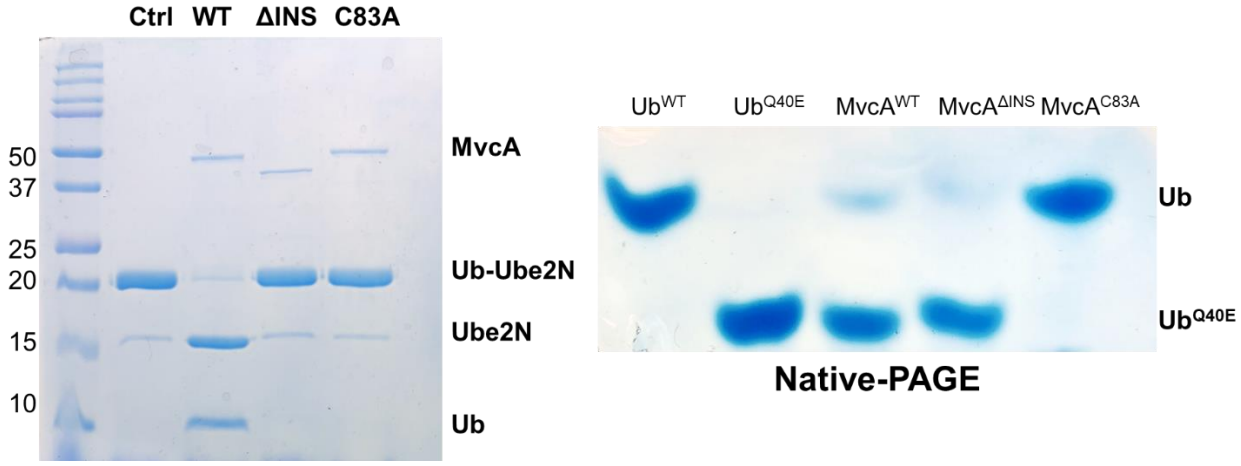


Figure 5-6: Analysis of MvcA insertion domain on deubiquitination and deamidation. Left panel: SDS-PAGE of ubiquitination reaction. Right panel: Native-PAGE of deamidation reaction.

5.3.3 Kinetic Analysis of MvcA

To examine the activity of MvcA in a more detailed manner and to compare it with canonical DUBs, a kinetic analysis of deubiquitination of Ub-Ube2N was performed. The results were fit to the Michaelis-Menten equation (Figure 5-7). With respect to deubiquitination, MvcA has a K_M of 54.68 μm and k_{cat} of 1.13 1/s. Interestingly, these values are comparable to those of standard DUBs that hydrolyze Ub linked via its C-terminus, including the DUB domain of SdeA.

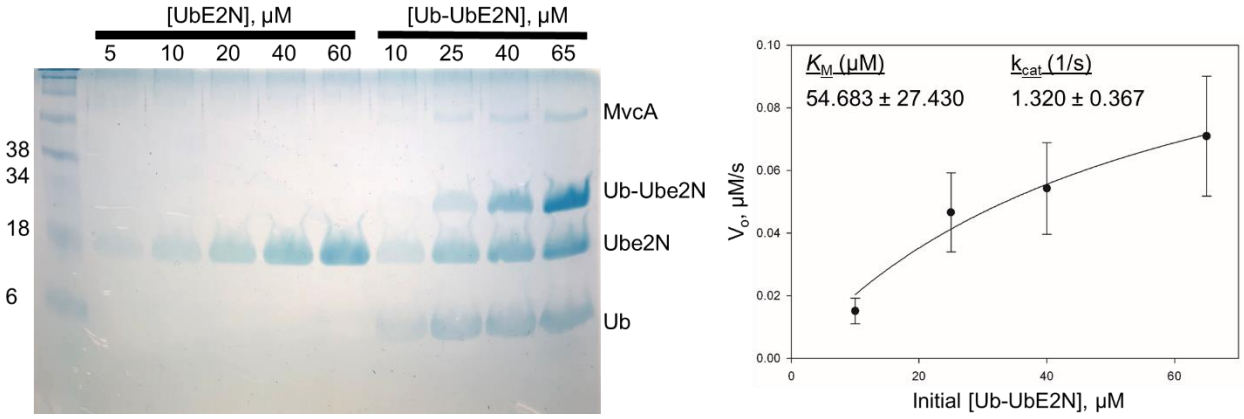


Figure 5-7: Michaelis-Menten kinetic analysis. A) SDS-PAGE gel of deubiquitination reactions with Ube2N standards. B) Plots of initial velocity versus initial [Ub-Ube2N], fit to the Michaelis-Menten equation with calculated kinetic parameters.

5.3.4 Lpg2149 is an Inhibitor of MavC and MvcA Activity

Previous studies identified Lpg2149 as an inhibitor of the deamidating activities of both MavC and MvcA.[76] In light of the findings that MavC and MvcA are a ubiquitin ligase/hydrolase pair, we hypothesized that Lpg2149 was also an inhibitor of these newly reported activities. To that end, Lpg2149 was incubated with MavC and MvcA at various concentrations. Strikingly, Lpg219 inhibited both ubiquitinating and deubiquitinating activities of MavC and MvcA respectively in a dose-dependent manner. Furthermore, the IC_{50} of Lpg2149 was a more potent inhibitor of MavC, with its IC_{50} calculated to be ~0.18 μm (apparent Ki of 0.060 μm) against MavC and ~1.51 μm (apparent Ki of 1.46 μm) against MvcA (Figure 5-8).

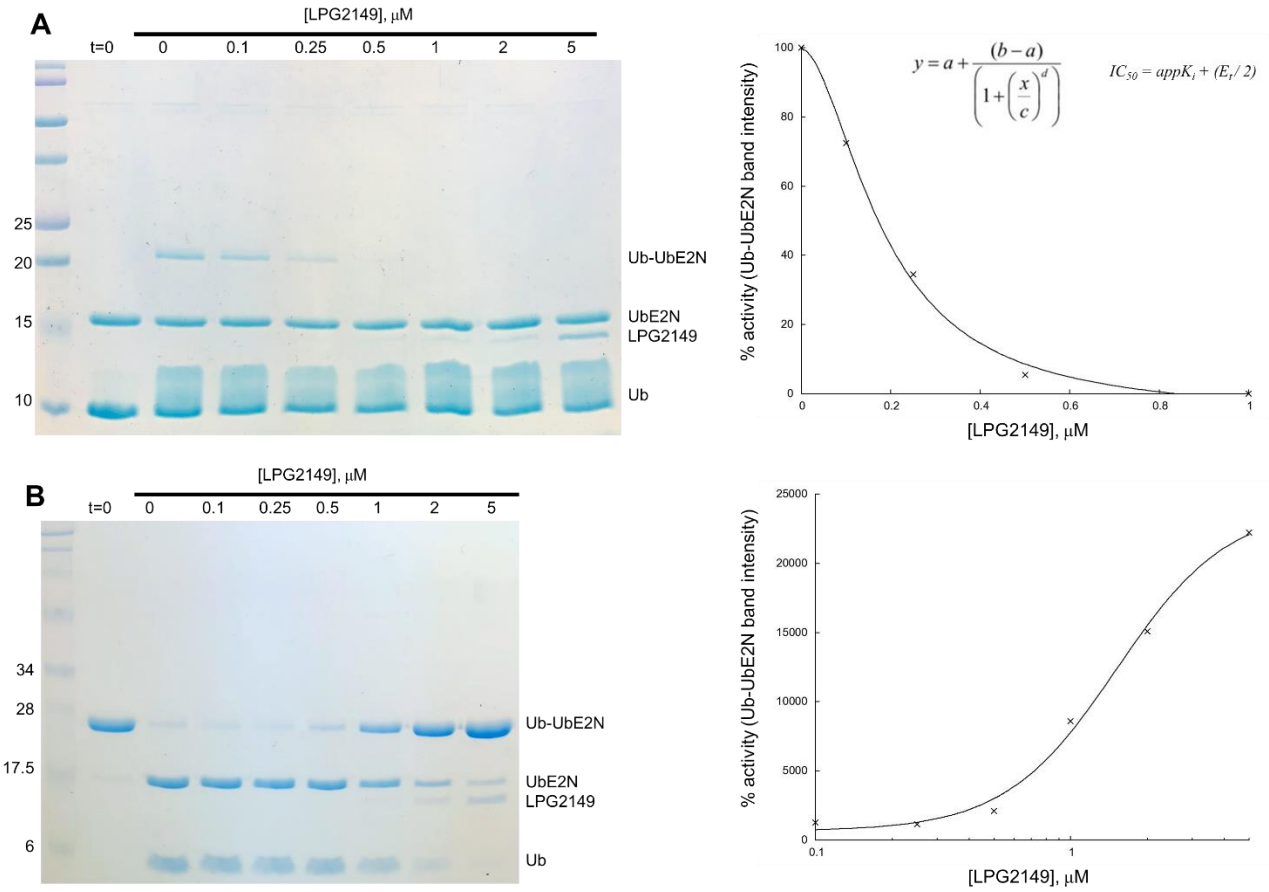


Figure 5-8: Analysis of lpg2149 inhibition of A) MavC ubiquitination activity and B) MvcA deubiquitination activity. Dose-response curves are also given along with relevant equations used to fit curves and calculate inhibition data.

5.4 Discussion

5.4.1 Regulation of SidEs

It is clear that *Legionella pneumophila* utilizes at least two types of metaeffectors to provide a check to the SidE family. SidJ acts to turn these enzymes off, and the Dups reverse their overall effects. By harnessing new biochemical assays, further studies of regulation of SidE enzymes is now facilitated. For example, one could imagine performing a detailed comparison of DupA and DupB in terms of their activity, substrate preference, and binding via this assay, which may shed light into why *Legionella* contains two seemingly redundant enzymes in its genome. Also, one could also examine SidJ's effects on different members of the SidE family in a more quantitative fashion, as well as perform further mechanistic studies on SidJ's glutamylation. It is

proposed that SidJ acts first by using ATP to generate an acyl-adenylated intermediate (Figure 5-1). This intermediate was detected in a previous experiment,[113] but it is still not clear which active site is responsible for which step in the reaction. Also it is evident that SidJ possesses some preference in its target residue, as Glu860 but not Glu862 of SdeA has been found to be glutamylated. However, whether or not SidJ can also modify other proteins and the effect of such modification remains to be seen.

5.4.2 Regulation of MavC/MvcA

Fascinatingly, MavC is also regulated by two metaeffector systems. One is MvcA, a paralog that catalyzes the reverse reaction of MavC, hydrolyzing the isopeptide crosslink through a similar catalytic core. Despite earlier studies showing that MvcA lacks the ability to bind to Ube2N, it is clear that the insertion domain is important for Ube2N recognition. Deletion of this domain causes MvcA to lose its ability to react with Ub-Ube2N but not Ub. It is possible that the insertion has a transient affinity towards Ube2N that is important for initial substrate recognition, explaining the lack of observed co-precipitation between the two proteins.

The efficiency of MvcA as a deubiquitinase was not known. These studies reveal that MvcA possesses a k_{cat}/K_M of approximately 0.020 $s^{-1}\mu m^{-1}$ which is markedly higher than the DUB domain of SdeA which was determined to be 0.0066 $s^{-1}\mu m^{-1}$ against its preferred substrate of K63-linked diubiquitin.[45] This strengthens the case that the deubiquitinating activity of MvcA is biologically relevant, as it occurs at a level comparable to that of other DUBs. Nonetheless, it is surprising that the Michaelis constant K_M of MvcA was found to be 54.68 μm, as this value is significantly greater than the amount of Ube2N that is typically present in the host cell (~2 μm). SdeA likely overcomes its relatively high K_M of ~200 μm by virtue of its spatial positioning. It is well known that SidE proteins localize to the LCV, and that K63 polyUb chains build up on the LCV surface. This increased local concentration may allow SdeA to play an impactful role as an effector. This then may raise the possibility that MvcA may employ a similar strategy to be able to counteract MavC's effect.

On a conceptual level, deubiquitination of MavC-catalyzed linkages by MvcA is reminiscent of the Dups, since it involves a pair of effectors with similar structure but opposite activities. Shin and colleagues explored the basis of this difference in activity, demonstrating that

weakening the affinity of DupA for Ub through mutagenesis caused it to weakly promote ubiquitin transfer to substrates. It was concluded that a longer residence time of Ub increased the likelihood of a hydrolysis reaction. While it is tempting to speculate that the same principles govern MvcA catalyzing an opposing reaction to MavC, additional work will be necessary to provide a more concrete answer.

Lpg2149's inhibitory role was further extended to the ubiquitinating and deubiquitinating activities of MavC and MvcA respectively. This is an intriguing example of not only a regulatory effector, but one that can also inhibit another regulatory effector in MvcA, further increasing the complexity of identified effector-effector interactions by *Legionella*. A noteworthy conclusion from the inhibition studies performed here is that this effector binds relatively tightly to MavC and MvcA, especially in comparison to their substrates. It is not surprising, therefore, that Lpg2149 was not found to be expressed by *L. pneumophila* until after infection had taken place and the pathogen had begun to divide. This is because expression of Lpg2149 at earlier stages would likely cause complex formation and inhibition of MavC and MvcA before being translocated into the host cell. The need to inhibit MvcA along with MavC is interesting, as MvcA appears to only serve a regulatory role itself. However, this inhibition of MvcA may serve to prevent deamidated Ub from building up in the host cell which is known to attenuate host ubiquitination signaling. Since deamidated Ub is not detected in cells infected with *Legionella*, it is reasonable to speculate that Lpg2149 may play a role in this. A measurement of Ub deamidation levels in cells infected with a ΔLpg2149 strain of *L. pneumophila* may provide insights into this idea. The structural basis of inhibition, and whether or not this is a competitive inhibitor also remains an open question.

5.4.3 Conclusions

Many effector and metaeffector pairs in the *Legionella* genome have been identified by biological studies as well as broader genetic screens, yet many of their their precise actions have remained unclear. Today a rich upsurge in reported results over the past few years has uncovered the molecular mechanisms of regulation of these ubiquitinating effectors. The biochemical studies presented in this chapter further our understanding of these effectors by providing insights into their effectiveness as enzymes or inhibitors, and also describe new methodologies for further investigations.

REFERENCES

(1) Hershko, A.; Ciechanover, A. The Ubiquitin System. *Annu. Rev. Biochem.* **1998**, *67* (1), 425–479. https://doi.org/10.1146/annurev.biochem.67.1.425.

(2) Komander, D.; Rape, M. The Ubiquitin Code. *Annu. Rev. Biochem.* **2012**, *81*, 203–229. https://doi.org/10.1146/annurev-biochem-060310-170328.

(3) Yau, R.; Rape, M. The Increasing Complexity of the Ubiquitin Code. *Nat. Cell Biol.* **2016**, *18* (6), 579–586. https://doi.org/10.1038/ncb3358.

(4) 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. https://doi.org/10.1126/science.2538923.

(5) Sloper-Mould, K. E.; Jemc, J. C.; Pickart, C. M.; Hicke, L. Distinct Functional Surface Regions on Ubiquitin. *J. Biol. Chem.* **2001**, *276* (32), 30483–30489. https://doi.org/10.1074/jbc.M103248200.

(6) Mukhopadhyay, D.; Riezman, H. Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling. *Science* **2007**, *315* (5809), 201–205. https://doi.org/10.1126/science.1127085.

(7) Henne, W. M.; Buchkovich, N. J.; Emr, S. D. The ESCRT Pathway. *Dev. Cell* **2011**, *21* (1), 77–91. https://doi.org/10.1016/j.devcel.2011.05.015.

(8) Hoege, C.; Pfander, B.; Moldovan, G.-L.; Pyrowolakis, G.; Jentsch, S. RAD6-Dependent DNA Repair Is Linked to Modification of PCNA by Ubiquitin and SUMO. *Nature* **2002**, *419* (6903), 135–141. https://doi.org/10.1038/nature00991.

(9) Deol, K. K.; Lorenz, S.; Strieter, E. R. Enzymatic Logic of Ubiquitin Chain Assembly. *Front. Physiol.* **2019**, *10*. https://doi.org/10.3389/fphys.2019.00835.

(10) Komander, D.; Clague, M. J.; Urbé, S. Breaking the Chains: Structure and Function of the Deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **2009**, *10* (8), 550–563. https://doi.org/10.1038/nrm2731.

(11) Clague, M. J.; Urbé, S.; Komander, D. Breaking the Chains: Deubiquitylating Enzyme Specificity Begets Function. *Nat. Rev. Mol. Cell Biol.* **2019**, *20* (6), 338–352. https://doi.org/10.1038/s41580-019-0099-1.

(12) Kwasna, D.; Abdul Rehman, S. A.; Natarajan, J.; Matthews, S.; Madden, R.; De Cesare, V.; Weidlich, S.; Virdee, S.; Ahel, I.; Gibbs-Seymour, I.; Kulathu, Y. Discovery and Characterization of ZUFSP/ZUP1, a Distinct Deubiquitinase Class Important for Genome Stability. *Mol. Cell* **2018**, *70* (1), 150–164.e6. https://doi.org/10.1016/j.molcel.2018.02.023.

(13) Brown, S.; Niimi, A.; Lehmann, A. R. Ubiquitination and Deubiquitination of PCNA in Response to Stalling of the Replication Fork. *Cell Cycle* **2009**, *8* (5), 689–692. https://doi.org/10.4161/cc.8.5.7707.

(14) Ronau, J. A.; Beckmann, J. F.; Hochstrasser, M. Substrate Specificity of the Ubiquitin and Ubl Proteases. *Cell Res.* **2016**, *26* (4), 441–456. https://doi.org/10.1038/cr.2016.38.

(15) Metzger, M. B.; Pruneda, J. N.; Klevit, R. E.; Weissman, A. M. RING-Type E3 Ligases: Master Manipulators of E2 Ubiquitin-Conjugating Enzymes and Ubiquitination. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2014**, *1843* (1), 47–60. https://doi.org/10.1016/j.bbamcr.2013.05.026.

(16) Lorenz, S. Structural Mechanisms of HECT-Type Ubiquitin Ligases. *Biol. Chem.* **2018**, *399* (2), 127–145. https://doi.org/10.1515/hsz-2017-0184.
(17) Ries, L. K.; Sander, B.; Deol, K. K.; Letzelter, M.-A.; Strieter, E. R.; Lorenz, S. Analysis of Ubiquitin Recognition by the HECT Ligase E6AP Provides Insight into Its Linkage Specificity. *J. Biol. Chem.* **2019**, *294* (15), 6113–6129. https://doi.org/10.1074/jbc.RA118.007014.
(18) Huibregtse, J. M.; Scheffner, M.; Beaudenon, S.; Howley, P. M. A Family of Proteins Structurally and Functionally Related to the E6-AP Ubiquitin-Protein Ligase. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (7), 2563–2567.
(19) Wenzel, D. M.; Lissounov, A.; Brzovic, P. S.; Klevit, R. E. UBCH7 Reactivity Profile Reveals Parkin and HHARI to Be RING/HECT Hybrids. *Nature* **2011**, *474* (7349), 105–108. https://doi.org/10.1038/nature09966.
(20) Davies, C. W.; Das, C. HHARI Is One HECT of a RING. *Structure* **2013**, *21* (6), 872–874. https://doi.org/10.1016/j.str.2013.05.003.
(21) Walden, H.; Rittinger, K. RBR Ligase–mediated Ubiquitin Transfer: A Tale with Many Twists and Turns. *Nat. Struct. Mol. Biol.* **2018**, *25* (6), 440–445. https://doi.org/10.1038/s41594-018-0063-3.
(22) VanDemark, A. P.; Hofmann, R. M.; Tsui, C.; Pickart, C. M.; Wolberger, C. Molecular Insights into Polyubiquitin Chain Assembly: Crystal Structure of the Mms2/Ubc13 Heterodimer. *Cell* **2001**, *105* (6), 711–720. https://doi.org/10.1016/S0092-8674(01)00387-7.
(23) Barandun, J.; Delley, C. L.; Weber-Ban, E. The Pupylation Pathway and Its Role in Mycobacteria. *BMC Biol.* **2012**, *10* (1), 95. https://doi.org/10.1186/1741-7007-10-95.
(24) 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. https://doi.org/10.1016/j.bbrc.2017.01.037.
(25) Fuchs, A. C. D.; Maldoner, L.; Wojtynek, M.; Hartmann, M. D.; Martin, J. Rpn11-Mediated Ubiquitin Processing in an Ancestral Archaeal Ubiquitination System. *Nat. Commun.* **2018**, *9* (1), 1–12. https://doi.org/10.1038/s41467-018-05198-1.
(26) Nunoura, T.; Takaki, Y.; Kakuta, J.; Nishi, S.; Sugahara, J.; Kazama, H.; Chee, G.-J.; Hattori, M.; Kanai, A.; Atomi, H.; Takai, K.; Takami, H. Insights into the Evolution of Archaea and Eukaryotic Protein Modifier Systems Revealed by the Genome of a Novel Archaeal Group. *Nucleic Acids Res.* **2011**, *39* (8), 3204–3223. https://doi.org/10.1093/nar/gkq1228.
(27) James, R. H.; Caceres, E. F.; Escasinas, A.; Alhasan, H.; Howard, J. A.; Deery, M. J.; Ettema, T. J. G.; Robinson, N. P. Functional Reconstruction of a Eukaryotic-like E1/E2/(RING) E3 Ubiquitylation Cascade from an Uncultured Archaeon. *Nat. Commun.* **2017**, *8* (1), 1–15. https://doi.org/10.1038/s41467-017-01162-7.
(28) Grau-Bové, X.; Sebé-Pedrós, A.; Ruiz-Trillo, I. The Eukaryotic Ancestor Had a Complex Ubiquitin Signaling System of Archaeal Origin. *Mol. Biol. Evol.* **2015**, *32* (3), 726–739. https://doi.org/10.1093/molbev/msu334.
(29) Ashida, H.; Kim, M.; Sasakawa, C. Exploitation of the Host Ubiquitin System by Human Bacterial Pathogens. *Nat. Rev. Microbiol.* **2014**, *12* (6), 399–413. https://doi.org/10.1038/nrmicro3259.

