# STRUCTURAL AND BIOCHEMICAL CHARACTERIZATION OF UNCONVENTIONAL BACTERIAL UBIQUITINATION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

David Jonathan Wasilko

December 2018

# STRUCTURAL AND BIOCHEMICAL CHARACTERIZATION OF UNCONVENTIONAL BACTERIAL UBIQUITINATION

David Jonathan Wasilko, Ph. D.

Cornell University 2018

*Legionella pneumophila*, a gram-negative intracellular pathogen, is the causative agent of a severe form of pneumonia termed Legionnaires' disease. Following engulfment by alveolar macrophages, *Legionella* secretes ~330 effector proteins into the host cell. These proteins facilitate the formation of the Legionella-Containing Vacuole (LCV), a replicative niche that allows *Legionella* to escape phagosomal degradation and subsequently permits intracellular bacterial replication. A growing number of effectors have been reported to alter host ubiquitin signaling, and further characterization of these effectors is crucial for understanding how host defenses are eluded.

Here, I present work that identifies the underlying mechanisms that govern E2 binding and ubiquitin transfer to a *Legionella*-encoded E3 ligase SidC. We have shown that the insertion (INS) domain of SidC is important for binding of ubiquitin-loaded E2. To define the key residues critical for this interaction, we solved the crystal structure of the SidC paralog SdcA in complex with an E2. We also crystallized and solved the structure of the catalytic domain of SidC in complex with its preferred E2 covalently conjugated to ubiquitin. This ternary structure reveals the molecular interactions that

lead to the transfer of ubiquitin from E2 to SidC. This study furthers our understanding of the mechanism by which this novel E3 ligase carries out its function.

The SidE family of effectors catalyzes a form of ubiquitination that functions independently of the canonical E1, E2, and E3 enzymatic cascade. However, the mechanism by which this reaction is carried out remained elusive. Here, we present the crystal structure of a fragment of the SidE family member SdeA that retains full ubiquitination activity. Our structure reveals that the ubiquitination catalytic module contains two distinct domains, a phosphodiesterase domain (PDE) and a mono-ADP-ribosyltransferase (mART) domain. Biochemical analysis shows that the conversion of ubiquitin to ADP-ribose-ubiquitin (ADPR-Ub) and the ligation of ADPR-Ub to substrates are two independent activities of SdeA. Furthermore, our crystal structures of a homologous PDE domain from the *L. pneumophila* effector SdeD in complex with both Ub and ADPR-Ub reveals an intriguing mechanism of how SdeA processes ADPR-Ub to release AMP and conjugates PR-Ub to serine residues of substrates.

## BIOGRAPHICAL SKETCH

Jon Wasilko was born in Clearfield, PA and was promptly moved to Rhode Island, eventually settling and growing up in Oakdale, CT. His first exposure to lab work came at about the age of three, when he routinely accompanied his father to his microbiology lab at Pfizer for after-hours and weekend experiments. As he grew up, he had a curiosity in several areas of science, which both of his parents nurtured. He completed a bachelor's degree in molecular biology at Penn State. His first forays into bench work were undertaken in Dr. Ming Tien's lab, where he immediately felt a sense of belonging that a career should be built on. After a brief meeting with his academic advisor (ostensibly to correct a scheduling mistake), in which the advisor offhandedly mentioned that he would have a good shot at getting into a Ph.D. program at Cornell, he could think of no other school that he wanted to attend for graduate work. Thus, he chose to pursue a Ph.D. at Cornell University in the field of Biochemistry, Molecular and Cell Biology. After completing his first two rotations he received an email from Dr. Yuxin Mao, who was looking for graduate students interested in biochemistry and crystallography. He completed his thesis work in Dr. Mao's lab, studying noncanonical ubiquitination enzymes found in the bacteria *Legionella pneumophila*. He plans to continue his training by making the jump to industry, looking for a different frame of reference through which to view his scientific work.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Yuxin Mao for his guidance and mentorship over the past four years. He has provided valuable instruction in how to approach and answer scientific questions. He has also greatly helped me to become better at communicating my work. I would also like to thank Dr. Fenghua Hu for her advice and critical feedback during our joint lab meetings. I would also like to thank my academic committee members, Dr. Holger Sondermann and Dr. William Brown. They have provided guidance and valuable input during my committee meetings.