(30) Perrett, C. A.; Lin, D. Y.; Zhou, D. Interactions of Bacterial Proteins with Host Eukaryotic Ubiquitin Pathways. *Front. Microbiol.* **2011**, *2*. https://doi.org/10.3389/fmicb.2011.00143.
(31) Molecular Characterization of LubX: Functional Divergence of the U-Box Fold by Legionella Pneumophila. *Structure* **2015**, *23* (8), 1459–1469. https://doi.org/10.1016/j.str.2015.05.020.
(32) Kamanova, J.; Sun, H.; Lara-Tejero, M.; Galán, J. E. The Salmonella Effector Protein SopA Modulates Innate Immune Responses by Targeting TRIM E3 Ligase Family Members. *PLoS Pathog.* **2016**, *12* (4). https://doi.org/10.1371/journal.ppat.1005552.
(33) Sheng, X.; You, Q.; Zhu, H.; Chang, Z.; Li, Q.; Wang, H.; Wang, C.; Wang, H.; Hui, L.; Du, C.; Xie, X.; Zeng, R.; Lin, A.; Shi, D.; Ruan, K.; Yan, J.; Gao, G. F.; Shao, F.; Hu, R. Bacterial Effector NleL Promotes Enterohemorrhagic E. Coli-Induced Attaching and Effacing Lesions by Ubiquitylating and Inactivating JNK. *PLOS Pathog.* **2017**, *13* (7), e1006534. https://doi.org/10.1371/journal.ppat.1006534.
(34) Price, C. T. D.; Al-Quadan, T.; Santic, M.; Rosenshine, I.; Abu Kwaik, Y. Host Proteasomal Degradation Generates Amino Acids Essential for Intracellular Bacterial Growth. *Science* **2011**, *334* (6062), 1553–1557. https://doi.org/10.1126/science.1212868.
(35) Pisano, A.; Albano, F.; Vecchio, E.; Renna, M.; Scala, G.; Quinto, I.; Fiume, G. Revisiting Bacterial Ubiquitin Ligase Effectors: Weapons for Host Exploitation. *Int. J. Mol. Sci.* **2018**, *19* (11). https://doi.org/10.3390/ijms19113576.
(36) Huibregtse, J.; Rohde, J. R. Hell's BELs: Bacterial E3 Ligases That Exploit the Eukaryotic Ubiquitin Machinery. *PLoS Pathog.* **2014**, *10* (8). https://doi.org/10.1371/journal.ppat.1004255.
(37) Hsu, F.; Luo, X.; Qiu, J.; Teng, Y.-B.; Jin, J.; Smolka, M. B.; Luo, Z.-Q.; Mao, Y. The Legionella Effector SidC Defines a Unique Family of Ubiquitin Ligases Important for Bacterial Phagosomal Remodeling. *Proc. Natl. Acad. Sci.* **2014**, *111* (29), 10538–10543. https://doi.org/10.1073/pnas.1402605111.
(38) Wasilko, D. J.; Huang, Q.; Mao, Y. Insights into the ubiquitin transfer cascade catalyzed by the Legionella effector SidC https://elifesciences.org/articles/36154 (accessed Mar 11, 2020). https://doi.org/10.7554/eLife.36154.
(39) Keszei, A. F. A.; Sicheri, F. Mechanism of Catalysis, E2 Recognition, and Autoinhibition for the IpaH Family of Bacterial E3 Ubiquitin Ligases. *Proc. Natl. Acad. Sci.* **2017**, *114* (6), 1311–1316. https://doi.org/10.1073/pnas.1611595114.
(40) Hermanns, T.; Hofmann, K. Bacterial DUBs: Deubiquitination beyond the Seven Classes. *Biochem. Soc. Trans.* **2019**, *47* (6), 1857–1866. https://doi.org/10.1042/BST20190526.
(41) Bailey-Elkin, B. A.; Kasteren, P. B. van; Snijder, E. J.; Kikkert, M.; Mark, B. L. Viral OTU Deubiquitinases: A Structural and Functional Comparison. *PLOS Pathog.* **2014**, *10* (3), e1003894. https://doi.org/10.1371/journal.ppat.1003894.
(42) Báez-Santos, Y. M.; St John, S. E.; Mesecar, A. D. The SARS-Coronavirus Papain-like Protease: Structure, Function and Inhibition by Designed Antiviral Compounds. *Antiviral Res.* **2015**, *115*, 21–38. https://doi.org/10.1016/j.antiviral.2014.12.015.
(43) Matthews, K.; Schäfer, A.; Pham, A.; Frieman, M. The SARS Coronavirus Papain like Protease Can Inhibit IRF3 at a Post Activation Step That Requires Deubiquitination Activity. *Virol. J.* **2014**, *11* (1), 209. https://doi.org/10.1186/s12985-014-0209-9.

(44) Deng, X.; Chen, Y.; Mielech, A. M.; Hackbart, M.; Kesely, K. R.; Mettelman, R. C.; O'Brien, A.; Chapman, M. E.; Mesecar, A. D.; Baker, S. C. Structure-Guided Mutagenesis Alters Deubiquitinating Activity and Attenuates Pathogenesis of a Murine Coronavirus. *bioRxiv* **2020**, 782409. https://doi.org/10.1101/782409.

(45) Sheedlo, M. J.; Qiu, J.; Tan, Y.; Paul, L. N.; Luo, Z.-Q.; Das, C. Structural Basis of Substrate Recognition by a Bacterial Deubiquitinase Important for Dynamics of Phagosome Ubiquitination. *Proc. Natl. Acad. Sci.* **2015**, *112* (49), 15090–15095. https://doi.org/10.1073/pnas.1514568112.

(46) Rytkönen, A.; Poh, J.; Garmendia, J.; Boyle, C.; Thompson, A.; Liu, M.; Freemont, P.; Hinton, J. C. D.; Holden, D. W. SseL, a Salmonella Deubiquitinase Required for Macrophage Killing and Virulence. *Proc. Natl. Acad. Sci.* **2007**, *104* (9), 3502–3507. https://doi.org/10.1073/pnas.0610095104.

(47) Makarova, K. S.; Aravind, L.; Koonin, E. V. A Novel Superfamily of Predicted Cysteine Proteases from Eukaryotes, Viruses and Chlamydia Pneumoniae. *Trends Biochem. Sci.* **2000**, *25* (2), 50–52. https://doi.org/10.1016/S0968-0004(99)01530-3.

(48) Pruneda, J. N.; Durkin, C. H.; Geurink, P. P.; Ovaa, H.; Santhanam, B.; Holden, D. W.; Komander, D. The Molecular Basis for Ubiquitin and Ubiquitin-like Specificities in Bacterial Effector Proteases. *Mol. Cell* **2016**, *63* (2), 261–276. https://doi.org/10.1016/j.molcel.2016.06.015.

(49) Wan, M.; Wang, X.; Huang, C.; Xu, D.; Wang, Z.; Zhou, Y.; Zhu, Y. A Bacterial Effector Deubiquitinase Specifically Hydrolyses Linear Ubiquitin Chains to Inhibit Host Inflammatory Signalling. *Nat. Microbiol.* **2019**, *4* (8), 1282–1293. https://doi.org/10.1038/s41564-019-0454-1.

(50) Kubori, T.; Kitao, T.; Ando, H.; Nagai, H. LotA, a Legionella Deubiquitinase, Has Dual Catalytic Activity and Contributes to Intracellular Growth. *Cell. Microbiol.* **2018**, *20* (7), e12840. https://doi.org/10.1111/cmi.12840.

(51) Ma, K.; Zhen, X.; Zhou, B.; Gan, N.; Cao, Y.; Fan, C.; Ouyang, S.; Luo, Z.-Q.; Qiu, J. The Bacterial Deubiquitinase Ceg23 Regulates the Association of Lys-63-Linked Polyubiquitin Molecules on the Legionella Phagosome. *J. Biol. Chem.* **2020**, *295* (6), 1646–1657. https://doi.org/10.1074/jbc.RA119.011758.

(52) Kubori, T.; Kitao, T.; Nagai, H. Emerging Insights into Bacterial Deubiquitinases. *Curr. Opin. Microbiol.* **2019**, *47*, 14–19. https://doi.org/10.1016/j.mib.2018.10.001.

(53) Cui, J.; Yao, Q.; Li, S.; Ding, X.; Lu, Q.; Mao, H.; Liu, L.; Zheng, N.; Chen, S.; Shao, F. Glutamine Deamidation and Dysfunction of Ubiquitin/NEDD8 Induced by a Bacterial Effector Family. *Science* **2010**, *329* (5996), 1215–1218. https://doi.org/10.1126/science.1193844.

(54) Sanada, T.; Kim, M.; Mimuro, H.; Suzuki, M.; Ogawa, M.; Oyama, A.; Ashida, H.; Kobayashi, T.; Koyama, T.; Nagai, S.; Shibata, Y.; Gohda, J.; Inoue, J.; Mizushima, T.; Sasakawa, C. The Shigella Flexneri Effector OspI Deamidates UBC13 to Dampen the Inflammatory Response. *Nature* **2012**, *483* (7391), 623–626. https://doi.org/10.1038/nature10894.

(55) Nishide, A.; Kim, M.; Takagi, K.; Himeno, A.; Sanada, T.; Sasakawa, C.; Mizushima, T. Structural Basis for the Recognition of Ubc13 by the Shigella Flexneri Effector OspI. *J. Mol. Biol.* **2013**, *425* (15), 2623–2631. https://doi.org/10.1016/j.jmb.2013.02.037.

(56) Qiu, J.; Sheedlo, M. J.; Yu, K.; Tan, Y.; Nakayasu, E. S.; Das, C.; Liu, X.; Luo, Z.-Q. Ubiquitination Independent of E1 and E2 Enzymes by Bacterial Effectors. *Nature* **2016**, *533* (7601), 120–124. https://doi.org/10.1038/nature17657.

(57) Luo, Z.-Q.; Isberg, R. R. Multiple Substrates of the Legionella Pneumophila Dot/Icm System Identified by Interbacterial Protein Transfer. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (3), 841–846. https://doi.org/10.1073/pnas.0304916101.

(58) Bardill, J. P.; Miller, J. L.; Vogel, J. P. IcmS-Dependent Translocation of SdeA into Macrophages by the Legionella Pneumophila Type IV Secretion System. *Mol. Microbiol.* **2005**, *56* (1), 90–103. https://doi.org/10.1111/j.1365-2958.2005.04539.x.

(59) Havey, J. C.; Roy, C. R. Toxicity and SidJ-Mediated Suppression of Toxicity Require Distinct Regions in the SidE Family of Legionella Pneumophila Effectors. *Infect. Immun.* **2015**, *83* (9), 3506–3514. https://doi.org/10.1128/IAI.00497-15.

(60) Jeong, K. C.; Sexton, J. A.; Vogel, J. P. Spatiotemporal Regulation of a Legionella Pneumophila T4SS Substrate by the Metaeffector SidJ. *PLOS Pathog.* **2015**, *11* (3), e1004695. https://doi.org/10.1371/journal.ppat.1004695.

(61) Palazzo, L.; Mikolčević, P.; Mikoč, A.; Ahel, I. ADP-Ribosylation Signalling and Human Disease. *Open Biol. 9* (4), 190041. https://doi.org/10.1098/rsob.190041.

(62) Simon, N. C.; Aktories, K.; Barbieri, J. T. Novel Bacterial ADP-Ribosylating Toxins: Structure and Function. *Nat. Rev. Microbiol.* **2014**, *12* (9), 599–611. https://doi.org/10.1038/nrmicro3310.

(63) Tsurumura, T.; Tsumori, Y.; Qiu, H.; Oda, M.; Sakurai, J.; Nagahama, M.; Tsuge, H. Arginine ADP-Ribosylation Mechanism Based on Structural Snapshots of Iota-Toxin and Actin Complex. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (11), 4267–4272. https://doi.org/10.1073/pnas.1217227110.

(64) Bhogaraju, S.; Kalayil, S.; Liu, Y.; Bonn, F.; Colby, T.; Matic, I.; Dikic, I. Phosphoribosylation of Ubiquitin Promotes Serine Ubiquitination and Impairs Conventional Ubiquitination. *Cell* **2016**, *167* (6), 1636–1649.e13. https://doi.org/10.1016/j.cell.2016.11.019.

(65) Kotewicz, K. M.; Ramabhadran, V.; Sjoblom, N.; Vogel, J. P.; Haenssler, E.; Zhang, M.; Behringer, J.; Scheck, R. A.; Isberg, R. R. A Single Legionella Effector Catalyzes a Multistep Ubiquitination Pathway to Rearrange Tubular Endoplasmic Reticulum for Replication. *Cell Host Microbe* **2017**, *21* (2), 169–181. https://doi.org/10.1016/j.chom.2016.12.007.

(66) Wang, Y.; Shi, M.; Feng, H.; Zhu, Y.; Liu, S.; Gao, A.; Gao, P. Structural Insights into Non-Canonical Ubiquitination Catalyzed by SidE. *Cell* **2018**, *173* (5), 1231–1243.e16. https://doi.org/10.1016/j.cell.2018.04.023.

(67) Akturk, A.; Wasilko, D. J.; Wu, X.; Liu, Y.; Zhang, Y.; Qiu, J.; Luo, Z.-Q.; Reiter, K. H.; Brzovic, P. S.; Klevit, R. E.; Mao, Y. The Mechanism of Phosphoribosyl-Ubiquitination Mediated by a Single Legionella Effector. *Nature* **2018**, *557* (7707), 729–733. https://doi.org/10.1038/s41586-018-0147-6.

(68) Dong, Y.; Mu, Y.; Xie, Y.; Zhang, Y.; Han, Y.; Zhou, Y.; Wang, W.; Liu, Z.; Wu, M.; Wang, H.; Pan, M.; Xu, N.; Xu, C.-Q.; Yang, M.; Fan, S.; Deng, H.; Tan, T.; Liu, X.; Liu, L.; Li, J.; Wang, J.; Fang, X.; Feng, Y. Structural Basis of Ubiquitin Modification by the Legionella Effector SdeA. *Nature* **2018**, *557* (7707), 674–678. https://doi.org/10.1038/s41586-018-0146-7.

(69) Kalayil, S.; Bhogaraju, S.; Bonn, F.; Shin, D.; Liu, Y.; Gan, N.; Basquin, J.; Grumati, P.; Luo, Z.-Q.; Dikic, I. Insights into Catalysis and Function of Phosphoribosyl-Linked Serine Ubiquitination. *Nature* **2018**, *557* (7707), 734–738. https://doi.org/10.1038/s41586-018-0145-8.

(70) Kim, L.; Kwon, D. H.; Kim, B. H.; Kim, J.; Park, M. R.; Park, Z.-Y.; Song, H. K. Structural and Biochemical Study of the Mono-ADP-Ribosyltransferase Domain of SdeA, a Ubiquitylating/Deubiquitylating Enzyme from Legionella Pneumophila. *J. Mol. Biol.* **2018**, *430* (17), 2843–2856. https://doi.org/10.1016/j.jmb.2018.05.043.

(71) Puvar, K.; Luo, Z.-Q.; Das, C. Uncovering the Structural Basis of a New Twist in Protein Ubiquitination. *Trends Biochem. Sci.* **2019**, *44* (5), 467–477. https://doi.org/10.1016/j.tibs.2018.11.006.

(72) Gan, N.; Nakayasu, E. S.; Hollenbeck, P. J.; Luo, Z.-Q. Legionella Pneumophila Inhibits Immune Signalling via MavC-Mediated Transglutaminase-Induced Ubiquitination of UBE2N. *Nat. Microbiol.* **2019**, *4* (1), 134. https://doi.org/10.1038/s41564-018-0282-8.

(73) Lawrence, T. The Nuclear Factor NF-κB Pathway in Inflammation. *Cold Spring Harb. Perspect. Biol.* **2009**, *1* (6). https://doi.org/10.1101/cshperspect.a001651.

(74) Griffin, M.; Casadio, R.; Bergamini, C. M. Transglutaminases: Nature's Biological Glues. *Biochem. J.* **2002**, *368* (Pt 2), 377–396. https://doi.org/10.1042/BJ20021234.

(75) Lorand, L.; Graham, R. M. Transglutaminases: Crosslinking Enzymes with Pleiotropic Functions. *Nat. Rev. Mol. Cell Biol.* **2003**, *4* (2), 140–156. https://doi.org/10.1038/nrm1014.

(76) Valleau, D.; Quaile, A. T.; Cui, H.; Xu, X.; Evdokimova, E.; Chang, C.; Cuff, M. E.; Urbanus, M. L.; Houliston, S.; Arrowsmith, C. H.; Ensminger, A. W.; Savchenko, A. Discovery of Ubiquitin Deamidases in the Pathogenic Arsenal of Legionella Pneumophila. *Cell Rep.* **2018**, *23* (2), 568–583. https://doi.org/10.1016/j.celrep.2018.03.060.

(77) Puvar, K.; Zhou, Y.; Qiu, J.; Luo, Z.-Q.; Wirth, M. J.; Das, C. Ubiquitin Chains Modified by the Bacterial Ligase SdeA Are Protected from Deubiquitinase Hydrolysis. *Biochemistry (Mosc.)* **2017**, *56* (36), 4762–4766. https://doi.org/10.1021/acs.biochem.7b00664.

(78) Puvar, K.; Saleh, A. M.; Curtis, R. W.; Zhou, Y.; R. Nyalapatla, P.; Fu, J.; Rovira, A. R.; Tor, Y.; Luo, Z.-Q.; Ghosh, A. K.; Wirth, M. J.; Chmielewski, J.; Kinzer-Ursem, T. L.; Das, C. Fluorescent Probes for Monitoring Serine Ubiquitination. *Biochemistry (Mosc.)* **2020**. https://doi.org/10.1021/acs.biochem.0c00067.

(79) Clague, M. J.; Heride, C.; Urbé, S. The Demographics of the Ubiquitin System. *Trends Cell Biol.* **2015**, *25* (7), 417–426. https://doi.org/10.1016/j.tcb.2015.03.002.

(80) Barrio, J. R.; Secrist, J. A.; Leonard, N. J. A Fluorescent Analog of Nicotinamide Adenine Dinucleotide. *Proc. Natl. Acad. Sci.* **1972**, *69* (8), 2039–2042. https://doi.org/10.1073/pnas.69.8.2039.

(81) Rovira, A. R.; Fin, A.; Tor, Y. Emissive Synthetic Cofactors: An Isomorphic, Isofunctional, and Responsive NAD+ Analogue. *J. Am. Chem. Soc.* **2017**, *139* (44), 15556–15559. https://doi.org/10.1021/jacs.7b05852.

(82) Michel, M. A.; Komander, D.; Elliott, P. R. Enzymatic Assembly of Ubiquitin Chains. *Methods Mol. Biol. Clifton NJ* **2018**, *1844*, 73–84. https://doi.org/10.1007/978-1-4939-8706-1_6.

(83) 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 (Mosc.)* **2014**, *53* (19), 3199–3217. https://doi.org/10.1021/bi5003162.

(84) Keusekotten, K.; Elliott, P. R.; Glockner, L.; Fiil, B. K.; Damgaard, R. B.; Kulathu, Y.; Wauer, T.; Hospenthal, M. K.; Gyrd-Hansen, M.; Krappmann, D.; Hofmann, K.; Komander, D. OTULIN Antagonizes LUBAC Signaling by Specifically Hydrolyzing Met1-Linked Polyubiquitin. *Cell* **2013**, *153* (6), 1312–1326. https://doi.org/10.1016/j.cell.2013.05.014.

(85) Klebl, B. M.; Pette, D. A Fluorometric Assay for Measurement of Mono-ADP-Ribosyltransferase Activity. *Anal. Biochem.* **1996**, *239* (2), 145–152. https://doi.org/10.1006/abio.1996.0309.

(86) Lorenz, S.; Bhattacharyya, M.; Feiler, C.; Rape, M.; Kuriyan, J. Crystal Structure of a Ube2S-Ubiquitin Conjugate. *PLOS ONE* **2016**, *11* (2), e0147550. https://doi.org/10.1371/journal.pone.0147550.

(87) Siepmann, T. J.; Bohnsack, R. N.; Tokgöz, Z.; Baboshina, O. V.; Haas, A. L. Protein Interactions within the N-End Rule Ubiquitin Ligation Pathway. *J. Biol. Chem.* **2003**, *278* (11), 9448–9457. https://doi.org/10.1074/jbc.M211240200.

(88) Stewart, M. D.; Ritterhoff, T.; Klevit, R. E.; Brzovic, P. S. E2 Enzymes: More than Just Middle Men. *Cell Res.* **2016**, *26* (4), 423–440. https://doi.org/10.1038/cr.2016.35.

(89) Eddins, M. J.; Carlile, C. M.; Gomez, K. M.; Pickart, C. M.; Wolberger, C. Mms2-Ubc13 Covalently Bound to Ubiquitin Reveals the Structural Basis of Linkage-Specific Polyubiquitin Chain Formation. *Nat. Struct. Mol. Biol.* **2006**, *13* (10), 915–920. https://doi.org/10.1038/nsmb1148.

(90) Ensminger, A. W.; Isberg, R. R. E3 Ubiquitin Ligase Activity and Targeting of BAT3 by Multiple Legionella Pneumophila Translocated Substrates. *Infect. Immun.* **2010**, *78* (9), 3905–3919. https://doi.org/10.1128/IAI.00344-10.

(91) Richards, A. M.; Von Dwingelo, J. E.; Price, C. T.; Abu Kwaik, Y. Cellular Microbiology and Molecular Ecology of Legionella–amoeba Interaction. *Virulence* **2013**, *4* (4), 307–314. https://doi.org/10.4161/viru.24290.

(92) Park, J. M.; Ghosh, S.; O'Connor, T. J. Combinatorial Selection in Amoebal Hosts Drives the Evolution of the Human Pathogen Legionella Pneumophila. *Nat. Microbiol.* **2020**, *5* (4), 599–609. https://doi.org/10.1038/s41564-019-0663-7.

(93) Gomez-Valero, L.; Rusniok, C.; Carson, D.; Mondino, S.; Pérez-Cobas, A. E.; Rolando, M.; Pasricha, S.; Reuter, S.; Demirtas, J.; Crumbach, J.; Descorps-Declere, S.; Hartland, E. L.; Jarraud, S.; Dougan, G.; Schroeder, G. N.; Frankel, G.; Buchrieser, C. More than 18,000 Effectors in the Legionella Genus Genome Provide Multiple, Independent Combinations for Replication in Human Cells. *Proc. Natl. Acad. Sci.* **2019**, *116* (6), 2265–2273. https://doi.org/10.1073/pnas.1808016116.

(94) Nikolajsen, C. L.; Dyrlund, T. F.; Poulsen, E. T.; Enghild, J. J.; Scavenius, C. Coagulation Factor XIIIa Substrates in Human Plasma IDENTIFICATION AND INCORPORATION INTO THE CLOT. *J. Biol. Chem.* **2014**, *289* (10), 6526–6534. https://doi.org/10.1074/jbc.M113.517904.

(95) [20] Processing of X-Ray Diffraction Data Collected in Oscillation Mode. **1997**, *276*, 307–326. https://doi.org/10.1016/S0076-6879(97)76066-X.

(96) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 Suite and Current Developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67* (4), 235–242. https://doi.org/10.1107/S0907444910045749.

(97) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66* (2), 213–221. https://doi.org/10.1107/S0907444909052925.

(98) Yao, Q.; Cui, J.; Wang, J.; Li, T.; Wan, X.; Luo, T.; Gong, Y.-N.; Xu, Y.; Huang, N.; Shao, F. Structural Mechanism of Ubiquitin and NEDD8 Deamidation Catalyzed by Bacterial Effectors That Induce Macrophage-Specific Apoptosis. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (50), 20395–20400. https://doi.org/10.1073/pnas.1210831109.

(99) Crow, A.; Hughes, R. K.; Taieb, F.; Oswald, E.; Banfield, M. J. The Molecular Basis of Ubiquitin-like Protein NEDD8 Deamidation by the Bacterial Effector Protein Cif. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (27), E1830-1838. https://doi.org/10.1073/pnas.1112107109.

(100) Shin, Y.-C.; Tang, S.-J.; Chen, J.-H.; Liao, P.-H.; Chang, S.-C. The Molecular Determinants of NEDD8 Specific Recognition by Human SENP8. *PLOS ONE* **2011**, *6* (11), e27742. https://doi.org/10.1371/journal.pone.0027742.

(101) Chica, R. A.; Gagnon, P.; Keillor, J. W.; Pelletier, J. N. Tissue Transglutaminase Acylation: Proposed Role of Conserved Active Site Tyr and Trp Residues Revealed by Molecular Modeling of Peptide Substrate Binding. *Protein Sci. Publ. Protein Soc.* **2004**, *13* (4), 979–991. https://doi.org/10.1110/ps.03433304.

(102) Murthy, S. N. P.; Iismaa, S.; Begg, G.; Freymann, D. M.; Graham, R. M.; Lorand, L. Conserved Tryptophan in the Core Domain of Transglutaminase Is Essential for Catalytic Activity. *Proc. Natl. Acad. Sci.* **2002**, *99* (5), 2738–2742. https://doi.org/10.1073/pnas.052715799.

(103) Iismaa, S. E.; Holman, S.; Wouters, M. A.; Lorand, L.; Graham, R. M.; Husain, A. Evolutionary Specialization of a Tryptophan Indole Group for Transition-State Stabilization by Eukaryotic Transglutaminases. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (22), 12636–12641. https://doi.org/10.1073/pnas.1635052100.

(104) Kubori, T.; Nagai, H. Bacterial Effector-Involved Temporal and Spatial Regulation by Hijack of the Host Ubiquitin Pathway. *Front. Microbiol.* **2011**, *2*. https://doi.org/10.3389/fmicb.2011.00145.

(105) Hubber, A.; Arasaki, K.; Nakatsu, F.; Hardiman, C.; Lambright, D.; De Camilli, P.; Nagai, H.; Roy, C. R. The Machinery at Endoplasmic Reticulum-Plasma Membrane Contact Sites Contributes to Spatial Regulation of Multiple Legionella Effector Proteins. *PLoS Pathog.* **2014**, *10* (7). https://doi.org/10.1371/journal.ppat.1004222.

(106) Al-Khodor, S.; Al-Quadan, T.; Abu Kwaik, Y. Temporal and Differential Regulation of Expression of the Eukaryotic-like Ankyrin Effectors of Legionella Pneumophila. *Environ. Microbiol. Rep.* **2010**, *2* (5), 677–684. https://doi.org/10.1111/j.1758-2229.2010.00159.x.

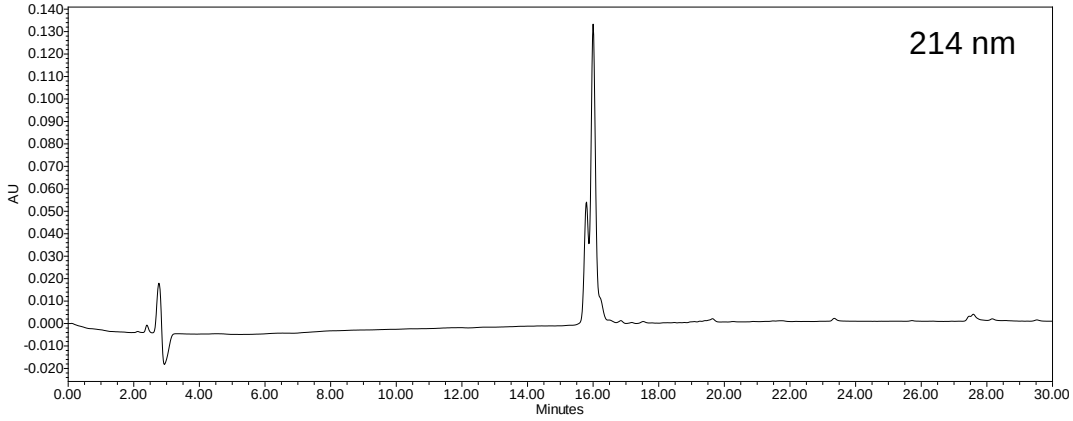
(107) Urbanus, M. L.; Quaile, A. T.; Stogios, P. J.; Morar, M.; Rao, C.; Di Leo, R.; Evdokimova, E.; Lam, M.; Oatway, C.; Cuff, M. E.; Osipiuk, J.; Michalska, K.; Nocek, B. P.; Taipale, M.; Savchenko, A.; Ensminger, A. W. Diverse Mechanisms of Metaeffector Activity in an Intracellular Bacterial Pathogen, Legionella Pneumophila. *Mol. Syst. Biol.* **2016**, *12* (12), 893. https://doi.org/10.15252/msb.20167381.

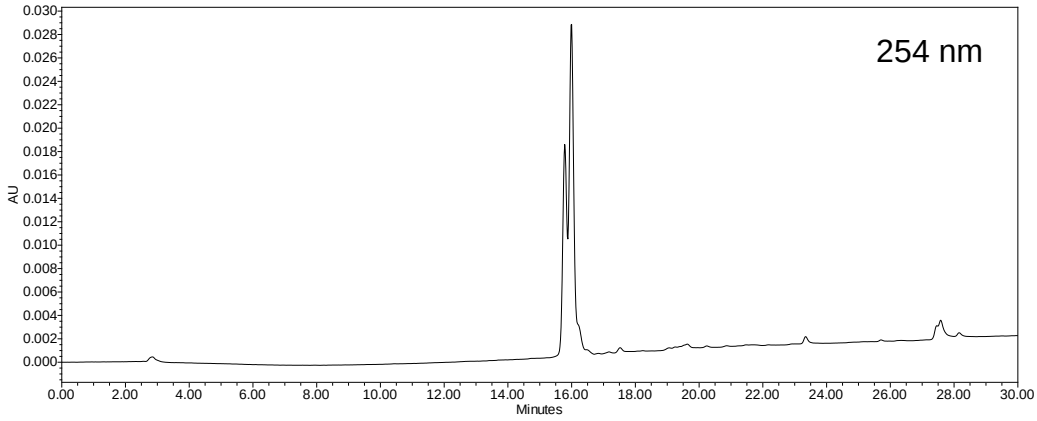
(108) Tan, Y.; Arnold, R. J.; Luo, Z.-Q. Legionella Pneumophila Regulates the Small GTPase Rab1 Activity by Reversible Phosphorylcholination. *Proc. Natl. Acad. Sci.* **2011**, *108* (52), 21212–21217. https://doi.org/10.1073/pnas.1114023109.
(109) Neunuebel, M. R.; Chen, Y.; Gaspar, A. H.; Backlund, P. S.; Yergey, A.; Machner, M. P. De-AMPylation of the Small GTPase Rab1 by the Pathogen Legionella Pneumophila. *Science* **2011**, *333* (6041), 453–456. https://doi.org/10.1126/science.1207193.
(110) Chen, Y.; Tascón, I.; Neunuebel, M. R.; Pallara, C.; Brady, J.; Kinch, L. N.; Fernández-Recio, J.; Rojas, A. L.; Machner, M. P.; Hierro, A. Structural Basis for Rab1 De-AMPylation by the Legionella Pneumophila Effector SidD. *PLOS Pathog.* **2013**, *9* (5), e1003382. https://doi.org/10.1371/journal.ppat.1003382.
(111) Joseph, A. M.; Pohl, A. E.; Ball, T. J.; Abram, T. G.; Johnson, D. K.; Geisbrecht, B. V.; Shames, S. R. The Legionella Pneumophila Metaeffector Lpg2505 (SusF) Regulates SidI-Mediated Translation Inhibition and GDP-Dependent Glycosyltransferase Activity. *bioRxiv* **2019**, 845313. https://doi.org/10.1101/845313.
(112) Qiu, J.; Yu, K.; Fei, X.; Liu, Y.; Nakayasu, E. S.; Piehowski, P. D.; Shaw, J. B.; Puvar, K.; Das, C.; Liu, X.; Luo, Z.-Q. A Unique Deubiquitinase That Deconjugates Phosphoribosyl-Linked Protein Ubiquitination. *Cell Res.* **2017**, *27* (7), 865–881. https://doi.org/10.1038/cr.2017.66.
(113) Black, M. H.; Osinski, A.; Gradowski, M.; Servage, K. A.; Pawłowski, K.; Tomchick, D. R.; Tagliabracci, V. S. Bacterial Pseudokinase Catalyzes Protein Polyglutamylation to Inhibit the SidE-Family Ubiquitin Ligases. *Science* **2019**, *364* (6442), 787–792. https://doi.org/10.1126/science.aaw7446.
(114) Gan, N.; Zhen, X.; Liu, Y.; Xu, X.; He, C.; Qiu, J.; Liu, Y.; Fujimoto, G. M.; Nakayasu, E. S.; Zhou, B.; Zhao, L.; Puvar, K.; Das, C.; Ouyang, S.; Luo, Z.-Q. Regulation of Phosphoribosyl Ubiquitination by a Calmodulin-Dependent Glutamylase. *Nature* **2019**, *572* (7769), 387–391. https://doi.org/10.1038/s41586-019-1439-1.
(115) Bhogaraju, S.; Bonn, F.; Mukherjee, R.; Adams, M.; Pfleiderer, M. M.; Galej, W. P.; Matkovic, V.; Lopez-Mosqueda, J.; Kalayil, S.; Shin, D.; Dikic, I. Inhibition of Bacterial Ubiquitin Ligases by SidJ–calmodulin Catalysed Glutamylation. *Nature* **2019**, *572* (7769), 382–386. https://doi.org/10.1038/s41586-019-1440-8.
(116) Sulpizio, A.; Minelli, M. E.; Wan, M.; Burrowes, P. D.; Wu, X.; Sanford, E. J.; Shin, J.-H.; Williams, B. C.; Goldberg, M. L.; Smolka, M. B.; Mao, Y. Protein Polyglutamylation Catalyzed by the Bacterial Calmodulin-Dependent Pseudokinase SidJ. *eLife* **2019**, *8*, e51162. https://doi.org/10.7554/eLife.51162.
(117) Boucher, D.; Larcher, J.-C.; Gros, F.; Denoulet, P. Polyglutamylation of Tubulin as a Progressive Regulator of in Vitro Interactions between the Microtubule-Associated Protein Tau and Tubulin. *Biochemistry (Mosc.)* **1994**, *33* (41), 12471–12477. https://doi.org/10.1021/bi00207a014.
(118) Tidow, H.; Nissen, P. Structural Diversity of Calmodulin Binding to Its Target Sites. *FEBS J.* **2013**, *280* (21), 5551–5565. https://doi.org/10.1111/febs.12296.
(119) Bhatnagar, R.; Batra, S. Anthrax Toxin. *Crit. Rev. Microbiol.* **2001**, *27* (3), 167–200. https://doi.org/10.1080/20014091096738.
(120) Friebe, S.; van der Goot, F. G.; Bürgi, J. The Ins and Outs of Anthrax Toxin. *Toxins* **2016**, *8* (3). https://doi.org/10.3390/toxins8030069.

(121) Satchell, K. J. F. Structure and Function of MARTX Toxins and Other Large Repetitive RTX Proteins. *Annu. Rev. Microbiol.* **2011**, *65* (1), 71–90. https://doi.org/10.1146/annurev-micro-090110-102943.
(122) Satchell, K. J. F. Multifunctional-Autoprocessing Repeats-in-Toxin (MARTX) Toxins of Vibrios. *Microbiol. Spectr.* **2015**, *3* (3). https://doi.org/10.1128/microbiolspec.VE-0002-2014.
(123) Shin, D.; Mukherjee, R.; Liu, Y.; Gonzalez, A.; Bonn, F.; Liu, Y.; Rogov, V. V.; Heinz, M.; Stolz, A.; Hummer, G.; Dötsch, V.; Luo, Z.-Q.; Bhogaraju, S.; Dikic, I. Regulation of Phosphoribosyl-Linked Serine Ubiquitination by Deubiquitinases DupA and DupB. *Mol. Cell* **2019**, *0* (0). https://doi.org/10.1016/j.molcel.2019.10.019.
(124) Wan, M.; Sulpizio, A. G.; Akturk, A.; Beck, W. H. J.; Lanz, M.; Faça, V. M.; Smolka, M. B.; Vogel, J. P.; Mao, Y. Deubiquitination of Phosphoribosyl-Ubiquitin Conjugates by Phosphodiesterase-Domain–containing Legionella Effectors. *Proc. Natl. Acad. Sci.* **2019**. https://doi.org/10.1073/pnas.1916287116.
(125) Gan, N.; Guan, H.; Huang, Y.; Yu, T.; Fu, J.; Nakayasu, E. S.; Puvar, K.; Das, C.; Wang, D.; Ouyang, S.; Luo, Z.-Q. Legionella Pneumophila Regulates the Activity of UBE2N by Deamidase-Mediated Deubiquitination. *EMBO J.* **2020**, *39* (4), e102806. https://doi.org/10.15252/embj.2019102806.

APPENDIX A: HPLC TRACES AND MASS SPECTRA OF SYNTHETIC PEPTIDE

Figure 1a
Peptide Information- All purified on 15-90% ACN:Water, 30 min. 0.1% TFA, C18 column
MSS:





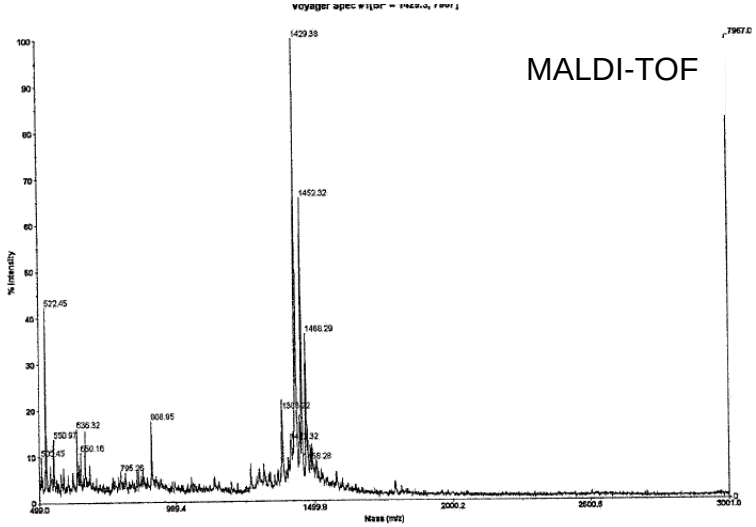
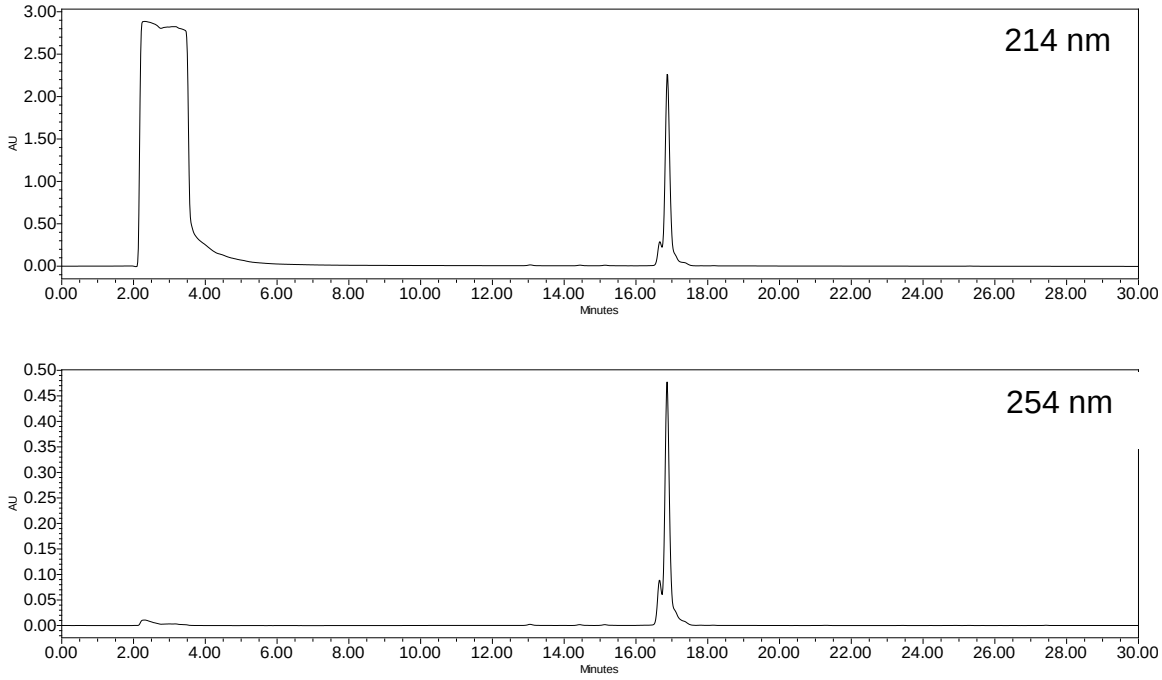


Figure 1b
MSA:



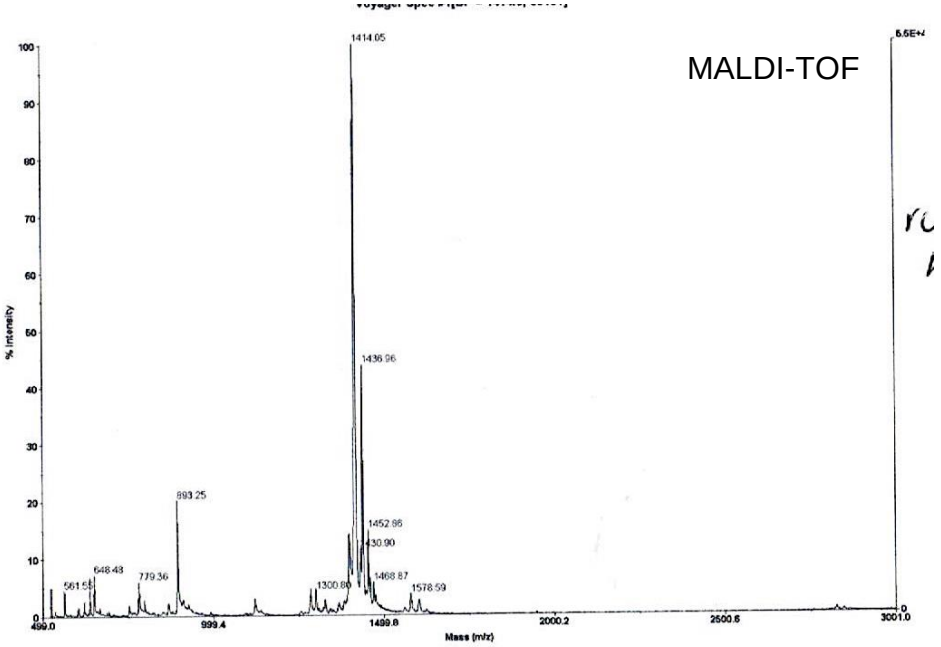
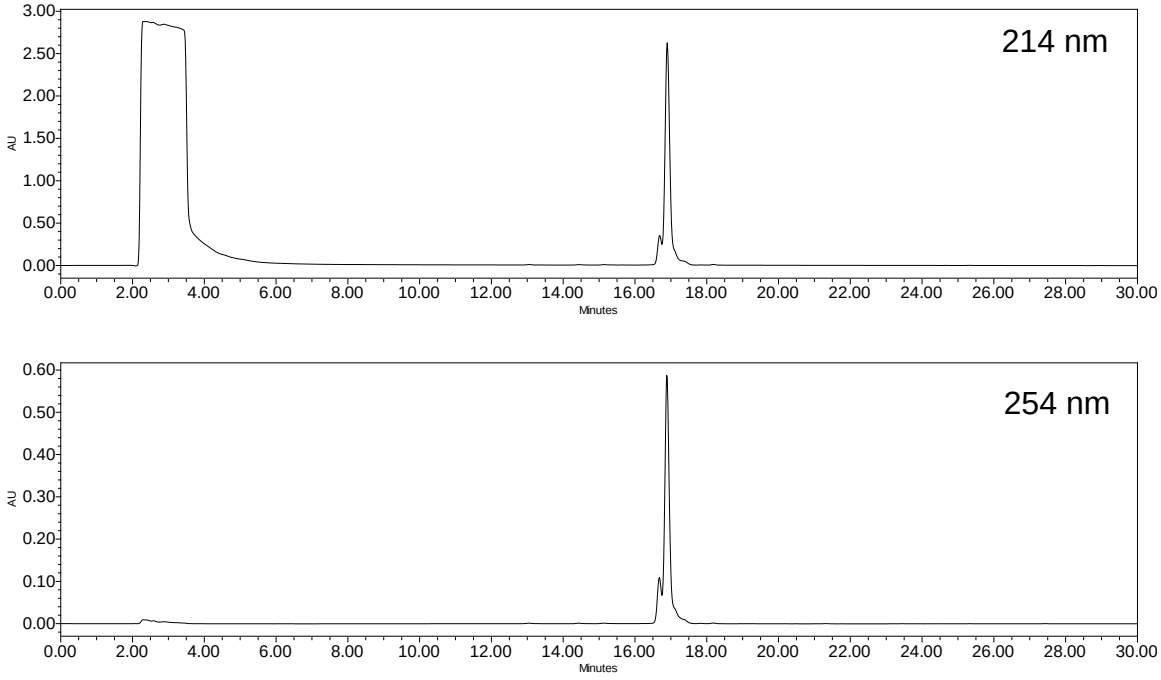


Figure 1c
MAS:



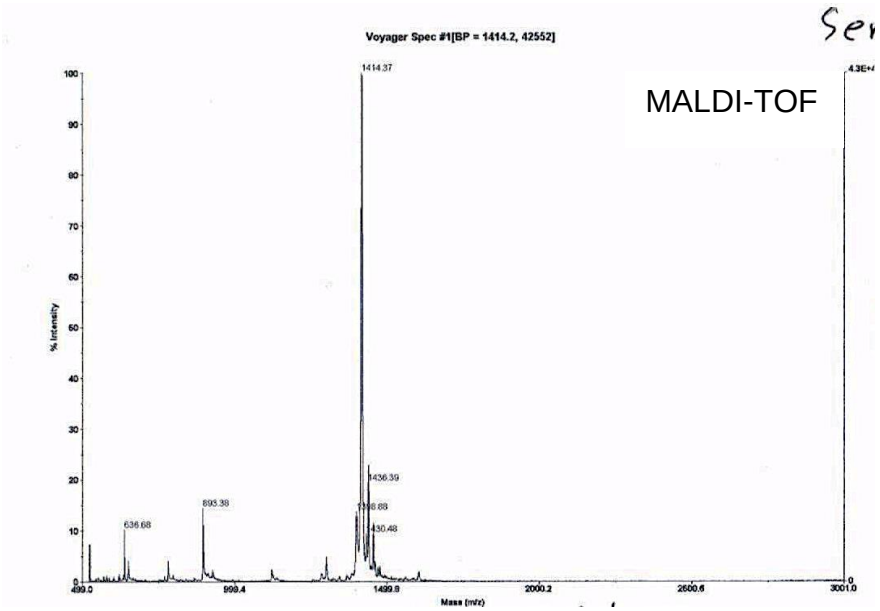
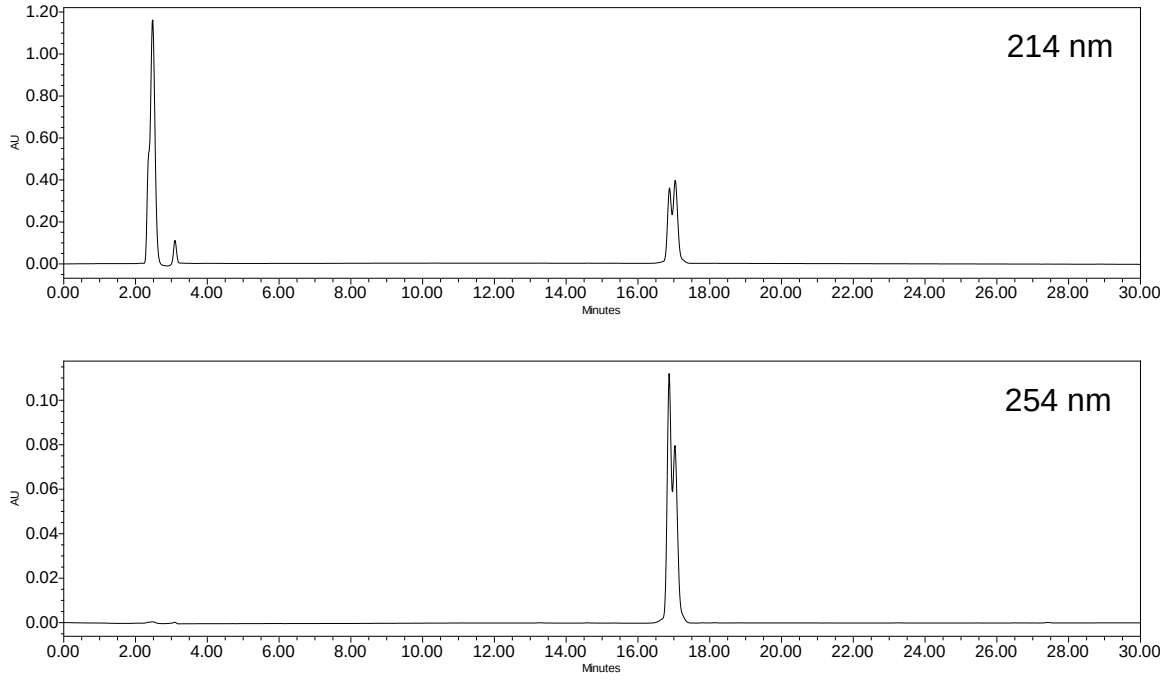