In addition to my committee, I would like to thank Dr. Debbie Nero and Mr. Jim Blankenship for their invaluable mentorship while I was a teaching assistant for their respective courses. Under their guidance, I learned a great deal about how to teach a course and gained a new perspective on the challenges and rewards of undergraduate teaching.

I am grateful for the lab members and classmates that I have had during my Ph.D. They have provided advice and encouragement throughout the trials and tribulations of graduate school. I would also like to acknowledge the friends that I have made here at Cornell. They have contributed to my wonderful overall graduate school experience.

I would also like to thank my family (some of whom think that I've spent much of the past five years exclusively gardening and making wine) for their support and encouragement throughout my time in graduate school.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# LIST OF ABBREVIATIONS

ADPR-Ub: ADP-ribose-Ubiquitin

AMP: Adenosine Monophosphate

DUB: Deubiquitinase

HECT: Homologous to E6AP C-Terminus

INS: Insertion Domain

LCV: *Legionella*-containing vacuole

mART: mono-ADP-Ribosyltransferase

NAD: Nicotinamide Adenine Dinucleotide

PDE: Phosphodiesterase

PI(4)P: Phosphitidylinositol-4-phosphate

PR-Ub: Phosphoribose-Ubiquitin

P4C: PI(4)P-Binding Domain of SidC

RBR: RING-Between-Ring

RING: Really Interesting New Gene

SidC: Substrate of Icm/Dot Transporter C

SidE: Substrate of Icm/Dot Transporter E

SNL: SidC N-terminal Ub Ligase

Ub: Ubiquitin

WT: Wild-Type

CHAPTER 1

# EXPLOITATION OF THE UBIQUITIN AND PHOSPHOINOSITIDE PATHWAYS BY THE *LEGIONELLA PNEUMOPHILA* EFFECTOR, SIDC

*1.1-1.4 excerpted from:*

Wasilko, D. J., & Mao, Y. (2016). Exploiting the ubiquitin and phosphoinositide pathways by the Legionella pneumophila effector, SidC. *Current genetics*, *62*(1), 105-108.

## 1.1 The ubiquitination pathway

Ubiquitination is a eukaryotic posttranslational modification important for a plethora of cellular pathways, including protein homeostasis, cell signaling, and membrane trafficking [1, 2]. This process is carried out through the action of three classes of enzymes: E1 activating enzymes, E2 conjugating enzymes, and E3 ubiquitin ligases [3]. Ubiquitin is activated and linked to the catalytic cysteine of E1 through a thioester linkage with the C-terminal glycine of ubiquitin and is then transferred to an E2 conjugating enzyme. The subsequent ubiquitination of a cellular substrate is mediated by the action of an E3 ubiquitin ligase, which catalyzes the formation of an isopeptide bond between the C-terminal glycine of ubiquitin and the ε-amino group of a substrate lysine. There are three main classes of E3 ubiquitin ligases: The RING (Really Interesting New Gene) family indirectly facilitates the transfer, the HECT (Homologous to E6AP C-Terminus) family utilizes a catalytic cysteine to form an E3-ubiquitin intermediate before ubiquitinating a substrate, and the RBR (Ring-Between-Ring) family uses a hybrid mechanism to carry out ligation [4-6]. While ubiquitination is well-established as a strictly eukaryotic process, many intracellular pathogens have adopted Bacterially-encoded E3 ubiquitin Ligases (BELs), likely through horizontal gene transfer [7, 8]. Although the disruption of the host ubiquitination pathway is

thought to play a crucial role in these infections, little is known about how disturbances in this pathway lead to a successful infection [9].