Figure 1d
MAT:



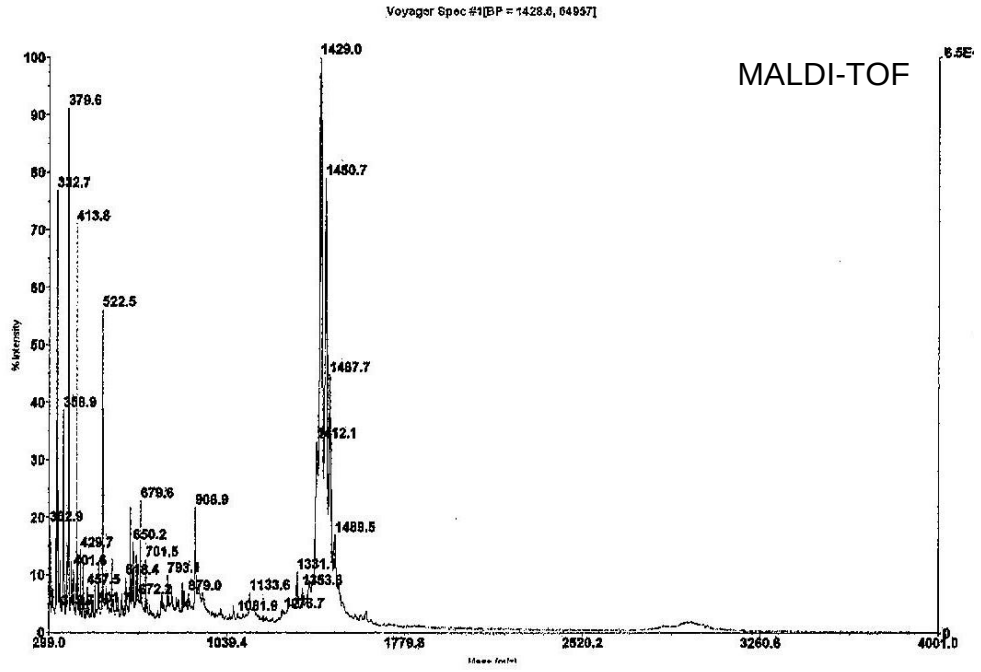
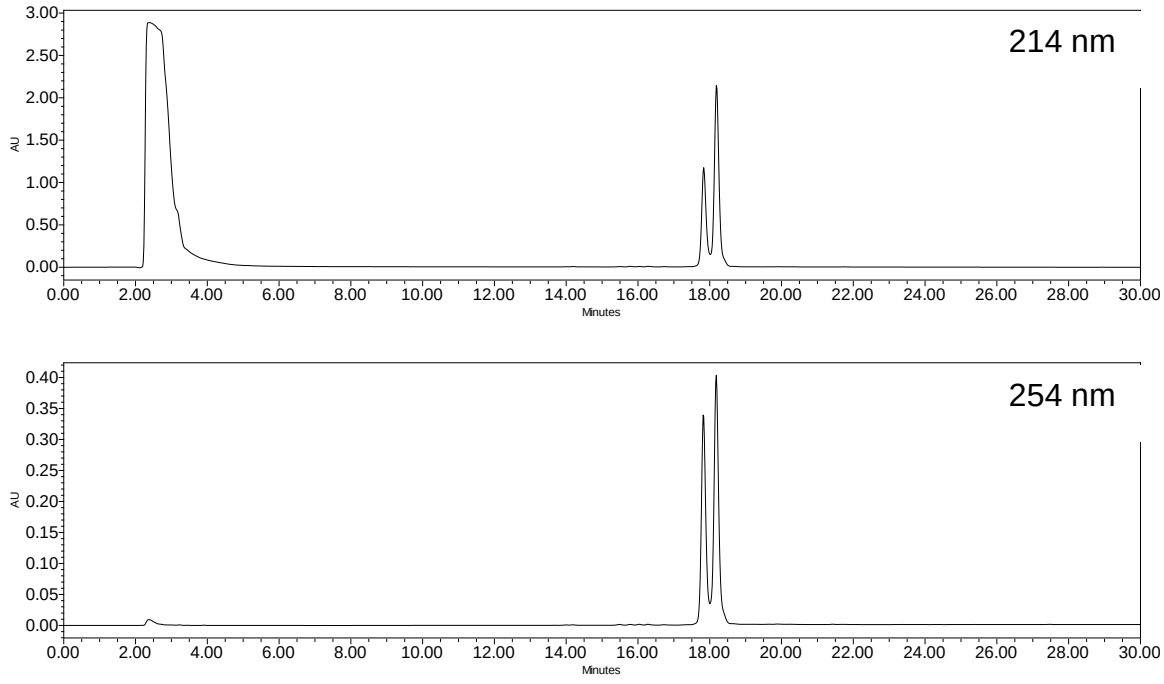


Figure 1e
MAY:



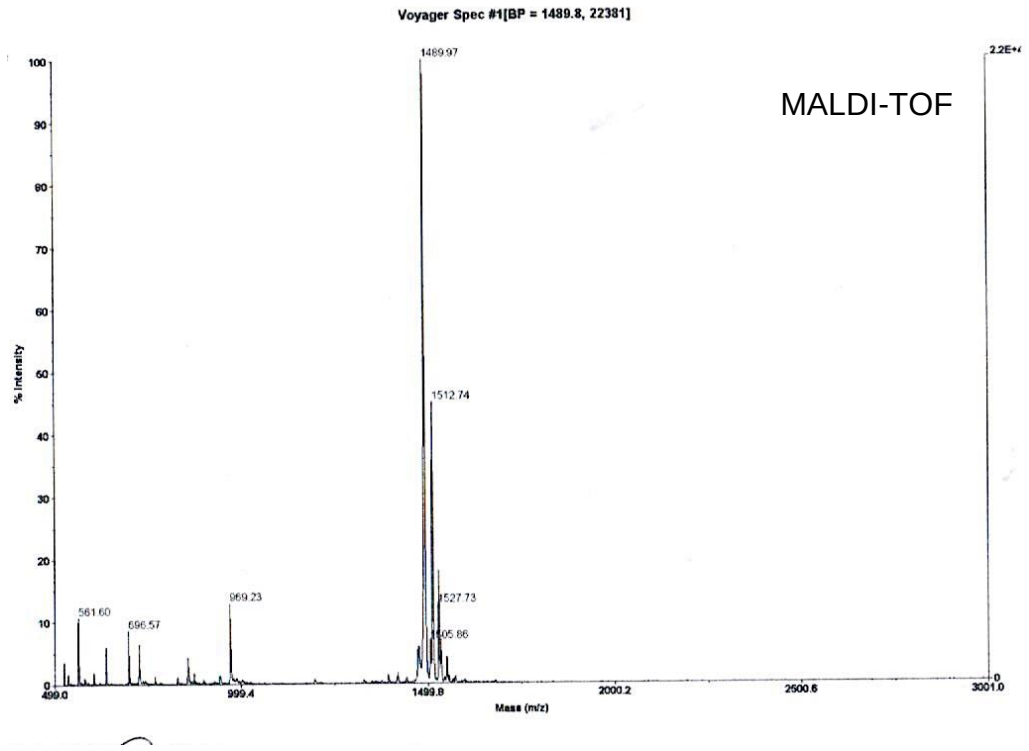


Figure 1. Purification and characterization of synthetic peptides used in this study. (a) MSS, (b) MSA, (c) MSS, (d) MAT, (e) MAY. HPLC traces and mass spectra are shown. Doublet from HPLC traces comes from the two isomers of fluorescein utilized for synthesis. Figures courtesy of Ryan Curtis (Chmielewski Lab, Purdue University).

APPENDIX B: MASS SPECTRA OF DIUBIQUITIN MODIFICATIONS

Figure 1A

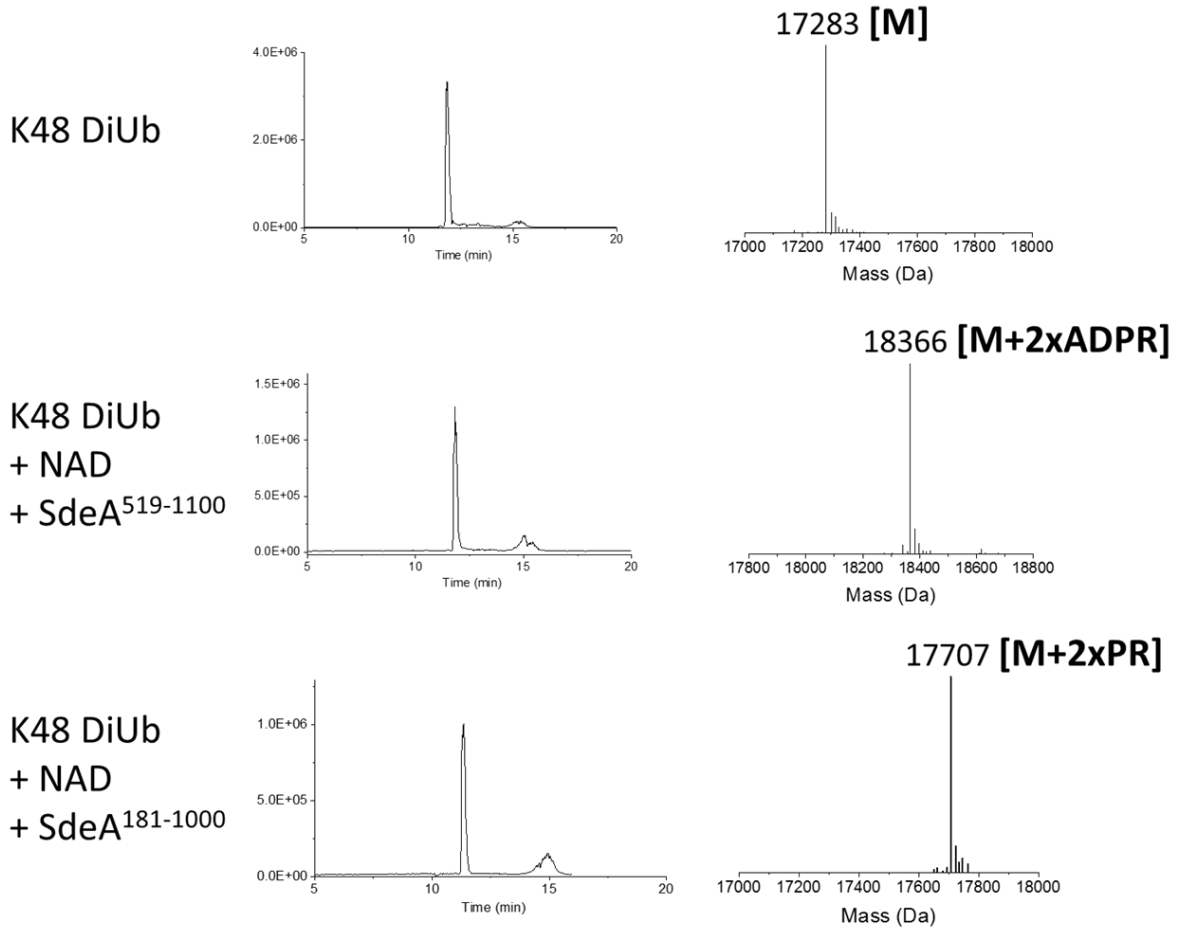


Figure 1B

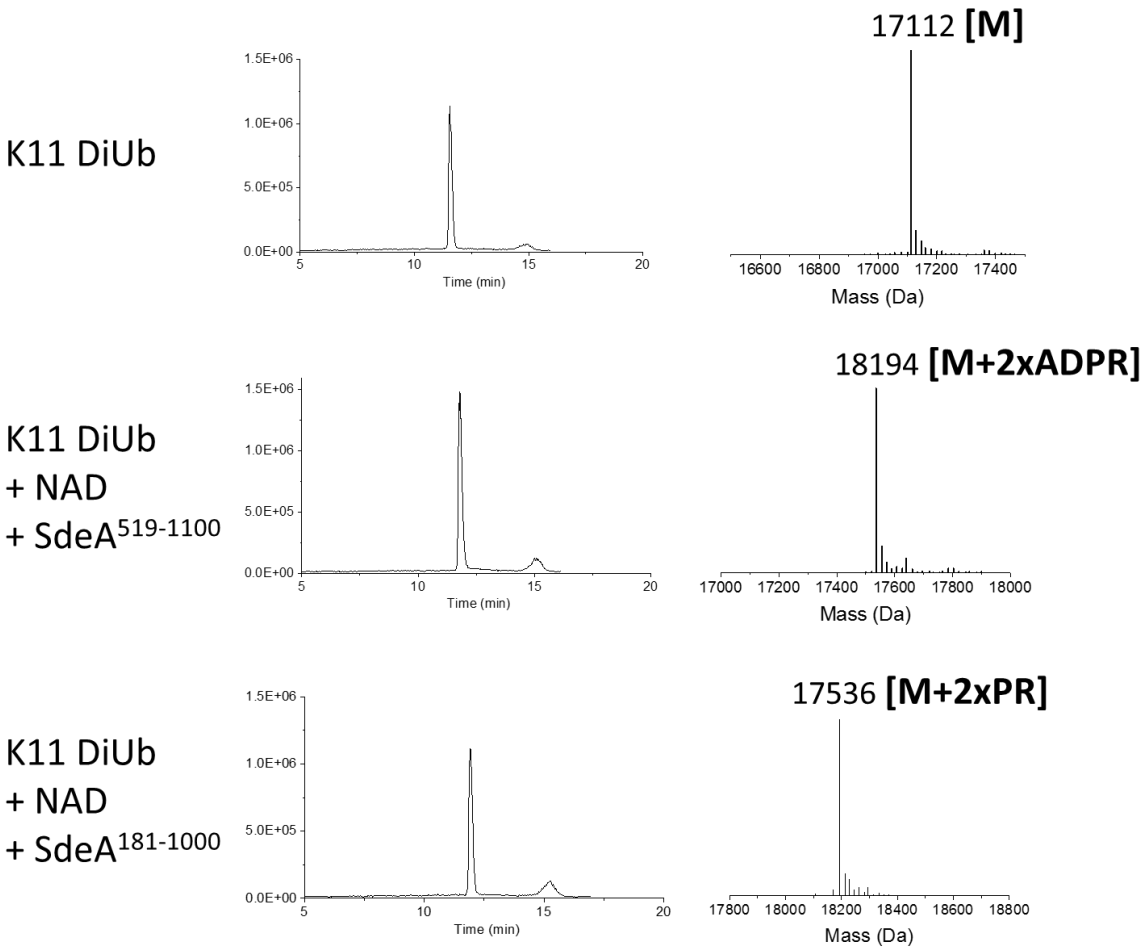

17112 [M]
K11 DiUb
1.5E+06
1.0E+06
5.0E+05
0.0E+00
5
10
15
20
Time (min)
16600
16800
17000
17200
17400
Mass (Da)
18194 [M+2xADPR]
K11 DiUb
+ NAD
+ SdeA519-1100
17000
17200
17400
17600
17800
18000
17536 [M+2xPR]
K11 DiUb
+ NAD
+ SdeA181-1000
17800
18000
18200
18400
18600
18800

Figure 1C

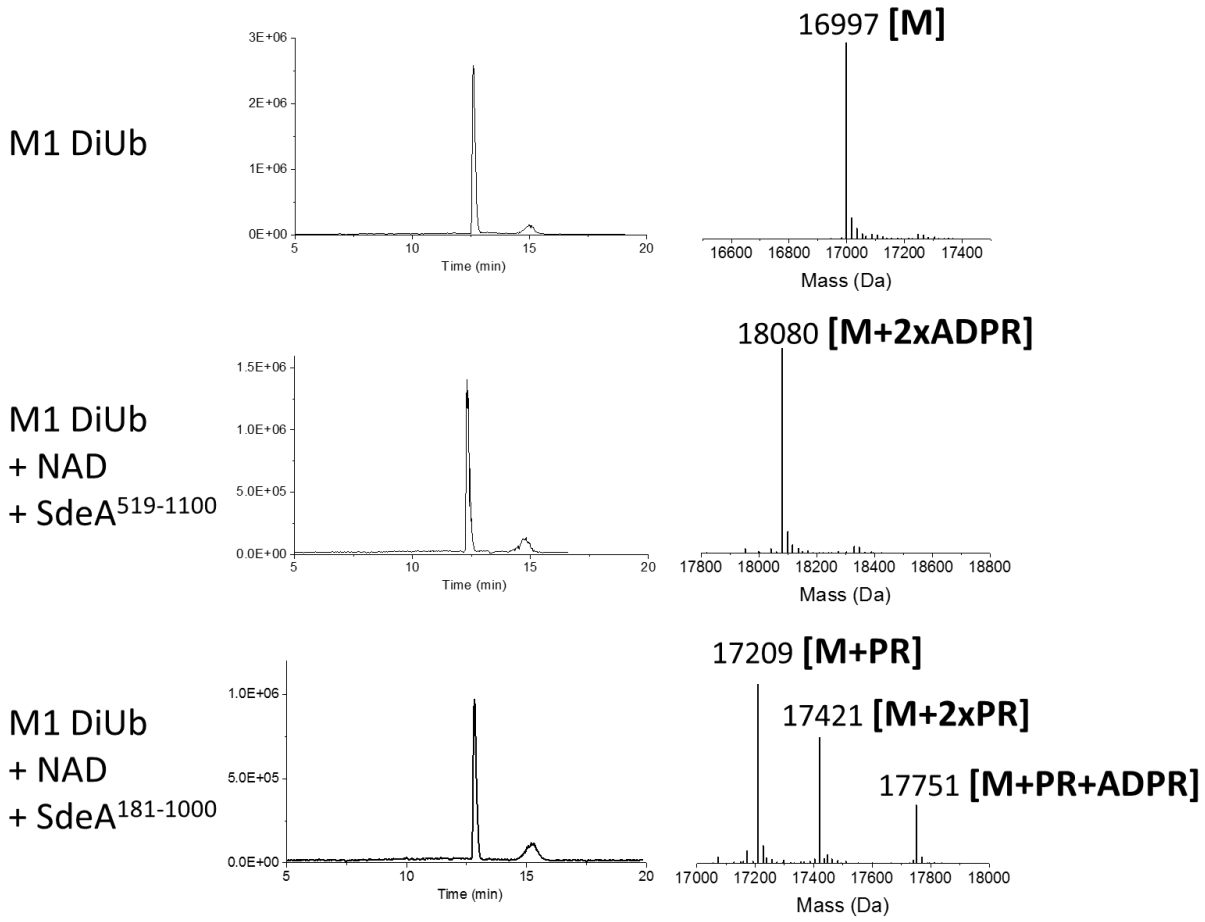

M1 DiUb
16997 [M]
Time (min)
Mass (Da)
M1 DiUb
+ NAD
+ SdeA519-1100
18080 [M+2xADPR]
Time (min)
Mass (Da)
M1 DiUb
+ NAD
+ SdeA181-1000
17209 [M+PR]
17421 [M+2xPR]
17751 [M+PR+ADPR]
Time (min)
Mass (Da)

Figure 1D

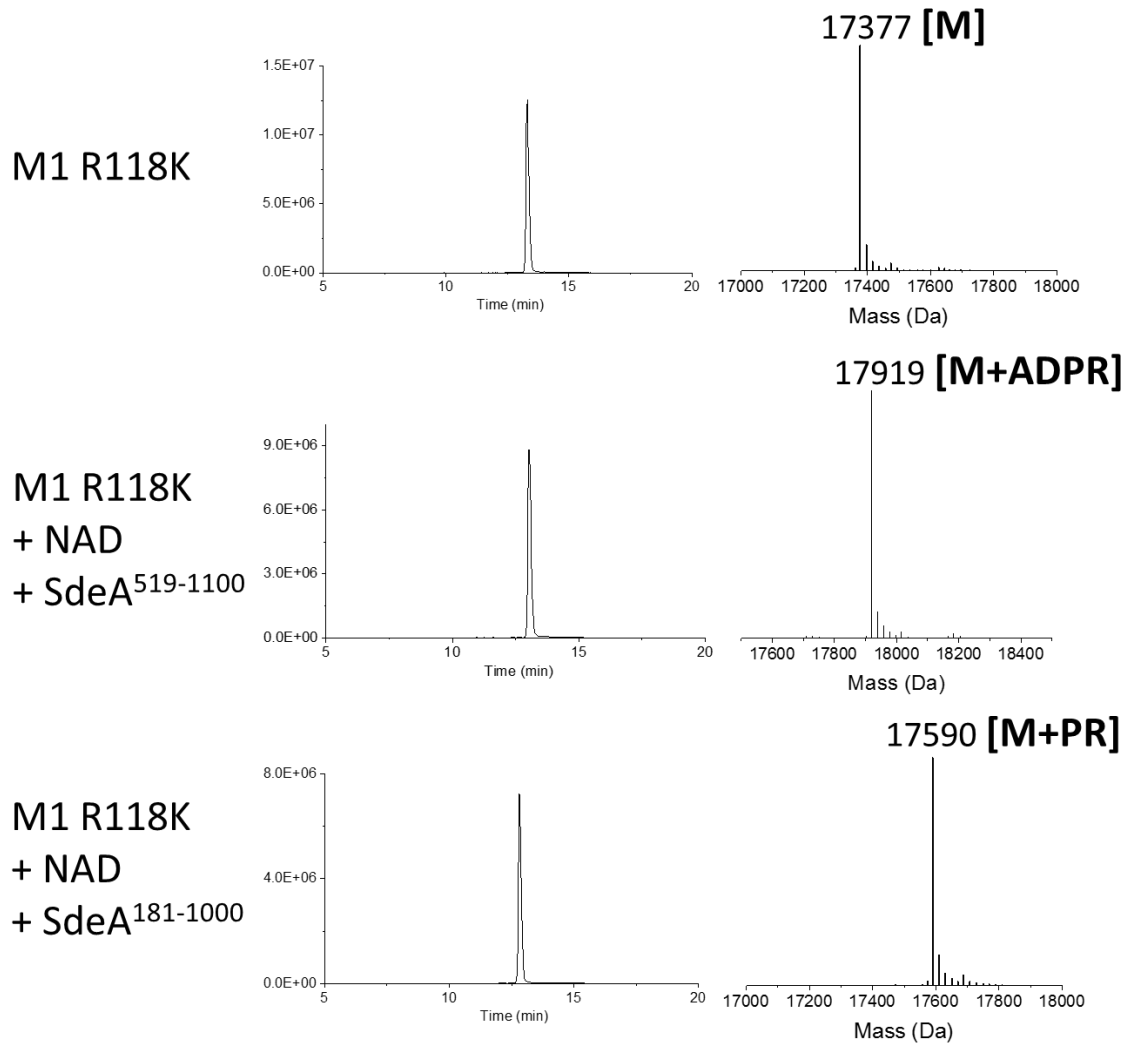

M1 R118K
17377 [M]
1.5E+07
1.0E+07
5.0E+06
0.0E+00
5
10
15
20
Time (min)
17000
17200
17400
17600
17800
18000
Mass (Da)
M1 R118K
+ NAD
+ SdeA519-1100
17919 [M+ADPR]
9.0E+06
6.0E+06
3.0E+06
0.0E+00
Time (min)
17600
17800
18000
18200
18400
Mass (Da)
M1 R118K
+ NAD
+ SdeA181-1000
17590 [M+PR]
8.0E+06
4.0E+06
0.0E+00
Time (min)
Mass (Da)

Figure 1E

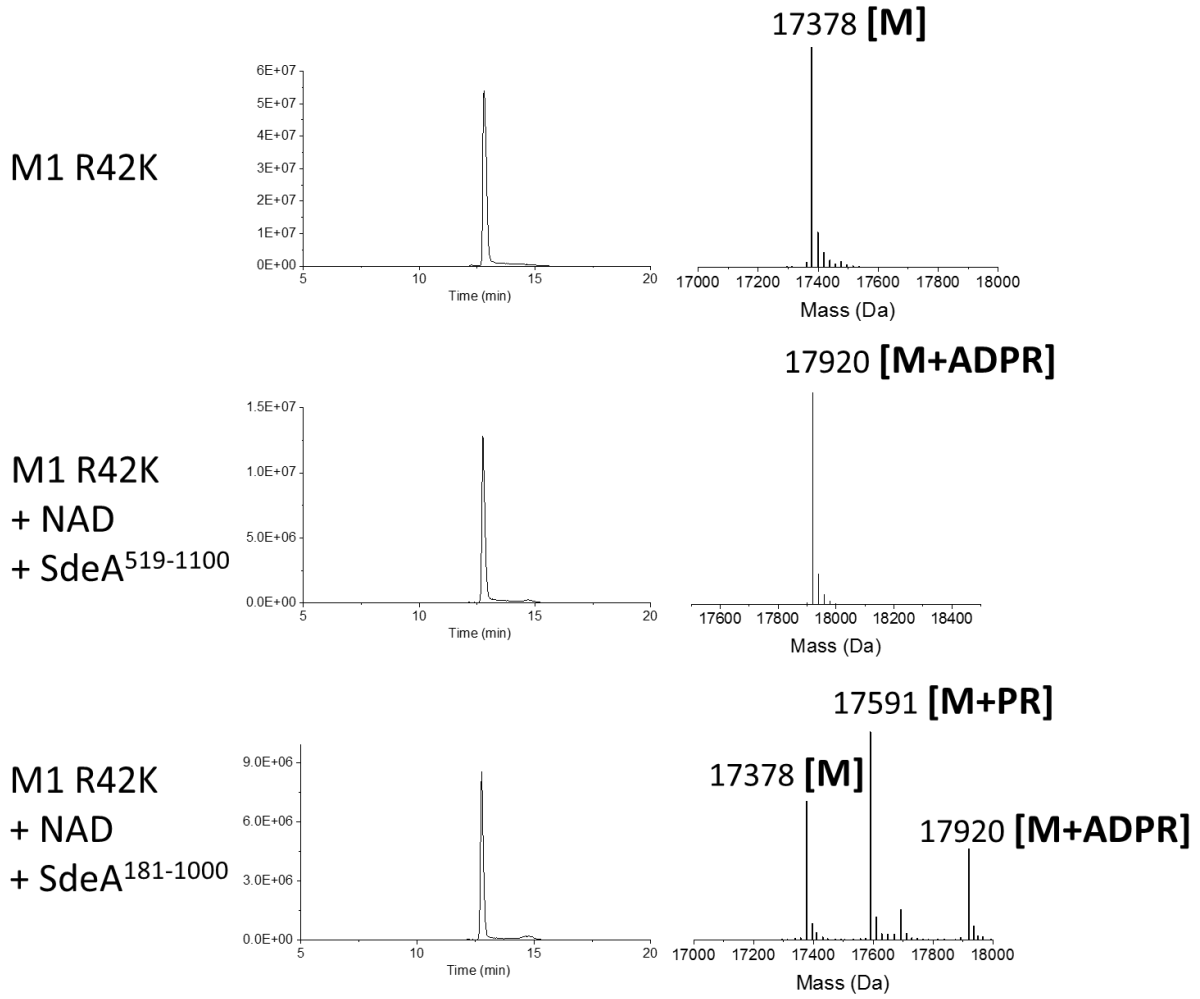
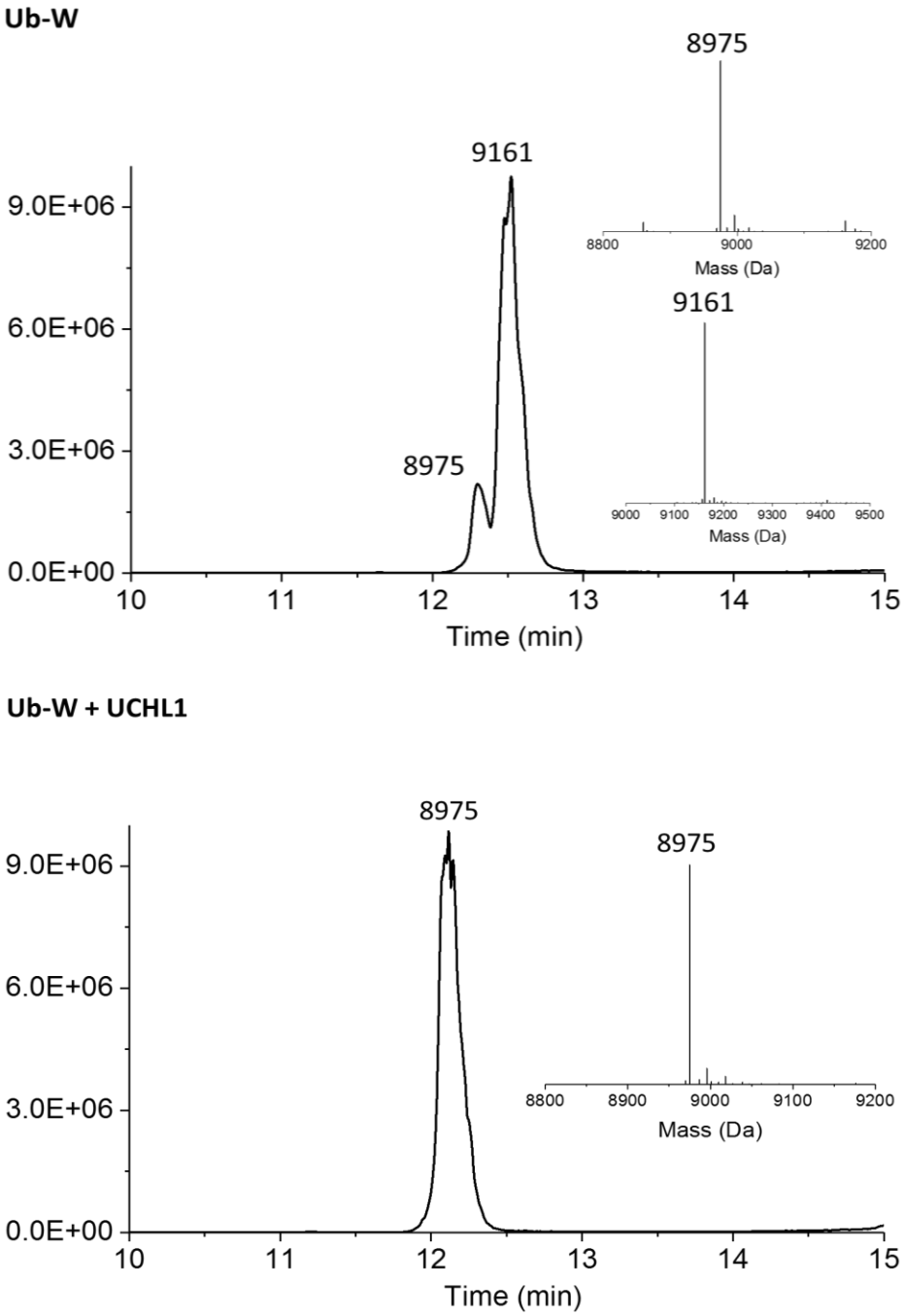
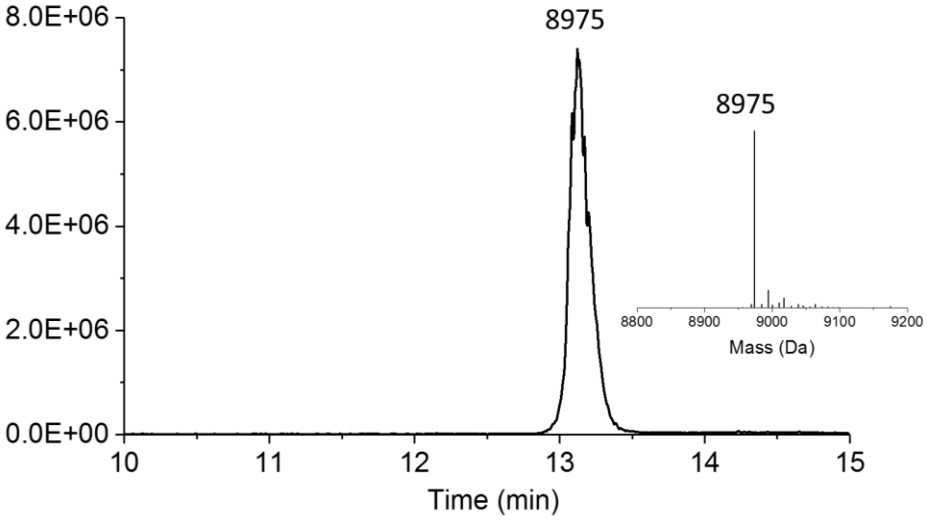


Figure 1. LC-MS of diubiquitin in the presence of SdeA constructs, showing a doubly ADP-ribosylated or phosphoribosylated species, or singly modified when target arginine residues are mutated. Linkage types tested include (A) K48, (B) K11, (C) M1, (D) M1 with R118K mutation, (E) M1 with R42K mutation.

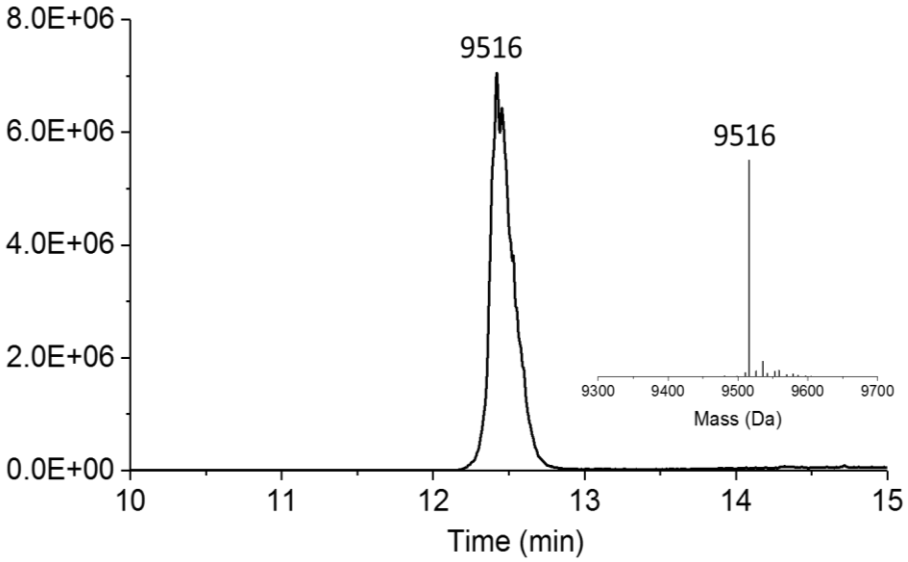
Figure 2


Ub-W
8975
9161
9.0E+06
6.0E+06
3.0E+06
0.0E+00
8975
8800
9000
9200
Mass (Da)
9161
9000
9100
9200
9300
9400
9500
Mass (Da)
10
11
12
13
14
15
Time (min)
Ub-W + UCHL1
8975
8975
9.0E+06
6.0E+06
3.0E+06
0.0E+00
8800
8900
9000
9100
9200
Mass (Da)
10
11
12
13
14
15
Time (min)

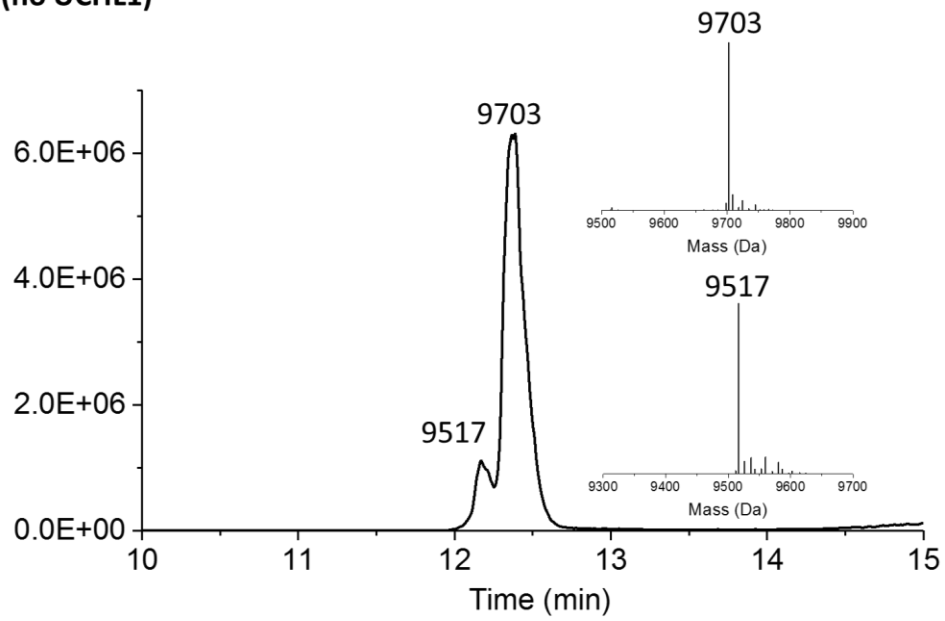
Ub-W + SdeA$^{519\text{-}1100}$ + UCHL1 (no NAD^{+})



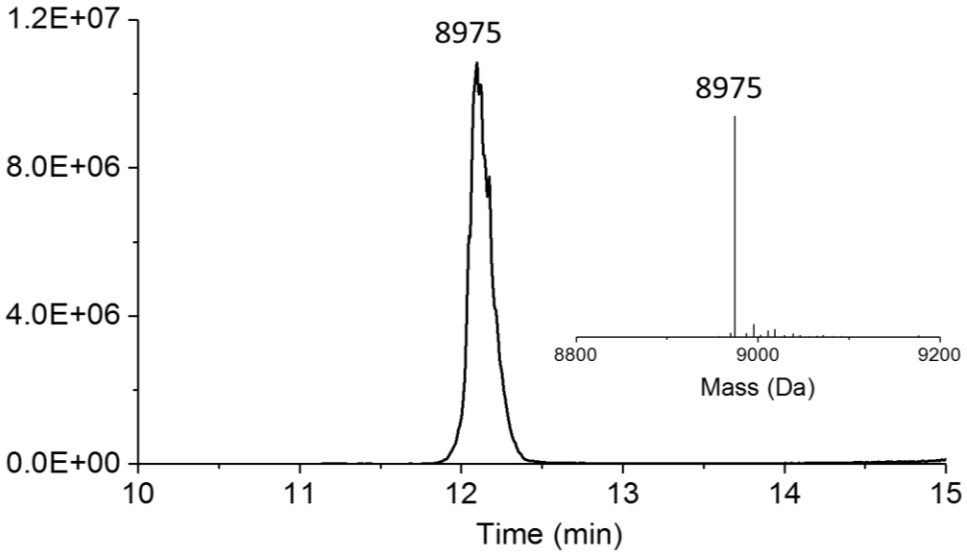
Ub-W + SdeA$^{519\text{-}1100}$ + NAD^{+} + UCHL1



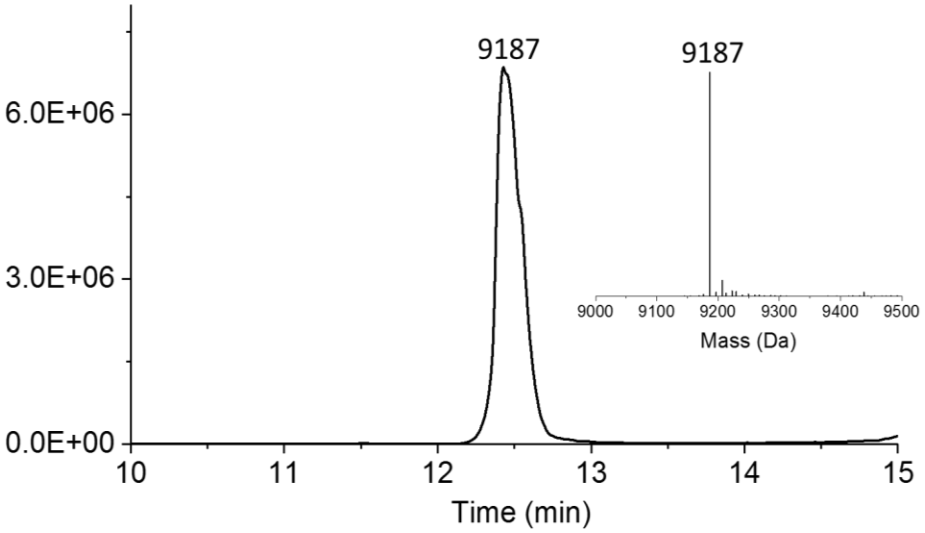
Ub-W + SdeA$^{519-1100}$ + NAD^{+}
(no UCHL1)



Ub-W + SdeA$^{181-1000}$
+ UCHL1
(no NAD^{+})



Ub-W + SdeA$^{181\text{-}1000}$ + NAD^{+} + UCHL1



Ub-W + SdeA$^{519\text{-}1100}$ + NAD^{+} (no UCHL1)

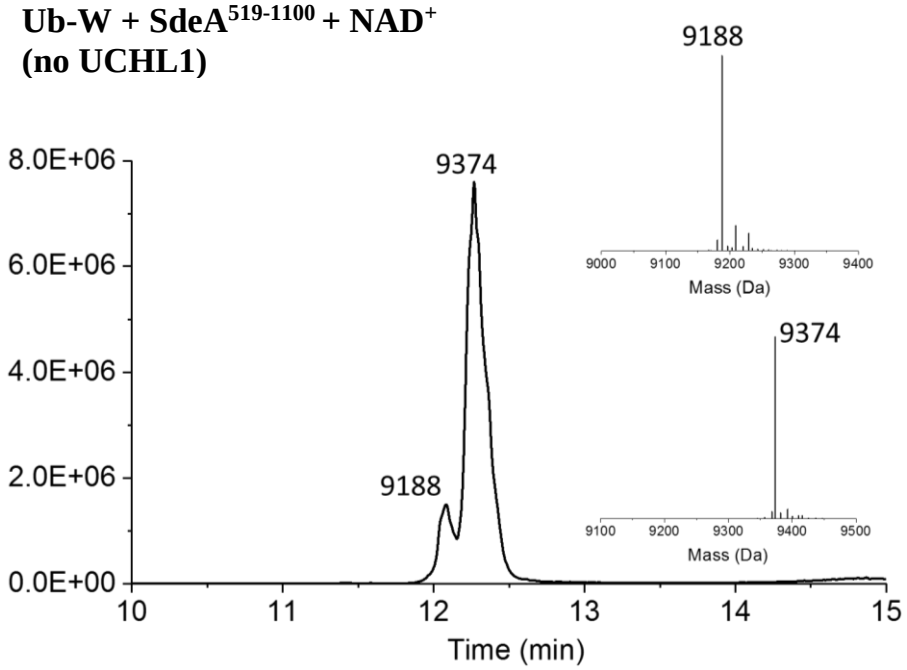


Figure 2. LC-MS of cleavage assay of Ub-W in the presence of UCHL1, with and without SdeA and NAD^{+} pre-incubation

Table A2. Analysis of Ub-W cleavage assay LC-MS data

	Obs. Mass	Theoretical Mass	Species	Δ Mass
Ub-W ctrl	9161	9162.52	GPLGS-Ub-W	-1.52
Ub-W + UCHL1	8975	8976.3	GPLGS-Ub	-1.3
Ub-W + SdeA$^{519-1100}$	9703	9703	GPLGS-Ub-W-ADPR	0
+ NAD^+ no UCHL1				
Ub-W + SdeA$^{519-1100}$	8975	8976.3	GPLGS-Ub	-1.3
+ UCHL1, no NAD^+				
Ub-W + SdeA$^{519-1100}$	9516	9517.3	GPLGS-Ub-ADPR	-1.3
+ NAD^+ + UCHL1				
Ub-W + SdeA$^{181-1100}$	9374	9374.52	GPLGS-Ub-W-PR	-0.52
+ NAD^+ no UCHL1				
Ub-W + SdeA$^{181-1100}$	8975	8976.3	GPLGS-Ub	-1.3
+ UCHL1, no NAD^+				
Ub-W + SdeA$^{181-1100}$	9187	9188.3	GPLGS-Ub-PR	-1.3
+ NAD^+ + UCHL1				

*** GST affinity purification of Ub-W resulted in the residues Gly, Pro, Leu, Gly, Ser (GPLGS) added to its N-terminus. This is a remnant of cloning, with the residues occurring after the cleavage site of PreScission™ Protease but before the start of the protein.

VITA

Kedar Puvar was born in Berwyn, Illinois to Dhirendra and Jayshree Puvar. He grew up in the Chicago area, attending Willow Creek Elementary, Lakeview Junior High and Downers Grove South High School. He developed an interest in Chemistry and Biology by Grade 9, hoping to one day become a researcher.

He pursued a Bachelor of Science in Biochemistry with a Minor in Studio Art at Valparaiso University, where he joined Phi Sigma Kappa Fraternity and conducted undergraduate research with Dr. Laura Rowe, designing a new fluorescent biosensor. He also spent two summers at the Ann & Robert H. Lurie Children's Hospital of Chicago where he helped test and develop new methods for the identification of MRSA strains. During those busy summers, after returning from the lab every evening he would also work as a tennis coach.

Upon graduating from Valpo, he joined the Department of Chemistry at Purdue University for his doctoral studies, joining the lab of Dr. Chitta Das where he worked on the structure and function of bacterial effectors that modify host proteins with ubiquitin. After graduating from Purdue, he will move to the Boston area to start a position as a Research Fellow at the Dana-Farber Cancer Institute in the lab of Dr. Eric Fischer.