### 1.2 Phosphoinositide signaling

Phosphoinositides (PIs) are an essential class of membrane lipids found on the cytosolic surface of membranes. PIs can be reversibly phosphorylated at the 3,4, and/or 5 positions of the inositol ring to generate seven distinct PI species. Each of these lipids has a unique distribution within the cell and contributes to the membrane identity of cellular organelles. Besides this "zip-code"-like function, PIs also play important roles in an ever-growing number of cellular processes, including membrane trafficking, signal transduction, and cell migration [10, 11]. Accumulating data has shown that PIs are a target of intracellular bacterial pathogens, which use a variety of mechanisms to subvert host PI metabolism [12-14]. Many intracellular pathogens secrete effector proteins that directly modify the PI levels in the host, alter the activity of host PI-metabolizing enzymes, or bind to PIs for their proper intracellular targeting and function [15-17].

### 1.3 *Legionella pneumophila*

A prime example of the exploitation of both the host PI and ubiquitin systems is *Legionella pneumophila,* a gram-negative intracellular pathogen whose natural host is freshwater amoeba. *Legionella* is an opportunistic pathogen of humans, causing a severe form of pneumonia termed Legionnaires' disease [18]. *Legionella* uses an intracellular multiplication/defective organelle trafficking (Dot/Icm) type IV secretion system to secrete approximately 300 effector proteins into the host cytosol to hijack many host cellular processes, leading to the formation of a replicative niche for the bacteria, termed the *Legionella*-containing vacuole (LCV) [19, 20]. While the

majority of the effector proteins remain to be characterized, one of these effectors, SidC (and its paralog SdcA), has been recently found to have multiple functions that serve to connect both the ubiquitin and PI pathways.

### 1.4 SidC

SidC was shown to associate with the LCV though a C-terminal PI(4)P-binding region and was suggested to play a role in the enrichment of both ER-derived vesicles and polyubiquitinated species on the LCV [13, 21-23]. We and others determined the structure of the N-terminal domain of SidC, which contained a novel fold with an unknown function [23-25]. With careful sequence and structural analysis, we found a cluster of residues on the surface of the protein that was conserved across all SidC homologs. We observed that this cluster contained three highly conserved residues (C46, H444, and D446) arranged in a manner reminiscent of the classical Cys-His-Asp catalytic triad commonly found in cysteine proteases and deubiquitinases. However, after extensive trials we failed to detect any activity related to these hydrolytic enzymes. Surprisingly, a significant change in the ubiquitination pattern of the total cell lysate was detected in an experiment in which SidC was co-expressed with HA tagged ubiquitin in HEK293T cells. This observation led us to hypothesize that SidC may function as an E3 ubiquitin ligase, which was verified by *in vitro* ubiquitination assays. In addition, we found that SidC preferentially catalyzed K11 and K33-linked polyubiquitin chains. Thus, we named this novel N-terminal domain the SNL domain (SidC N-terminal ubiquitin Ligase).

### 1.5 SidE family-mediated phosphoribosyl-ubiquitination

Recently, a novel form of ubiquitination has been discovered in *Legionella*, carried out by the SidE family of proteins [26]. Termed phosphoribosyl-ubiquitination (PR-ubiquitination), this

pathway uses $NAD^+$ as a cofactor to catalyze the addition of a phosphoribose on Arg 42 of ubiquitin which is then linked to a substrate serine residue. The C-terminal glycine used for ligation in conventional ubiquitin is not essential for this reaction, and PR-ubiquitin is incapable of being used by canonical enzymes, as the modification blocks a hydrophobic patch centered around Ile44 that is critical for traditional ubiquitination [27]. In place of the canonical E1, E2, and E3 enzymes, the SidE family catalyzes this ubiquitination using two domains: a mono-ADP ribosyltransferase (mART) and a phosphodiesterase (PDE). Additionally, the SidE family contains a canonical deubiquitinase (DUB) at the N-terminus which is not essential for the PR-ubiquitination activity, and a coiled-coil domain at the C-terminus which is important for proper folding of the mART domain and is potentially important for translocation into host cells [28, 29]. The targets of PR-ubiquitination include ER-associated Rab proteins, including Rab1, Rab6A, Rab30, and Rab33b, although a complete list of substrates and the function of this ubiquitination pathway are both lacking [26]. Interestingly, ubiquitin chains that have been modified by the SidE family cannot be hydrolyzed by canonical DUBs [30].