PUBLICATIONS

Puvar, K.; Zhou, Y.; Qiu, J.; Luo, Z.-Q.; Wirth, M. J.; Das, C. Ubiquitin Chains Modified by the Bacterial Ligase SdeA Are Protected from Deubiquitinase Hydrolysis. *Biochemistry* **2017**, *56* (36), 4762–4766. https://doi.org/10.1021/acs.biochem.7b00664.

Puvar, K.; Luo, Z.-Q.; Das, C. Uncovering the Structural Basis of a New Twist in Protein Ubiquitination. *Trends in Biochemical Sciences* **2019**, *44* (5), 467–477. https://doi.org/10.1016/j.tibs.2018.11.006.

Gan, N.; Guan, H.; Huang, Y.; Yu, T.; Fu, J.; Nakayasu, E. S.; Puvar, K.; Das, C.; Wang, D.; Ouyang, S.; Luo, Z.-Q. Legionella Pneumophila Regulates the Activity of UBE2N by Deamidase-Mediated Deubiquitination. *The EMBO Journal* **2020**, *39* (4), e102806. https://doi.org/10.15252/embj.2019102806.

Puvar, K.; Saleh, A. M.; Curtis, R. W.; Zhou, Y.; R. Nyalapatla, P.; Fu, J.; Rovira, A. R.; Tor, Y.; Luo, Z.-Q.; Ghosh, A. K.; Wirth, M. J.; Chmielewski, J.; Kinzer-Ursem, T. L.; Das, C. Fluorescent Probes for Monitoring Serine Ubiquitination. *Biochemistry* **2020**. https://doi.org/10.1021/acs.biochem.0c00067.

Trends in Biochemical Sciences

CellPress REVIEWS

Review

Uncovering the Structural Basis of a New Twist in Protein Ubiquitination

Kedar Puvar,[1] Zhao-Qing Luo,[2,*] and Chittaranjan Das[1,*]

Members of the SidE effector family from *Legionella pneumophila* represent a new paradigm in the ubiquitin world. These enzymes catalyze ubiquitination of target proteins via a mechanism different from that of conventional E1-E2-E3 biochemistry and play important roles in *L. pneumophila* virulence. They combine mono-ADP-ribosylation and phosphodiesterase activities to attach ubiquitin onto substrates, in great contrast to the orthodox pathway. A series of recent structural and mechanistic studies have clarified the action of these enzymes. Herein, we summarize the key insights into the structure and function of these proteins, emphasizing their modular nature, and discuss the biochemical implications of these proteins as well as areas of further exploration.

Highlights

The SidE family of *Legionella pneumophila* effectors can ubiquitinate host protein substrates on Ser residues via a novel mechanism considerably different from that of the canonical E1-E2-E3 pathway.

X-ray crystallographic data of these proteins reveal a unique three-domain arrangement and identify Ub and cofactor binding sites.

Biochemical studies have provided insights into the SidE mechanism and also discovered that flexible regions of proteins containing Ser residues can serve as ubiquitination targets.

Ubiquitination and the SidE Family

The post-translational modification known as ubiquitination, vital for cellular function and signaling in eukaryotic organisms, is defined as the attachment of the small protein ubiquitin (Ub) through its C terminus onto the lysine residue of a protein substrate via an isopeptide bond [1]. Ub is attached to the substrate by the coordinated, ATP-driven action of three enzymes termed as the Ub activating E1, Ub conjugating E2, and Ub E3 ligase that catalyze a cascade of reactions involving activation and sequential transfer of Ub, ultimately resulting in isopeptide linkage of Ub with target Lys residues. Addition of a single Ub via this E1-E2-E3 cascade results in monoubiquitination, which is often further elaborated by successive addition of more Ub groups to produce polyubiquitin chains in which the monomeric Ub units are isopeptide linked via one of the seven internal Lys residues, or the N-terminal Met of a preceding unit with the C terminus of a succeeding one. Chains formed between different residues represent distinct biological signals specifying distinct biological outcomes, forming the foundation of a complex signaling network based on Ub modification [2]. This modification plays a major role in a variety of cellular functions, including proteostasis, immunity, and trafficking, among others. Because of this major signaling role in cellular immunity, the Ub system is often the target of interference and manipulation by invading pathogens seeking to evade host responses [3]. This mode of host–pathogen interplay has been validated through the discovery of a wide array of bacterial effectors that interact with the Ub system [4]. It is to be noted that bacteria lack a Ub system of their own; nevertheless, effectors in bacterial genomes have been found to perform actions including deubiquitination [5,6], Ub ligation [7], and Ub modification [8].

Intriguingly, some bacterial proteins have evolved to interact with Ub through unconventional mechanisms never before seen in nature. The most striking example of this in the literature to date is the noncanonical ubiquitination of host substrates by the *Legionella pneumophila* SidE effector family. The SidE family contains four large, modular, and highly conserved proteins, SidE, SdeA, SdeB, and SdeC, that are required for optimal *Legionella* virulence. While the importance of this effector family in virulence has been known for more than a decade, their biochemical function remained elusive until recently. A key study reported that these four

[1]Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, IN 47906, USA
[2]Purdue Institute of Immunology, Inflammation, and Infectious Diseases and the Department of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, IN 47906, USA

*Correspondence:
luoz@purdue.edu (Z.-Q. Luo) and
cdas@purdue.edu (C. Das).

proteins catalyzed ubiquitination of host substrates using an unprecedented mechanism that differs significantly from the established route of ubiquitination in eukaryotes and represents the first and thus far only known example of ubiquitination occurring independently of the well-known E1-E2-E3 cascade [9]. The initial discovery of this new mechanism has since sparked intense worldwide research efforts in the past 2 years that have yielded a bounty of biological and biochemical insights into this process [10,11].

In this review article, we discuss the structural basis of this novel ubiquitination process, focusing on newly presented crystallographic data on the SidE proteins and their substrates (Table 1), and provide an overview of their separate domains and their interplay [12–16]. We also examine how these enzymes recognize their substrates and provide possible future directions and applications of this process, including its relevance to *Legionella* pathogenesis.

Early Work and Characterization of SidE Enzymatic Activity

Of all the known pathogenic bacteria, *L. pneumophila* has the largest repertoire of effectors, maintaining an arsenal of more than 330 individual proteins that is delivered into the host cell through its Dot/Icm secretion system [17,18]. While many of these effectors appear to be functionally redundant, a strain of *L. pneumophila* lacking the SidE family exhibits impaired growth of the bacteria within host cells [19]. Furthermore, expression of SidE proteins in yeast and mammalian cells leads to cytotoxic effects [20].

Table 1. Summary of Crystal Structures in PDB of SdeA and Homologs, with Bound Ligands and Residue Numbers

Protein	Construct	Domain(s)	Ligand(s)	PDB ID
SdeA	1–193	DUB	Ub-VME[a]	5CRA
SdeA	1–193 C118A	DUB		5CRB
SdeA	1–163	DUB (partial)		5CRC
SidE	222–1057	mART, PDE, CC (partial)		5ZQ2
SidE	222–1057	mART, PDE, CC (partial)	Ub (R42A)	5ZQ5
SidE	222–1057	mART, PDE, CC (partial)	Ub (R42A), NAD^+	5ZQ7
SidE	222–1057	mART, PDE, CC (partial)	Ub (R42A), ADPR	5ZQ6
SidE	222–589	PDE	Ub	5ZQ3
SidE	222–589	PDE	Ub, ADPR	5ZQ4
SdeA	231–1190	mART, PDE, CC (partial)		5YIM
SdeA	231–1190	mART, PDE, CC (partial)	Ub	5YIK
SdeA	231–1190	mART, PDE, CC (partial)	Ub, NADH	5YIJ
SdeA	211–910	PDE, mART		6B7Q
SdeD				6B7P
SdeD	1–341	PDE	Ub	6B7M
SdeD	1–341	PDE	ADPR-Ub, Ub	6B7O
SdeA	213–907	PDE, mART		6G0C
SdeA	765–905	mART (partial)		5YSJ
SdeA	765–905 E860/E862A	mART (partial)		5YSK
SdeA	765–905 E860/E862A	mART (partial)	NAD^+	5YSI

[a]VME, vinyl methyl ester.

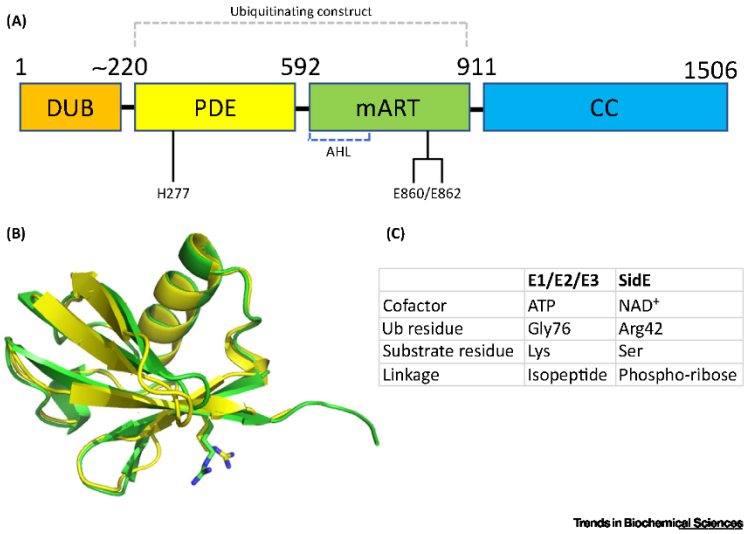
In pursuit of biochemical functions of this effector family, it was first noticed that the first roughly 200 residues of the SidE proteins harbored a deubiquitinase (DUB) domain (hereafter SdeA^{DUB}), capable of removing isopeptide-linked Ub from protein substrates as well as Ub chains with a distinct preference for Lys63-linked polyubiquitin chains [21]. The structure of SdeA^{DUB} revealed a core fold that was first characterized in the yeast desumoylase Ulp1 and human deneddylase (DEN1) [22]. This so-called Ulp1 fold is also shared by other bacterial DUBs and deneddylases in the CE-clan of prokaryotic proteases [23,24]. However, a strain of *L. pneumophila* lacking the SidE family regained full virulence when made to express SdeA with an inactive DUB mutation (C118A). This finding suggested that the DUB domain was not needed for bacterial virulence in the host systems used for these experiments.

In pursuit of further biochemical characterization, SidE proteins were shown to possess an unprecedented, all-in-one ubiquitinating activity outside the DUB domain, an activity that requires NAD^+ and a putative catalytic motif typically found in bacterial mono-ADP-ribosyltransferase (mART) toxins [9]. Indeed, Ub is ADP ribosylated at Arg42 forming ADP-ribosylated Ub (ADPR-Ub) as a reactive intermediate [9]. Further biochemical characterization revealed a two-step mechanism for SidE enzymes [10,11]. The ADPR-Ub intermediate produced in the first step is further processed by a novel phosphodiesterase (PDE) domain leading to phosphoribosyl (PR)-linked Ub conjugation to target Ser residues. This review article focuses on the ubiquitinating activity of these proteins, which is distinct from the DUB activity and important for bacterial virulence.

The SidE Ligase Catalytic Machinery

The core of the SidE family ubiquitination machinery comprises two catalytic units that work in sequence to catalyze PR-linked Ub attachment onto substrate Ser residues (Figure 1). In stark contrast to the conventional ATP-driven E1-E2-E3 cascade, the SdeA ubiquitination catalysis requires the nucleotide cofactor NAD^+. Overall, the underlying chemistry is strikingly different between the SidE all-in-one ubiquitination machinery and the three-enzyme eukaryotic system. The SidE ligase machinery carries out Ub activation using a mono-ADP-ribosyltransferase reaction whereby the ADP ribose group of the nucleotide cofactor is covalently added to Arg42 of Ub, forming ADPR-Ub. This enzymatic chemistry follows the same mechanism as seen in the well-known cholera toxin group of bacterial toxins that modify host proteins through mono-ADP-ribosylation of Arg [25]. The similarity in enzymatic chemistry is due to sharing of the characteristic RSE catalytic motif with the bacterial toxins [Arg for positioning of NAD^+ in the active site, Ser for stabilizing NAD^+ in its binding pocket, and Glu of the Glu(Gln)-X-Glu triad for promoting nucleophilic attack by the acceptor Arg], examples of which include Iota toxin from *Clostridium perfringens* [25], HopU1 of *Pseudomonas syringae* [26], and scabin from *Streptomyces scabies* [27]. The SidE mART domain (SidE^{mART}) spans approximately between residues 600 and 900 of the protein. An important distinguishing feature of SidE^{mART} is its ability to selectively accept Ub as the substrate for ADP-ribosylation, which implies the presence of Ub recognition elements as a part of the mART catalytic unit. The mART toxins catalyze ADP-ribosylation, resulting in a stable modification of the host protein, whereas SidE^{mART} catalyzes the formation of a species that undergoes further reaction.

Between the DUB and mART domains lies a distinct region that shares sequence similarity with another *Legionella* effector, the nucleotide binding protein known as lpg1496 of unknown function, and also with other known bacterial PDEs [10]. This PDE domain lies approximately between residues 200 and 600 of the protein (SidE^{PDE}). Outside the *L. pneumophila* genome, the closest similarity of this domain is with a cGMP dinucleotide PDE that shares 23% sequence identity, including the key conserved catalytic residues, a pair of His's and a Glu (His277,



	E1/E2/E3	SidE
Cofactor	ATP	NAD^+
Ub residue	Gly76	Arg42
Substrate residue	Lys	Ser
Linkage	Isopeptide	Phospho-ribose

Figure 1. Diagram Depicting the General Domain Organization of SidE Proteins. (A) Key catalytic residues are indicated (residue numbers pertain to SdeA). (B) Structural alignment of ubiquitin (Ub) (green ribbon) and NEDD8 (yellow) with R42 residues shown in sticks. (C) Comparison of key aspects of E1-E2-E3 and SidE-mediated ubiquitination reactions. AHL, α-helical lobe; CC, coiled coil; DUB, deubiquitinase; mART, mono-ADP-ribosyltransferase; PDE, phosphodiesterase.

His407 and Glu340, SdeA numbering) [28]. Despite being referred to as a PDE, the actual reaction catalyzed by this domain from the point of view of ubiquitination chemistry is akin to a phosphotransferase activity (rather than a hydrolase activity) wherein a substituted phosphate group as phospho-ribosyl-Ub is transferred from the ADPR-Ub donor to a substrate hydroxyl group. Mechanistically this is similar, in part, to the catalytic activity of certain bacterial His kinases [29,30]. In fact, the enzymatic chemistry in $SidE^{PDE}$ catalysis does involve a substituted phospho-His intermediate where a His is phosphorylated transiently before the attack of the substrate Ser residue [14].

While the sequential activity of both domains is required for substrate ubiquitination, each domain can function independently of the other, with *in vitro* experiments showing that the mART domain by itself is capable of producing ADPR-Ub and the PDE domain is able to use the intermediate in *trans* to ubiquitinate substrates or produce PR-Ub [12]. However, close interactions between the two domains as found in the full-length core construct $SidE^{Core}$ (defined as the SidE segment encompassing the PDE and mART domains) are required for maximal activity [12,14]. Conserved residues at the interface of the mART and PDE domains stabilize interdomain packing that provides better catalytic efficiency of $SidE^{Core}$.

Beyond the mART domain (residues 900 to end), lies a C-terminal coiled-coil (CC) domain that ultimately is not required for catalytic activity *in vitro* but may play a role ultimately in localization of the enzymes in host cells, it also carries the translocation signals recognized by the Dot/Icm

protein secretion system. It has been shown that a portion of the CC domain is used in interactions with the components of the translocation machinery [15], presumably for injection of the effector into the host cytoplasm.

Overview of Structural Data

Four independent studies have been thus far published in 2018 reporting crystal structures of different core constructs of SidE proteins, mainly SdeA, either in its apo form or in complex with various ligands, including NAD^+ (and NADH), Ub (and its R42A mutant), and the intermediate ADPR-Ub (Table 1) [12–15]. In addition, a recent paper on the crystal structure of the mART domain has contributed additional insights into the mechanism of catalysis by the SidE ligase [16]. The crystal structures confirmed the biochemical prediction of SidE proteins possessing distinct mART, PDE, and CC domains (Figure 2) [10]. In addition to core constructs, some of these structures included parts of the CC domain, which appears to be conformationally flexible relative to the PDE-mART core, adopting different orientations in a length-different manner (different truncations of the CC domain were captured, but not the entire predicted domain). Although not absolutely essential for ubiquitination activity, it may play a role in relative positioning of the PDE domain with respect to the mART domain, thereby supporting appropriate interdomain packing for optimal activity. In one construct, the SidE CC domain forms a domain-swapped dimer in crystal and in solution [12]. Nevertheless, the PDE and mART domains are invariantly placed next to each other in the structure. The active sites of the two domains face in different directions with a substantial spatial separation, indicating the independence of the two catalytic centers (Figure 2B).

mART Domain and the α-Helical Lobe

The structure of SdeA mART domain features two distinct lobes typically seen in bacterial mART enzymes, a primarily α-helical lobe (AHL; between approximately residues 600 and 750) and a main lobe (mART core) comprising a β sandwich structure [25]. The AHL in other bacterial mART proteins lies in close proximity to the core, creating a binding pocket for NAD^+ that is lined with residues from both lobes. In the crystal structures of two SdeA constructs lacking the CC domain, the AHL is observed protruding out from the remainder of the mART core connected by a flexible hinge in what seems to be an unproductive arrangement of the two lobes. Since these constructs are catalytically active, AHL likely moves closer to the mART core during catalysis. However, in constructs containing parts of the CC region the AHL is observed in a canonical arrangement as seen in bacterial mARTs. Inclusion of the CC domain does seem to stabilize a productive form of mART domain in solution, with the full-length SdeA construct showing higher activity than the constructs lacking the CC region [14]. The key catalytic residues of the RSE motif are present in the main lobe, which share substantial sequence similarity with the bacterial mART toxin HopU1 (~21%) [26]. In productive arrangement, the NAD^+ binding pocket is located within a cleft between the two lobes. Three co-crystal structures (NAD^+ and NADH bound with mART domain) reveal key contacts between the nucleotide and residues lining the NAD^+ binding pocket [12,15,16] (Figure 3, Key Figure). NAD^+ is held in a position poised for reaction, with the nucleotide adopting a strained-ring conformation in which the nicotinamide group (NAM) is forced to stack against the AMP group. Similar strained conformation of NAD^+ has been observed as a key catalytic feature in mART structures of other bacterial toxins, where the strain conformation has been proposed to facilitate the nucleophilic attack of the substrate Arg residue through an SN1 mechanism, described as SN1-strain alleviation mechanism [25]. Ub has been captured bound to the mART domain in two co-crystal structures, also with bound

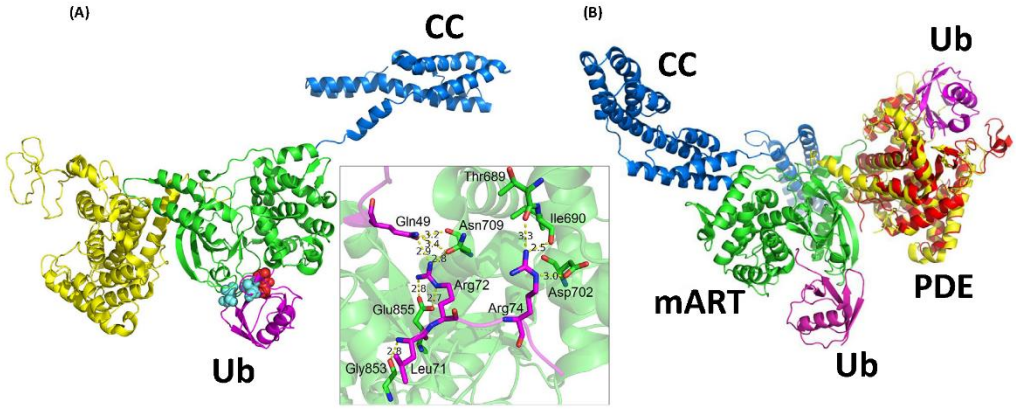


Figure 2. Ubiquitin Recognition by the SidE Ligase. (A) Structure of SidE in complex with ubiquitin (Ub) (magenta) at mono-ADP-ribosyltransferase (mART) site with key interacting residues of Ub highlighted. The hydrophobic patch of Ub is depicted in cyan, and key Arg residues are depicted in red. Inset depicts polar/charge interactions between the SidE mART domain and Ub. (B) Structure of SdeA in complex with Ub at mART site (magenta) aligned with structure of phosphodiesterase (PDE) domain of SdeD (red) in complex with ADPR-Ub (magenta). The SdeD *Legionella pneumophila* effector has a domain homologous to the SidE PDE domain with the same catalytic residues but lacks an mART domain. CC, coiled coil.

nucleotide [12,15]. Interestingly, Ub captured in the co-crystal structures appears to represent a pre-catalytic substrate recognition state with the actual nucleophile Arg42 facing away from the C1′ ribose attack center (the ribose group attached to NAM in NAD^+, $ribose^{Nic}$; Figure 3A,B) with a distance of 11 Å between the $N^{\omega}H$ group and the C1′ atom. Instead, Arg72 is closer to the C1′-ribose center as if it were the true nucleophile. Arg72 does play an important role in substrate recognition; however, during the actual reaction, it must be repositioned to allow Arg42 to take its place proximal to the reaction center. It thus appears that Ub recognition might proceed in at least two distinct steps: an initial engagement that places Arg72 proximal to the $C1\text{-}ribose^{NAM}$ atom, followed by a rearrangement to bring the nucleophilic Arg42 adjacent to the ribose moiety. Additional structural work is required to capture the correct catalytic state of Ub in its binding site in SdeA mART domain.

Ub Recognition by SidE mART

The Ub-bound structures reveal an overall position for Ub that is consistent with reaction with NAD^+ for the ADP-ribosylation reaction [12,15]. Ub approaches the mART center by facing the NAD^+ cleft and is perched on a concave surface contributed by both the AHL and main lobes. The AHL lobe plays a critical role in Ub recognition alongside the catalytic motif of the main lobe, with major interactions occurring between Ub C-terminal tail residues Arg72 and Arg74 and two acidic pockets in SdeA comprised of the catalytic Glu860/Glu862 (main lobe) and the AHL Asp691/Asp707, respectively. Additional interactions of the Ub's I44 patch (a hydrophobic patch of three side chains Leu8/Ile44/Val70, widely used in Ub recognition by eukaryotic enzymes in the Ub pathway) with a complementary hydrophobic patch on the main lobe (Val822/Ile826/Phe827) contribute to the interface and the observed orientation of Ub in its binding site (Figure 2A). As noted, the actual catalytic state would require a reorientation of Ub to bring R42 to face the NAD^- binding pocket. Mutation of Arg72 or Arg74 results in significant

Key Figure

The Proposed Mechanism of SidE-Catalyzed Ubiquitination

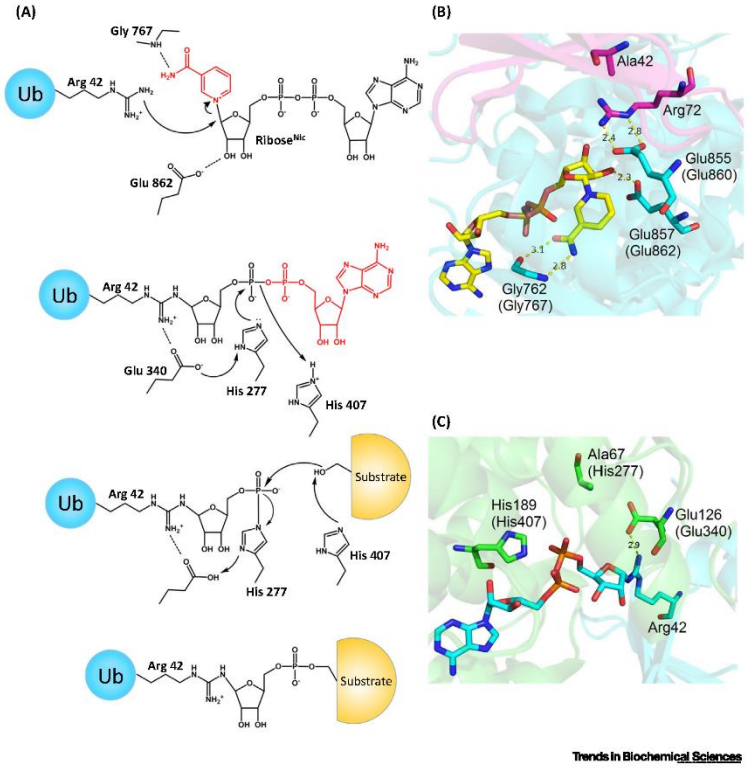


Figure 3. (A) Key residues are included (residue numbers pertain to SdeA). (B) Structure of SidE bound to NAD^+ and ubiquitin (Ub) R42A, depicting a likely pre-catalytic state analogous to ADP-ribosylation of R42, corresponding to the scheme in (A). (C) Structure of SdeD bound to ADP-ribosylated ubiquitin (ADPR-Ub), analogous to the phosphotransfer step of the proposed mechanism catalyzed by the phosphodiesterase domain. SdeA residue numbers are included in parentheses. Nic, nicotinamide.

loss of mART activity, and the interaction involving Arg72 may serve as a determinant of selectivity of SidE mART for Ub over NEDD8. In fact, mutation of the Ala72 to Arg in NEDD8 results in appreciable ADP-ribosylation of NEDD8 [10]. Thus, interactions with Arg72 may contribute to substrate selection by SidE mART domain.

The last two Gly residues (Gly75-Gly76) and the C-terminal carboxylate group are not engaged and are solvent exposed [12,15], which explains previous biochemical results showing Ala

substitution at the Gly75-Gly76 segment and addition of another Ub group, as in di-Ub, are tolerated in mART function [9,31].

PDE Domain

The structure of the PDE domain reveals discernible similarities with other PDE enzymes while at the same time revealing unique features required for a phosphotransfer reaction. A key feature specific to the SidE family is a cap lobe, which is necessary for Ub binding and critical for activity. In SdeA, it had been known that mutation of His277 or His407 caused a loss of PDE activity [10]. When His407 was changed to Asn, reaction with Ub and NAD^+ caused the accumulation of a stable intermediate between His277 and the phosphate group of PR-Ub [14]. These observations strongly suggest that H277 acts as a nucleophile to attack ADPR-Ub, further supported by a hydrogen bonding interaction with nearby Glu340 in the crystal structure, with the latter serving to stabilize the deprotonated form of the imidazole group (the nucleophile) [15]. A possible double role for H407 has been proposed: first, as a proton donor to promote the departure of the AMP group of ADPR-Ub, leading to the H277-PR-Ub intermediate; and second, as a mediator of the attack of a substrate Ser residue on the same intermediate (Figure 3A) [13,14]. This proposed mechanism up to AMP departure is reminiscent of many bacterial histidine kinases [29]; in SidE catalysis, the His-PR bond is further subjected to an additional nucleophilic attack by a substrate Ser residue.

Ub (ADPR-Ub) Recognition

One study resolved the structure of the PDE domain of SidE (222–589) in complex with Ub, revealing that a stretch of N-terminal residues Lys6-Thr9 of Ub were a major site of interaction [12]. Another structure of the PDE domain bound to ADP-ribose found that it also forms extensive contacts with its ADP moiety, with the orientation of the catalytically relevant His residues in agreement with the proposed mechanism mentioned above (Figure 3C). A co-crystal structure of ADPR-Ub bound to a related *Legionella* effector SdeD also agreed with these results [13]. SdeD, while sharing a similar PDE domain including the same three catalytic residues (His-His-Glu) with the SidE family, lacks DUB, mART, and CC domains; therefore, its role in a biological setting remains unclear. It is noteworthy that the Ub binding sites of the mART and PDE domains are not near one another (Figure 2B), and further studies are required to examine the mechanism by which ADPR-Ub moves to the PDE domain for the second step of catalysis.

CC Domain

An enigmatic region of the SidE proteins has been the C-terminal domain. Interestingly, structural data have revealed that it likely plays a major role in AHL stabilization. A comparison of crystal structures of SidE family members shows that the AHL is positioned away from the mART core in structures where the CC domain is absent [12–15] (Figure 4A). This conformation is likely unfavorable for Ub binding. Biochemical analysis shows that the CC domain is required for optimum mART activity, possibly via stabilization of the AHL in a productive arrangement with respect to the mART core, supported by structural data showing that a combination of electrostatic and hydrophobic interactions hold together the AHL and the CC domain (Figure 4B), as well as another set of interactions between the mART core and the CC domain [12,15] (Figure 4C). However, further studies are required to determine the full extent of CC-mART and possible CC-PDE domain interactions.

Another role for the CC domain was investigated through binding assays. These assays showed that $SdeA^{1101\ 1350}$ was able to bind tightly to IcmS-IcmW-DotL, a *Legionella* adaptor protein complex involved in secretion of effectors [15,32,33].

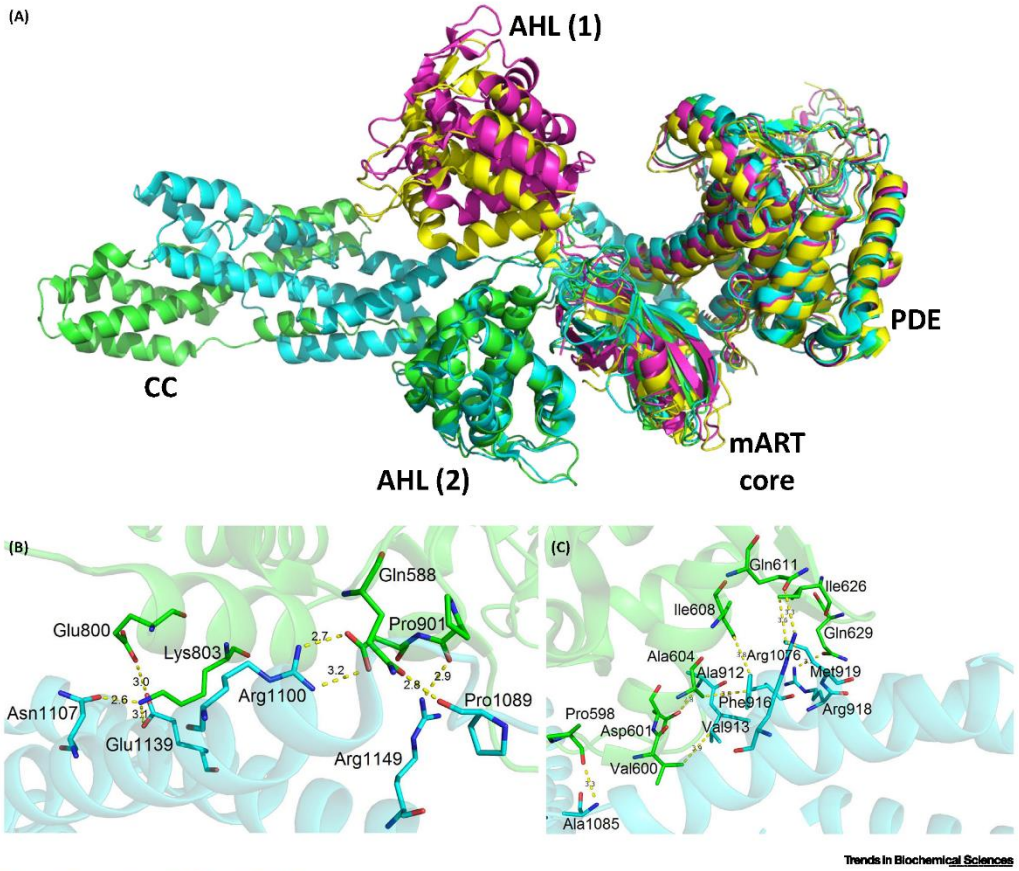


Figure 4. Comparison of SidE and SdeA Structures. (A) Structural alignment of various constructs of SidE proteins (green, SidE$^{222-1057}$; cyan, SdeA$^{231-1190}$; magenta, SdeA$^{211-910}$; yellow, SdeA$^{213-997}$), with domains labeled. AHL (1) and AHL (2) refer to two different orientations of the α-helical lobe (AHL) domain observed in crystal structures of the SidE constructs, with the latter corresponding to a productive orientation. (B) Interactions between mono-ADP-ribosyltransferase (mART) core and coiled-coil (CC) domain of SdeA. (C) Interactions between mART AHL and CC domain.

Interestingly, one study resolved a crystal structure of SdeA with two Ub moieties bound to the CC domain [15]. No previous evidence exists of the CC domain directly able to bind Ub; therefore, the biological and biochemical relevance of this phenomenon will need to be revisited in further study. The CC domain appears to also a play a role in promoting dimer formation of SidE and some SdeA constructs, the significance of which is not known. Two of the studies found that SidE$^{222-1057}$ and SdeA^{FL} constructs form dimers in solution [12,14], where a third study found that SdeA$^{231-1190}$ is a monomer [15]. The basis for this conflict in SdeA may be truncation of the CC domain, as a truncated SidE construct behaved as a monomer in solution compared to the longer construct.

Substrate Recognition

Recently, great advances have been made in our understanding of how SidE proteins identify their protein substrates. Initially, five substrates had been identified: Rab1a, Rab30, Rab33b, Rab6a, and Rtn4 [9,11]. In addition, the SidE proteins self-ubiquitinate. Two other proteins, Rab5a and Rac1, were intriguingly found not to be ubiquitinated, despite considerable structural similarity to the four other Rab substrates [9]. Mass spectrometric analysis identified Ser154 of Rab33b as a site of ubiquitination [10]. However, mutation of Ser154 to Ala failed to prevent ubiquitination, suggesting the presence of additional sites. New data have shown that for substrate ubiquitination, SidE proteins do not appear to have a structural preference [12,14]. Instead, they are able to ubiquitinate Ser residues that are part of any unstructured region that can fit into the PDE groove containing the catalytic site. For example, Ser residues near the N terminus of the known substrate proteins were found to be major targets. This hypothesis was further validated by the fact that one can transform a non-substrate protein such as Rac1 into a robust substrate by adding a flexible Ser-containing region [12]. Such manipulation even allowed structurally dissimilar proteins such as Ub and glutathione S-transferase to be ubiquitinated by SdeA. According to this model, the originally identified Ser154 of Rab33b is likely targeted due to its location in the flexible G4 loop region of the protein [34], but the primary site is the N terminus, which presents a larger unstructured region, with many more Ser residues. Furthermore, by screening Rtn4 peptides, hydrophobic residues around the target Ser appeared to improve ubiquitination efficiency [14]. These new insights suggest that SidE proteins' localization rather than structure will dictate their true biological targets. The ability to add a monoubiquitination site onto a wide variety of proteins or peptides through a flexible Ser-containing region may also prove a valuable chemical biology tool in future studies.

Concluding Remarks

One eminent feature of the hundreds of effectors of *L. pneumophila* is that few of them are necessary for virulence in commonly used host systems [35]. Members of the SidE family display clear functional redundancy because the bacterium displays a defect in intracellular growth in protozoan hosts only when all members of the family were deleted, and such a defect can be complemented by any member of the family [19,20]. Deletion of *sidE*s caused a delay in the conversion of its vacuolar membrane into an endoplasmic reticulum (ER)-like compartment, which is the result of ubiquitination of multiple proteins involved in ER structure and function. The fact that SidEs attack structurally diverse proteins associated with the ER such as Rab small GTPases and reticulon 4 (Rtn4) suggests a large repertoire of their cellular targets [11]. Further studies are necessary to obtain a more complete list of substrates of SidEs by taking advantage of Ub variants that cannot be used by the canonical mechanism (see Outstanding Questions). Ubiquitination of Rtn4 results in the formation of ER tubules and its enrichment on the membrane of the bacterial vacuole, but the effects on the activity Rab proteins are less clear because modified Rab33b only is marginally defective in GTP hydrolysis [9]. Thus, it is unclear how ubiquitination induced by SidEs changes the function of their targets and how such changes benefit the bacterium. Another intriguing question is whether ubiquitination driven by ADP-ribosylation also exists in eukaryotes and other pathogens, or it is exclusive to the SidE family. Indeed, eukaryotes including humans possess a family of ADP-ribosylating enzymes named ectoARTs that contain the RSE motif and that are related to the RSE bacterial mART toxins [9]. The better-studied poly(ADP-ribose) polymerase family in eukaryotes is also related, albeit more distantly [36]. However, it has yet to be determined whether these eukaryotic enzymes act only alone or are part of a novel ubiquitination scheme. Answers to these questions will surely expand our current appreciation of the role of Ub in cell signaling and disease development.

Outstanding Questions

Theoretically, any protein with a flexible Ser-containing region may be a substrate for ubiquitination by SidE effectors. What, then, are the primary host protein targets of SidE? What is the biological outcome of PR-linked ubiquitination? Are other hydroxyl-containing residues such as Thr or Tyr targeted?

How does the intermediate ADPR-Ub move from the mART to the relatively distant PDE domain to undergo phosphotransfer?

What is the biological function of SdeD, a homologous effector altogether lacking an mART or CC domain?

Does the SidE ubiquitination mechanism exist as a eukaryotic cellular process?

Acknowledgments

We kindly acknowledge the NIH (1R01GM126296) for financial support.

References

1. Hershko, A. and Ciechanover, A. (1998) The ubiquitin system. *Annu. Rev. Biochem.* 67, 425–479
2. Komander, D. and Rape, M. (2012) The ubiquitin code. *Annu. Rev. Biochem.* 81, 203–229
3. Bhavsar, A.P. *et al.* (2007) Manipulation of host-cell pathways by bacterial pathogens. *Nature* 449, 827–834
4. Zhou, Y. and Zhu, Y. (2015) Diversity of bacterial manipulation of the host ubiquitin pathways. *Cell Microbiol.* 17, 26–34
5. Mevissen, T.E.T. and Komander, D. (2017) Mechanisms of deubiquitinase specificity and regulation. *Annu. Rev. Biochem.* 86, 159–192
6. Ronau, J.A. *et al.* (2016) Substrate specificity of the ubiquitin and Ubl proteases. *Cell Res.* 26, 441–456
7. Huibregtse, J. and Rohde, J.R. (2014) Hell's BELs: bacterial E3 ligases that exploit the eukaryotic ubiquitin machinery. *PLoS Pathog.* 10, e1004255
8. Cui, J. *et al.* (2010) Glutamine deamidation and dysfunction of ubiquitin/NEDD8 induced by a bacterial effector family. *Science* 329, 1215–1218
9. Qiu, J. *et al.* (2016) Ubiquitination independent of E1 and E2 enzymes by bacterial effectors. *Nature* 533, 120–124
10. Bhogaraju, S. *et al.* (2016) Phosphoribosylation of ubiquitin promotes serine ubiquitination and impairs conventional ubiquitination. *Cell* 167, 1636–1649.e13
11. Kotewicz, K.M. *et al.* (2017) A single *Legionella* effector catalyzes a multistep ubiquitination pathway to rearrange tubular endoplasmic reticulum for replication. *Cell Host Microbe* 21, 169–181
12. Wang, Y. *et al.* (2018) Structural insights into non-canonical ubiquitination catalyzed by SidE. *Cell* 173, 1231–1243.e16
13. Akturk, A. *et al.* (2018) Mechanism of phosphoribosyl-ubiquitination mediated by a single *Legionella* effector. *Nature* 557, 729–733
14. Kalayil, S. *et al.* (2018) Insights into catalysis and function of phosphoribosyl-linked serine ubiquitination. *Nature* 557, 734–738
15. Dong, Y. *et al.* (2018) Structural basis of ubiquitin modification by the *Legionella* effector SdeA. *Nature* 557, 674–678
16. Kim, L. *et al.* (2018) Structural and biochemical study of the mono-ADP-ribosyltransferase domain of SdeA, a ubiquitylating/deubiquitylating enzyme from *Legionella pneumophila*. *J. Mol. Biol.* 430, 2843–2856
17. Luo, Z.-Q. and Isberg, R.R. (2004) Multiple substrates of the *Legionella pneumophila* Dot/Icm system identified by interbacterial protein transfer. *Proc. Natl. Acad. Sci. U. S. A.* 101, 841–846
18. Ensminger, A.W. (2016) *Legionella pneumophila*, armed to the hilt: justifying the largest arsenal of effectors in the bacterial world. *Curr. Opin. Microbiol.* 29, 74–80
19. Bardill, J.P. *et al.* (2016) IcmS-dependent translocation of SdeA into macrophages by the *Legionella pneumophila* type IV secretion system. *Mol. Microbiol.* 56, 90–103
20. Havey, J.C. and Roy, C.R. (2015) Toxicity and SidJ-mediated suppression of toxicity require distinct regions in the SidE family of *Legionella pneumophila* effectors. *Infect. Immun.* 83, 3506–3514
21. Sheedlo, M.J. *et al.* (2015) Structural basis of substrate recognition by a bacterial deubiquitinase important for dynamics of phagosome ubiquitination. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15090–15095
22. Reverter, D. *et al.* (2005) Structure of a complex between Nedd8 and the Ulp/Senp protease family member Den1. *J. Mol. Biol.* 345, 141–151
23. Rawlings, N.D. *et al.* (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 40, D343–D350
24. Pruneda, J.N. *et al.* (2016) The molecular basis for ubiquitin and ubiquitin-like specificities in bacterial effector proteases. *Mol. Cell* 63, 261–276
25. Tsurumura, T. *et al.* (2013) Arginine ADP-ribosylation mechanism based on structural snapshots of iota-toxin and actin complex. *Proc. Natl. Acad. Sci. U. S. A.* 110, 4267–4272
26. Jeong, B. *et al.* (2011) Structure function analysis of an ADP-ribosyltransferase type III effector and its RNA-binding target in plant immunity. *J. Biol. Chem.* 286, 43272–43281
27. Lyons, B. *et al.* (2016) Scabin, a novel DNA-acting ADP-ribosyltransferase from Streptomyces scabies. *J. Biol. Chem.* 291, 11198–11215
28. Rinaldo, S. *et al.* (2015) Structural basis of functional diversification of the HD-GYP domain revealed by the *Pseudomonas aeruginosa* PA4781 protein, which displays an unselective bimetallic binding site. *J. Bacteriol.* 197, 1525–1535
29. Klumpp, S. and Krieglstein, J. (2002) Phosphorylation and dephosphorylation of histidine residues in proteins. *Eur. J. Biochem.* 269, 1067–1071
30. Matte, A. *et al.* (1998) How do kinases transfer phosphoryl groups? *Structure* 6, 413–419
31. Puvar, K. *et al.* (2017) Ubiquitin chains modified by the bacterial ligase SdeA are protected from deubiquitinase hydrolysis. *Biochemistry* 56, 4762–4766
32. Cambronne, E.D. and Roy, C.R. (2007) The *Legionella pneumophila* IcmSW complex interacts with multiple Dot/Icm effectors to facilitate type IV translocation. *PLoS Pathog.* 3, e188
33. Kwak, M.-J. *et al.* (2017) Architecture of the type IV coupling protein complex of *Legionella pneumophila*. *Nat. Microbiol.* 2, 17114
34. Muratcioglu, S. *et al.* (2015) GTP-dependent K-Ras dimerization. *Structure* 23, 1325–1335
35. Qiu, J. and Luo, Z.-Q. (2017) *Legionella* and *Coxiella* effectors: strength in diversity and activity. *Nat. Rev. Microbiol.* 15, 591–605
36. Cohen, M.S. and Chang, P. (2018) Insights into the biogenesis, function, and regulation of ADP-ribosylation. *Nat. Chem. Biol.* 14, 236–243

Biochemistry

Communication

pubs.acs.org/biochemistry

Ubiquitin Chains Modified by the Bacterial Ligase SdeA Are Protected from Deubiquitinase Hydrolysis

Kedar Puvar,[†] Yiyang Zhou,[†] Jiazhang Qiu,[‡,§] Zhao-Qing Luo,[‡] Mary J. Wirth,[†] and Chittaranjan Das*[,†]

[†]Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, Indiana 47906, United States

[‡]Purdue Institute of Immunology, Inflammation, and Infectious Diseases and Department of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, Indiana 47906, United States

[§]Key Laboratory of Zoonosis, Ministry of Education, Institute of Zoonosis, College of Veterinary Medicine, Jilin University, Xi'an Road 5333, Changchun, Jilin 130062, China

Supporting Information

ABSTRACT: The SidE family of *Legionella pneumophila* effectors is a unique group of ubiquitin-modifying enzymes. Along with catalyzing NAD^+-dependent ubiquitination of certain host proteins independent of the canonical E1/E2/E3 pathway, they have also been shown to produce phosphoribosylated free ubiquitin. This modified ubiquitin product is incompatible with conventional E1/E2/E3 ubiquitination processes, with the potential to lock down various cellular functions that are dependent on ubiquitin signaling. Here, we show that in addition to free ubiquitin, Lys63-, Lys48-, Lys11-, and Met1-linked diubiquitin chains are also modified by SdeA in a similar fashion. Both the proximal and distal ubiquitin moieties are targeted in the phosphoribosylation reaction. Furthermore, this renders the ubiquitin chains unable to be processed by a variety of deubiquitinating enzymes. These observations broaden the scope of SdeA's modulatory functions during *Legionella* infection.

Ubiquitin (Ub) signaling regulates many key processes in eukaryotic cells.[1] This occurs by post-translationally modifying a substrate protein with the conserved 76-residue protein ubiquitin. Traditionally, ubiquitination has been known to occur via a coordinated three-step mechanism, requiring the E1 (ubiquitin-activating), E2 (ubiquitin-conjugating), and E3 (ubiquitin-ligating) enzymes along with the cofactor ATP, ultimately resulting in an isopeptide bond between the C-terminus of Ub and an acceptor lysine residue of the substrate protein. Often, polyubiquitin chains are generated on substrate proteins via successive addition of Ub groups to the first one.[1] Ubiquitin possesses seven Lys residues that, along with the N-terminal amino group of Met1, serve as points through which Ub groups in polyubiquitin chains may be connected. The simplest of this polyubiquitin chains are homotypic polymers that are produced by linking through the same Lys residue; accordingly, the homotypic chains are named after the linking Lys, such as Lys63- or Lys48-linked chains. In addition, complex chain types involving mixed linkages and branching generate an immense diversity of polyubiquitin chains serving different signaling functions. For example, polyubiquitin chains linked via Lys48 of ubiquitin (K48) are known to be a signal for proteasomal degradation, while K63-linked chains are implicated in lysosomal degradation, cellular trafficking, and DNA repair.[1] Importantly, ubiquitination is a reversible process. The hydrolysis of Ub isopeptide linkages is achieved by a group of enzymes collectively known as deubiquitinases, or deubiquitinases (DUBs).[1−3] Several categories of human DUBs have been identified to date; these include the USP, UCH, OTU, MJD, and MINDY[4] families, cleaving ubiquitin linkages through a cysteine protease mechanism, and the JAMM family, utilizing a zinc metalloprotease mechanism.[5,6] Precise processing of ubiquitin chains by DUBs serves an important regulatory function in eukaryotes and is crucial for normal cellular function.