The SidE family consists of four large members (SidE, SdeA, SdeB, and SdeC) and five smaller members, identified based on the presence of a homologous PDE domain. Much of the work done has used the SidE family member SdeA as a model for the family, but there could be slight variations between the four large family members. As of yet, there has been no reported role for the smaller SidE family members. Also, there have been reports of regulators of SidE family activity, however more studies are needed to gain a better perspective on the control of PR-ubiquitination [31, 32]. Many of the details for this family of proteins, including a role for PR-ubiquitination in a *Legionella* infection, have yet to be determined.

CHAPTER 2

# PI(4)P-MEDIATED REGULATION AND E2 RECOGNITION BY SIDC

*Portions excerpted from:*

Luo, X., Wasilko, D. J., Liu, Y., Sun, J., Wu, X., Luo, Z. Q., & Mao, Y. (2015). Structure of the Legionella virulence factor, SidC reveals a unique PI (4) P-specific binding domain essential for its targeting to the bacterial phagosome. *PLoS pathogens*, *11*(6), e1004965.

## 2.1 A structural basis for regulation

Further structural studies of SidC led us to reveal the nearly full-length structure of SidC [33]. We found that the C-terminal PI(4)P-binding P4C domain assumed a four antiparallel α-helical bundle structure. A positively charged pocket formed at one end of this bundle was confirmed as the specific PI(4)P binding site. Intriguingly, the P4C domain masked the ubiquitin ligase active site though a hydrophobic interaction with a patch of residues near the catalytic cysteine (Figure 2.2A and B). This "closed" conformation not only rendered SidC a lower ubiquitin ligase activity compared to the isolated SNL domain, but also a lower affinity for PI(4)P binding compared to the P4C domain alone. We found that SidC presented an increased ligase activity upon PI(4)P binding, presumably due to the "opening" of the catalytic site induced by PI(4)P. This regulatory mechanism seemed to be conserved in all SidC orthologs except that of *Legionella longbeachae*. In that species, the leucine residue that mediated the interaction between the SNL and P4C domains was altered to a glutamic acid. This substitution likely disrupted the hydrophobic interaction and gave *L. longbeachae* SidC a preference for the "open" conformation. In agreement with this observation, *L. longbeachae* SidC was shown to have ~3 fold increase in affinity for PI(4)P binding when compared to *L. pneumophila* SidC [34].

## 2.2 E2 recognition by the SidC INS domain

A comparison of the two SidC structures showed that a major conformational change occurs at the two flexible linkers between the SNL and INS domains. The SNL domain is rotated along a hinge (residue S224 and S328) by about 30° in SidC871 compared to the orientation in SidC542 (Figure 2.1A). This hinge bending motion is reminiscent of a similar displacement between the N- and the C-lobes of the HECT family ubiquitin ligases [35, 36]. These observations imply that the INS domain may interact with E2s in the ubiquitin conjugating reaction. To test this hypothesis, we used size exclusion chromatography to analyze the interaction between E2 and SidC. The catalytically inactive C85S mutant UbcH7 was stably charged with ubiquitin (UbcH7~Ub) and incubated with SidC542. The protein mixture was then separated by size exclusion chromatography and elution fractions were analyzed by SDS-PAGE. UbcH7~Ub, but not UbcH7 alone was able to form a stable complex with SidC542 as indicated by the co-fractionation of SidC542 with UbcH7~Ub, but not with UbcH7. In contrast, although the SidC542 ΔINS mutant behaved well as a mono-dispersed protein [23], it failed to interact with Ub-UbcH7 (Figure 2.1E). This suggests that the INS domain mediates the binding with E2s. Furthermore, like the E2~Ub-HECT E3 ternary complex [37], the covalently linked ubiquitin likely interacts with an area near the E3 catalytic site, thus stabilizing the E2~Ub-SidC ternary complex. Consistent with these results, the SidC542 ΔINS mutant, which failed to bind E2, is completely inactive in ubiquitination assays (Figure 2.2).

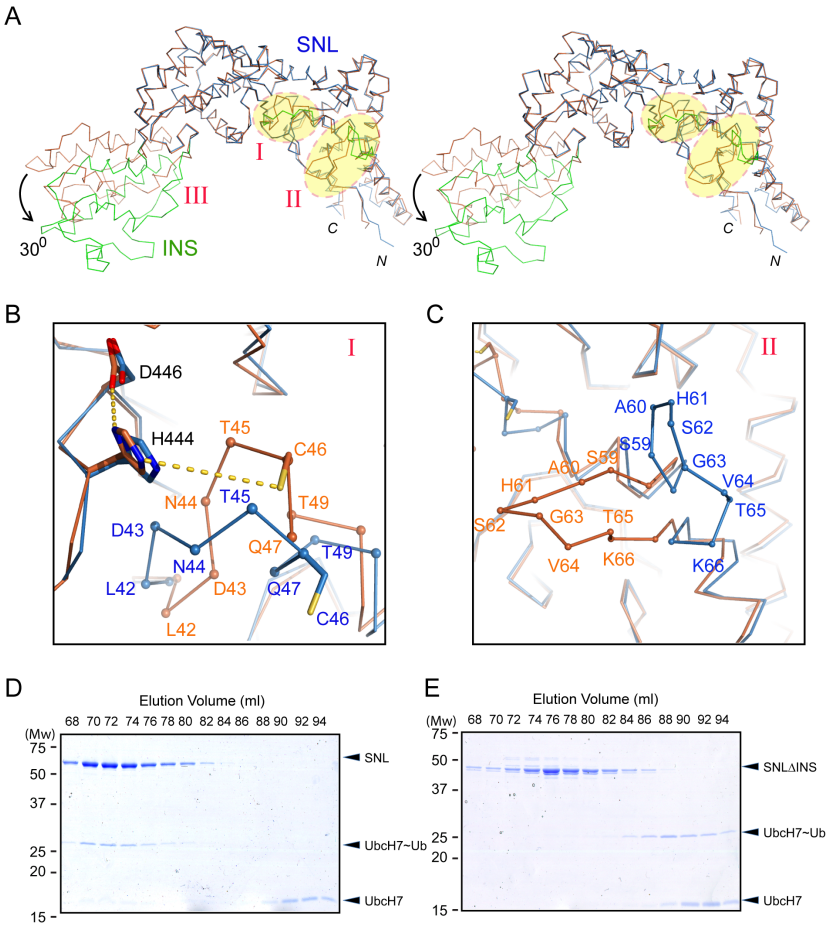


**Figure 2.1: Conformational dynamics of the SNL and INS domains.** (A) Stereo view of the Cα trace of the SNL (blue) and INS (green) domains of SidC in overlay with our previously reported SNL-INS domain structure (light brown; PDB ID: 4TRH). The three areas that have major conformational changes are labeled with I, II, and III, respectively. The INS domain in SidC871 is bent by about $30^0$ relative to the SNL domain. (B) Zoomed in view of the conformational changes at the catalytic site. C46 is shifted away from H444 and D446 in SidC871 (blue) compared with the SidC542 structure (brown). (C) Zoomed in view of the conformational change at the non-conserved loop (residue 59-66). (D) The SNL domain (SidC542) forms stable complex with UbcH7~Ub. SDS-gel of the size exclusion chromatography fractions from the sample containing SNL domain with ubiquitin-charged UbcH7. (E) The INS domain is involved in the binding of the SNL domain with UbcH7~Ub. SDS-gel of the size exclusion chromatography fractions from the sample containing the SNLΔINS domain with ubiquitin-charged UbcH7. UbcH7~Ub did not co-migrate with the SNLΔINS domain.