While prokaryotes do not possess a ubiquitin system of their own, remarkably there are several examples in nature of pathogens that are able to control host ubiquitin signaling for their benefit,[7,8] through both ubiquitin ligation and deubiquitination. The infectious bacterial organism *Legionella pneumophila*, responsible for Legionnaire's disease,[9] is one such pathogen. It exerts its control by injecting hundreds of effector proteins via the Dot/Icm type IV secretion system into the cytoplasm of the host cell during infection.[10] Among these effectors is the SidE family, which includes four proteins, the paralogs SdeA, SdeB, SdeC, and SidE. This family is required for optimal *Legionella* replication within the host. SdeA was recently observed to possess two noteworthy functions. First, it was able to catalyze deubiquitination[11] through its N-terminal domain (the $SdeA^{DUB}$ domain, spanning the first 200-residue segment). The second function, and pertaining to the focus of this study, is that it ubiquitinated certain human substrate proteins in a manner completely different from the canonical E1/E2/E3 cascade,[12] with NAD^+ instead of ATP as a necessary cofactor and ADP-ribosylated Arg42 of Ub (Ub-ADPR) as the activated Ub intermediate. The Ub-ADPR intermediate is produced by a mono-ADP ribosyl transferase domain present within the segment of residues ∼500−1000 of the effector ($SdeA^{mART}$). A further study[13] showed that this intermediate is subsequently attacked by a substrate Ser residue, resulting in ubiquitination of the Ser residue through a phosphor−ribosyl (PR) linkage between the target and Arg42 of Ub. This

Received: July 13, 2017
Revised: August 14, 2017
Published: August 15, 2017

DOI: 10.1021/acs.biochem.7b00664
Biochemistry XXXX, XXX, XXX−XXX

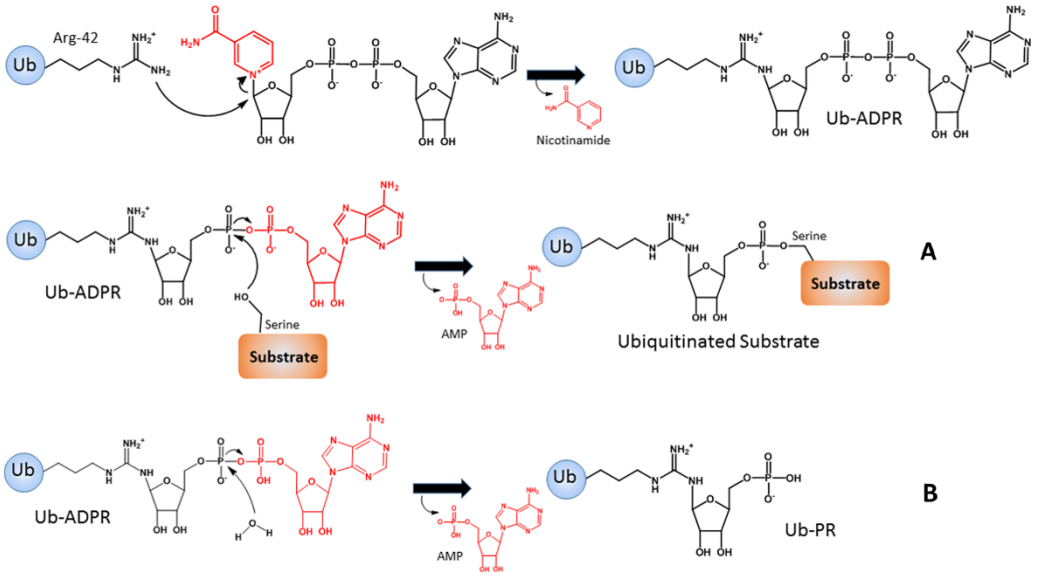


Figure 1. Proposed two-step mechanism for SdeA-catalyzed ubiquitination of the substrate Rab33b (A) as well as generation of Ub-PR (B) upon hydrolysis of the Ub-ADPR intermediate. Reaction byproducts and leaving groups are colored red.

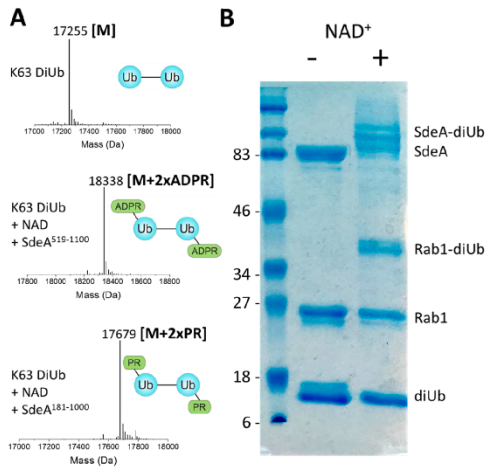


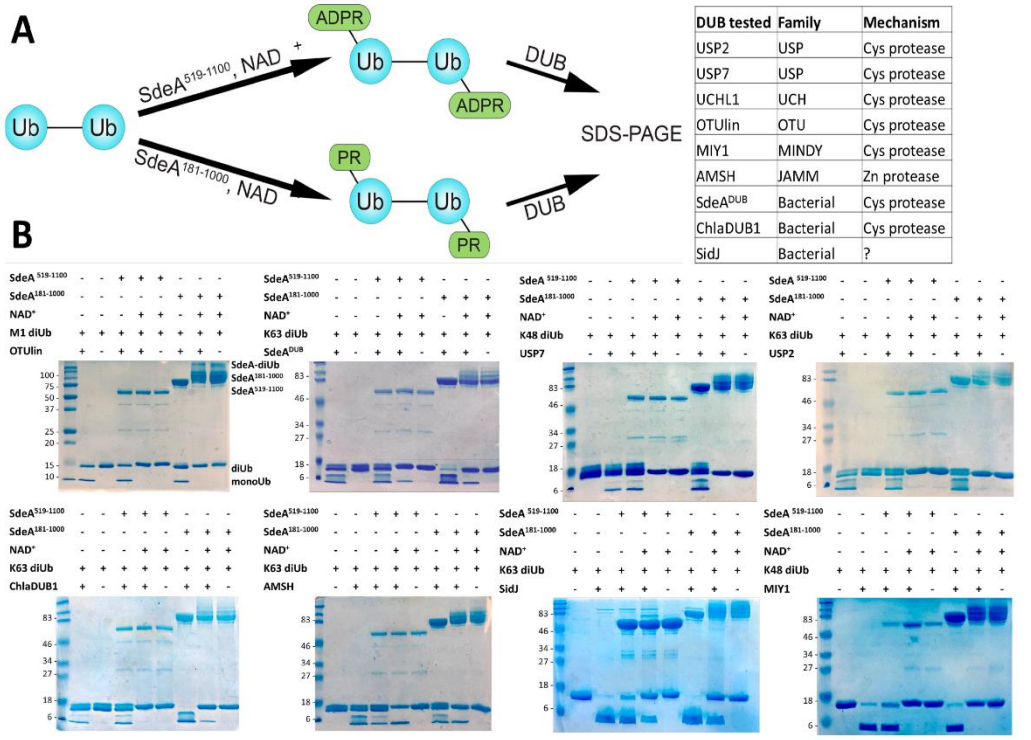
Figure 2. Reactions of diubiquitin with SdeA. (A) MS of diUb incubated with SdeA constructs and NAD^+ (see Figure S1 for LC–MS data of other diubiquitin chain types and Table S1 for MS data analysis). (B) Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) of diUb and $SdeA^{181-1000}$ incubated with the substrate Rab1 in the presence or absence of NAD^+.

ubiquitination step is catalyzed by a phosphotransferase domain (dubbed the PDE domain)[13] harbored within the segment of residues ~200–500 of the protein. These studies showed ubiquitination taking place in any fashion besides E1/E2/E3 and through a linker besides the isopeptide bond for the first time. The Ub-ADPR intermediate is also subject to hydrolysis, with water cleaving the bond between the phosphate groups, in a reaction that is also catalyzed by the same PDE domain (Figure 1). The hydrolysis results in the generation of free phosphoribosylated ubiquitin (Ub-PR) (Figure 1). Both outcomes are accompanied by the release of AMP as a byproduct. Ub-PR was shown not to participate in any of the three steps of the canonical ubiquitination pathway.[13] However, the ADP-ribosylating and PDE activity of SdeA against ubiquitin chains as well as their effects on DUBs has remained unexplored. An understanding of these effects is of great interest due to the fact that polyubiquitinated proteins are known to associate with the LCV (*Legionella*-containing vacuole) during infection.[14]

In this study, we demonstrate that Ub chains of different linkage types are modified by SdeA. We also show that both Ub monomers of diubiquitin (diUb) chains undergo modification, suggesting that SdeA is nonselective in its ADP-ribosylating activity as well as its PDE activity. These modified ubiquitin chains were then tested against several different DUBs. We report that ADP ribosylation as well as phosphoribosylation of diUb by SdeA leads to a significant protection against DUBs that engage Arg42 of Ub during hydrolysis. We explored the basis for this effect through selective modification of the proximal versus the distal ubiquitin species and structural modeling. These data suggest that SdeA may prevent the processing of existing ubiquitin chains by host DUBs.

In a recent study,[15] Kotewicz et al. demonstrated that K63-linked tetra-Ub chains were ADP-ribosylated by SdeA. However, which ubiquitin monomers underwent modification remained unknown. To investigate this, we tested the activity of SdeA *in vitro* against a diUb substrate. K63-, K48-, K11-, and M1-linked DiUb were incubated with two constructs of SdeA: the ubiquitinating construct $SdeA^{181-1000}$ and $SdeA^{519-1100}$, which contains the mART domain but lacks the PDE domain.

DOI: 10.1021/acs.biochem.7b00664
Biochemistry XXXX, XXX, XXX–XXX



DUB tested	Family	Mechanism
USP2	USP	Cys protease
USP7	USP	Cys protease
UCHL1	UCH	Cys protease
OTUlin	OTU	Cys protease
MIY1	MINDY	Cys protease
AMSH	JAMM	Zn protease
SdeA^{DUB}	Bacterial	Cys protease
ChlaDUB1	Bacterial	Cys protease
SidJ	Bacterial	?

Figure 3. DUB cleavage assays of modified vs unmodified diUb species. (A) Overall scheme of assays and list of DUBs used. (B) SDS−PAGE gels of reactions between DUB and diUb that had been pretreated with SdeA constructs in the presence or absence of NAD^+. For UCHL1, Ub-W was used as a substrate and reactions were analyzed by LC−MS. See Figure S2 for complete LC−MS results and Table S2 for detailed analysis of MS data.

The reaction mixture was analyzed using top-down liquid chromatography and mass spectrometry (LC−MS) (Figure 2A and Figure S1) to determine whether one or both ubiquitin moieties were modified. The masses of the reaction products indicated that a doubly modified diUb species was generated (Table S1). Therefore, SdeA does not distinguish between free Ub and Ub that is part of a polyubiquitin chain. In addition, within the time scale of our reaction, no preference for either the proximal (C-terminus-engaged) or distal (Lys-engaged) moiety was observed, due to an absence of peaks for singly ADP-ribosylated or phosphoribosylated diUb species.

There appeared to be no preference for chain type, with K11-, K48-, K63-, and M1-linked chains all serving as viable substrates, being converted to their modified form within the time of our observation. We also found that SdeA is able to transfer diUb to Rab1, a previously reported substrate, as well as to itself (self-modification, a common feature of ubiquitin ligases) (Figure 2B).

It has also been suggested that modified Ub chains are resistant to hydrolysis from the DUB domain of SdeA.[15] To better understand this process and determine whether this effect could also be extended to other DUBs, the activities of the DUB domain of SdeA, the related bacterial DUB ChlaDUB1,[16] and the recently described *L. pneumophila* DUB SidJ[17] along with that of the human DUBs USP2, USP7, OTUlin, AMSH, and UCHL1 and the yeast DUB MIY1 were tested against unmodified, ADP-ribosylated, and phosphoribosylated ubiquitin substrates. K63-linked diUb was used as the substrate for USP2, AMSH, SdeA, and ChlaDUB1; K48-linked diUb was used for USP7, and M1-linked diUb was used for OTUlin. Ubiquitin with a C-terminal tryptophan (Ub-W) was used as a substrate for UCHL1 because of its preference for cleaving small peptides from the C-terminus of Ub rather than diUb linkages.[18,19] Substrates were incubated with SdeA in the presence or absence of NAD^+ for 1 h before the addition of DUBs (Figure 3A).

The occurrence of ADP ribosylation or phosphoribosylation protected the substrate from DUB hydrolysis in all cases except for that of AMSH and UCHL1 (Figure 3B, Figure S2, and Table S2). Their activity may have been unhindered because of a lack of steric clashes between the -ADPR/-PR group of ubiquitin and the enzyme. Unexpectedly, for several of the DUBs examined, the presence of SdeA$^{181-1000}$ in the absence of NAD^+ caused an apparent increase in DUB activity compared to trials in which SdeA$^{519-1100}$ or no SdeA was included. It was expected that SdeA would compete with the DUB for diUb binding and therefore cause an inhibitory effect, but the opposite was observed. This may be due to the longer SdeA construct binding to the diUb substrate and holding it in a conformation that is more accessible for the DUB to act. This acceleration of activity was observed in the cases of SdeA^{DUB}, ChlaDUB1, and USP7, with this effect being most prominent in

DOI: 10.1021/acs.biochem.7b00664
Biochemistry XXXX, XXX, XXX−XXX

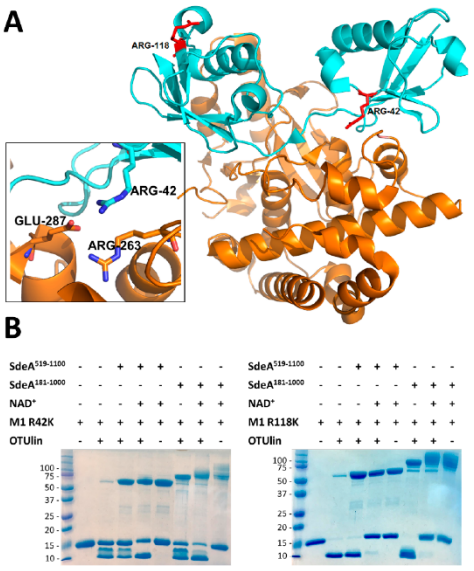


Figure 4. Mutational analysis. (A) Crystal structure of OTUlin bound to M1 diUb,[18] retrieved from the Protein Data Bank, with R42 and R118 of diUb colored red. The inset focuses on the two OTUlin residues that interact with R42. (B) SDS−PAGE of cleavage assays with indicated mutant M1 diUb substrates.

the case of ChlaDUB1. It is known that SdeA$^{181-1000}$ will ubiquitinate itself upon being incubated with Ub and NAD$^+$.[12] An upward shift in the mass of SdeA, indicative of autoubiquitination, continued to be observed when diUb instead of Ub was included. The addition of NAD$^+$ alone caused little to no inhibition of DUBs (Figure S3A). Decreasing the NAD$^+$ concentration to 100 μM resulted in a similar level of Rab1 ubiquitination as well as similar protective effects from DUB activity (Figure S3B,C).

According to structural data, R42 of M1-linked diUb is located at the DUB−substrate interface, whereas the R118 residue (proximal acceptor) is farther from the enzyme being exposed to solvent[20] (Figure 4A). Therefore, we expected the addition of an -ADPR or -PR group to R42 (distal acceptor) rather than R118 to cause a disruption to DUB binding and a corresponding loss of activity. To further investigate the basis for protection of diUb from hydrolysis, we generated two mutants of M1-linked diUb, with each mutant having either R42 or R118 changed to a lysine residue. Consistent with our hypothesis, the R118K mutant behaved like wild-type (WT) diUb in that the DUB activity of OTUlin was lost after modification with SdeA. The R42K mutant, in contrast, was cleaved by OTUlin with or without modification at a comparable level (Figure 4B). We noticed that the R42K mutant was a poorer substrate for OTUlin than the wild type or R118K mutant was, highlighting the importance of the distal R42 for binding (Figure S4). To compensate for the loss of activity upon Arg42 to Lys mutation, the reaction time was increased in the assay involving the R42K mutant. These data provide an example of the structural basis for the protection of Ub chains from DUB hydrolysis. Via the addition of steric bulk in the form of -ADPR or -PR, the interface between the enzyme and Ub chain is disrupted. Selective ADP ribosylation and phosphoribosylation by SdeA of certain Ub moieties of chains may serve as a useful and simple enzymatic tool in future studies for probing proximal versus distal binding of uncharacterized ubiquitin-interacting enzymes such as DUBs.

The role of SdeA during *Legionella* infection has been the subject of intense study. A growing list of functions have been described, including ubiquitination of ER-associated Rab proteins,[12] ADP ribosylation, and phosphoribosylation of free ubiquitin possibly leading to subsequent loss of activity of E1/E2/E3 machinery,[13] and remodeling of host ER tubules.[15] The results presented here offer new insights into the role of SdeA in controlling the host's ubiquitin system. Our data suggest that along with free Ub, Ub that is part of a chain is not spared from modification by SdeA, with all Ub monomers serving as a target. SdeA-catalyzed modification of polyubiquitin chains can be extensive and most likely protects the chain from hydrolysis by most host DUBs as well as the DUB module of SdeA itself. The LCV surface is decorated with polyubiquitinated proteins during infection, a modification that seems to serve an important role in the life cycle of the bacterium.[14,21] In such a scenario, protection of polyubiquitin chains from host DUBs might prove advantageous to *L. pneumophila*. At least one other example of Ub chain modification has been observed, phosphorylation at Ser65 by PINK1.[22] Interestingly, this also impairs hydrolysis by DUBs. Future studies will elucidate the cellular effects of SdeA's activity, including the protection of Ub chains, as well as the structural basis of SdeA's Ub binding ability.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.7b00664.

Materials and Methods, Figures S1−S3, and Tables S1 and S2 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: cdas@purdue.edu.

ORCID

Chittaranjan Das: 0000-0002-0567-7753

Funding

Funding was provided by National Institutes of Health Grants R21AI105714 (Z.-Q.L.), R01AI127465 (Z.-Q.L. and C.D.), R01GM103401 (C.D.), and P41GM108569 (M.J.W.). Portions of this work were also supported by the PI4D Team of Science Incentive Grant (to Z.-Q.L. and C.D.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the Das lab for assistance in sample preparation as well as Andy Mesecar and Yogesh Kulathu for their gifts of USP7 and MIY1, respectively.

■ REFERENCES

(1) Komander, D., and Rape, M. (2012) *Annu. Rev. Biochem. 81*, 203−229.

DOI: 10.1021/acs.biochem.7b00664
Biochemistry XXXX, XXX, XXX−XXX

(2) Nijman, S. M., Luna-Vargas, M. P., Velds, A., Brummelkamp, T. R., Dirac, A. M., Sixma, T. K., and Bernards, R. (2005) *Cell 123*, 773−786.
(3) Ronau, J. A., Beckmann, J. F., and Hochstrasser, M. (2016) *Cell Res. 26*, 441−456.
(4) Abdul Rehman, S. A., Kristariyanto, Y. A., Choi, S.-Y., Nkosi, P. J., Weidlich, S., Labib, K., Hofmann, K., and Kulathu, Y. (2016) *Mol. Cell 63*, 146−155.
(5) Sato, Y., Yoshikawa, A., Yamagata, A., Mimura, H., Yamashita, M., Ookata, K., Nureki, O., Iwai, K., Komada, M., and Fukai, S. (2008) *Nature 455*, 358−362.
(6) Shrestha, R. K., Ronau, J. A., Davies, C. W., Guenette, R. G., Strieter, E. R., Paul, L. N., and Das, C. (2014) *Biochemistry 53*, 3199−3217.
(7) Bhavsar, A. P., Guttman, J. A., and Finlay, B. B. (2007) *Nature 449*, 827−834.
(8) Angot, A., Vergunst, A., Genin, S., and Peeters, N. (2007) *PLoS Pathog. 3*, No. e3.
(9) Newton, H. J., Ang, D. K. Y., van Driel, I. R., and Hartland, E. L. (2010) *Clin. Microbiol. Rev. 23*, 274−298.
(10) Luo, Z.-Q., and Isberg, R. R. (2004) *Proc. Natl. Acad. Sci. U. S. A. 101*, 841.
(11) Sheedlo, M. J., Qiu, J., Tan, Y., Paul, L. N., Luo, Z.-Q., and Das, C. (2015) *Proc. Natl. Acad. Sci. U. S. A. 112*, 15090−15095.
(12) Qiu, J., Sheedlo, M. J., Yu, K., Tan, Y., Nakayasu, E. S., Das, C., Liu, X., and Luo, Z.-Q. (2016) *Nature 533*, 120−124.
(13) Bhogaraju, S., Kalayil, S., Liu, Y., Bonn, F., Colby, T., Matic, I., and Dikic, I. (2016) *Cell 167*, 1636−1649.
(14) Ivanov, S. S., and Roy, C. R. (2009) *Cell. Microbiol. 11*, 261−278.
(15) Kotewicz, K. M., Ramabhadran, V., Sjoblom, N., Vogel, J. P., Haenssler, E., Zhang, M., Behringer, J., Scheck, R. A., and Isberg, R. R. (2017) *Cell Host Microbe 21*, 169−181.
(16) Misaghi, S., Balsara, Z. R., Catic, A., Spooner, E., Ploegh, H. L., and Starnbach, M. N. (2006) *Mol. Microbiol. 61*, 142−150.
(17) Qiu, J., Yu, K., Fei, X., Liu, Y., Nakayasu, E. S., Piehowski, P. D., Shaw, J. B., Puvar, K., Das, C., Liu, X., and Luo, Z.-Q. (2017) *Cell Res. 27*, 865−881.
(18) Das, C., Hoang, Q. Q., Kreinbring, C. A., Luchansky, S. J., Meray, R. K., Ray, S. S., Lansbury, P. T., Ringe, D., and Petsko, G. A. (2006) *Proc. Natl. Acad. Sci. U. S. A. 103*, 4675−4680.
(19) Boudreaux, D. A., Maiti, T. K., Davies, C. W., and Das, C. (2010) *Proc. Natl. Acad. Sci. U. S. A. 107*, 9117−9122.
(20) Keusekotten, K., Elliott, P. R., Glockner, L., Fiil, B. K., Damgaard, R. B., Kulathu, Y., Wauer, T., Hospenthal, M. K., Gyrd-Hansen, M., Krappmann, D., Hofmann, K., and Komander, D. (2013) *Cell 153*, 1312−1326.
(21) Isaac, D. T., and Isberg, R. (2014) *Future Microbiol. 9* (3), 343−59.
(22) Wauer, T., Swatek, K. N., Wagstaff, J. L., Gladkova, C., Pruneda, J. N., Michel, M. A., Gersch, M., Johnson, C. M., Freund, S. M., and Komander, D. (2015) *EMBO J. 34*, 307−325.

DOI: 10.1021/acs.biochem.7b00664
Biochemistry XXXX, XXX, XXX−XXX