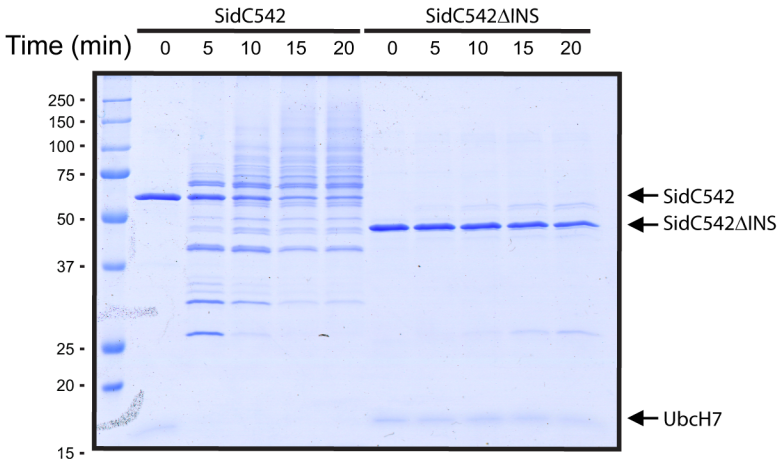


**Figure 2.2: In vitro ubiquitination assays with SidC542 and SidC542ΔINS.** Polyubiquitinated species accumulated in reactions with SidC542, however, polyubiquitinated protein bands are absent in reactions with Sidc542ΔINS.

### 2.3 The interface between the P4C and SNL domains

Our crystal structure revealed that the P4C packs against the SNL domain and blocks the accessibility of the E3 catalytic site. The interaction between the P4C and SNL domains is mainly mediated by extensive hydrophobic interactions between residues L629 and I633 of the P4C domain and a hydrophobic patch consisting of residues V50, I53, A60, I140, and P141 near the ligase catalytic site on the SNL domain Figure 2.3A and B). To characterize the nature of this intramolecular interaction, we created the L629R mutant with the speculation that the bulky and charged arginine residue would destroy both surface complementarity and the hydrophobic nature at the interface. Thus, this mutation would displace the P4C domain and keep SidC in an open conformation. Indeed, SidC743 C46A/L629R, but not SidC743 C46A formed a stable complex with UbcH7~Ub as shown by size exclusion chromatography analysis (Figure 2.3C and D). The formation of the ternary complex is likely due to the open conformation caused by this mutation, which allows for the binding of the ubiquitin moiety to the catalytic site. These data suggest that the intramolecular interaction between the P4C and the SNL domain is not due to crystal packing.

Instead, it likely represents an intramolecular regulatory mechanism in which the displacement of the P4C domain from the catalytic site switches the enzyme into an open conformation to release the occlusion at the ligase active site of SidC.

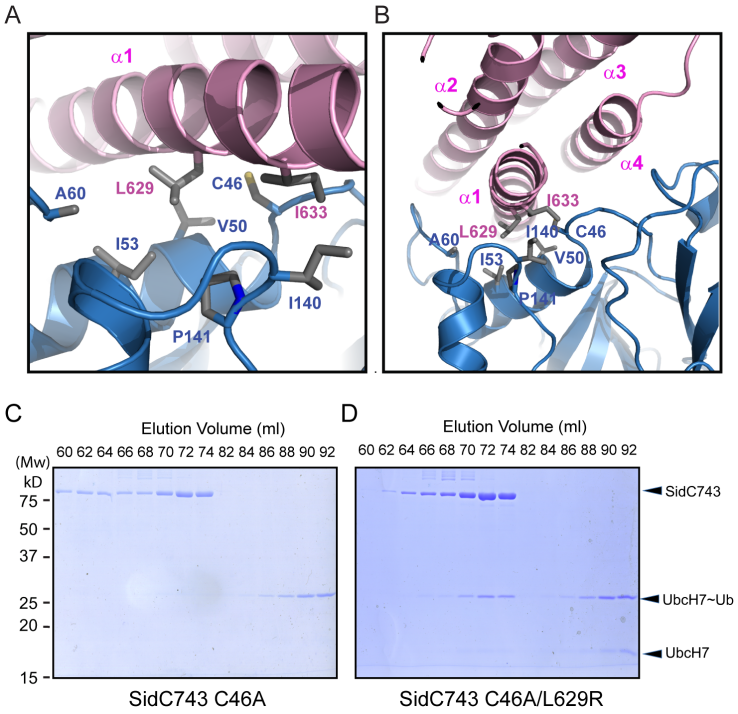


**Figure 2.3: The interface between the P4C domain and the SNL domain.** (A) and (B) Two orthogonal views of the interface between the P4C and SNL domains. Hydrophobic residues at the interface are shown in sticks. The P4C domain is colored in pink and the SNL domain in blue. (C) SidC743 C46A does not form a stable complex with ubiquitin-charged UbcH7 as indicated by SDS-PAGE gel analysis of fractions from size exclusion chromatography experiment. (D) SidC743 C46A/L629R forms stable complex with UbcH7~Ub as demonstrated by the co-migration of SidC743 C46A/L629R with UbcH7~Ub on the size exclusion column.

## 2.4 Summary

Our results indicate the discovery of the first effector protein that merges two independent cellular pathways, by uniting a unique ubiquitin E3 ligase activity and the recognition of a specific PI in a single peptide. These findings allow us to propose a model of SidC action (Figure 2.4). After *Legionella* infection, SidC is secreted in the closed state. Upon binding to PI(4)P on the LCV via the P4C domain, the ubiquitin ligase active site becomes accessible, leading to an increase in

E3 ubiquitin ligase activity. In this model, the precise molecular mechanism for ubiquitin ligation remains to be explored. More importantly, the specific substrates that are ubiquitinated by SidC are still unknown. Future experiments are certainly needed to address how host ER vesicle trafficking is rerouted by the ubiquitin ligase activity of SidC.

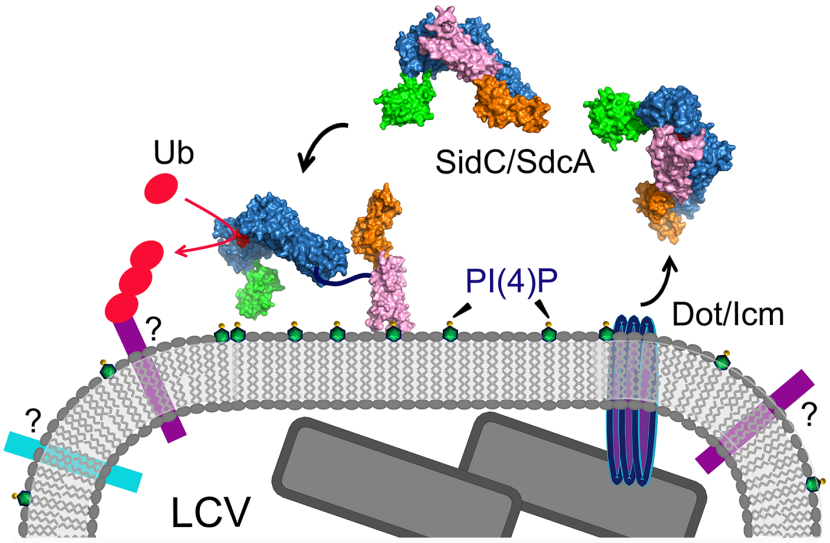


**Figure 2.4: A schematic model of SidC functions at the LCV surface.** SidC/SdcA is translocated into the cytosol through the Dot/Icm apparatus. SidC anchors to the LCV membrane through the binding to PI(4)P by its P4C domain. The binding of PI(4)P allows the P4C domain to move away from the SNL domain and exposes the ubiquitin ligase active site. The activated SidC/SdcA ubiquitinates unknown factors, presumably proteins residing on the LCV that determine the identity of the bacterial phagosome.

# CHAPTER 3

# INSIGHTS INTO THE UBIQUITIN TRANSFER CASCADE CATALYZED BY THE *LEGIONELLA* EFFECTOR SIDC

Wasilko, D. J., Huang, Q., & Mao, Y. (2018). Insights into the ubiquitin transfer cascade catalyzed by the Legionella effector SidC. *eLife*, *7*, e36154.

## 3.1 Abstract

The causative agent of Legionnaires' disease, *Legionella pneumophila*, delivers more than 330 virulent effectors to its host to establish an intracellular membrane-bound organelle called the *Legionella* containing vacuole. Among the army of *Legionella* effectors, SidC and its paralog SdcA have been identified as novel bacterial ubiquitin (Ub) E3 ligases. To gain insight into the molecular mechanism of SidC/SdcA as Ub ligases, we determined the crystal structures of a binary complex of the N-terminal catalytic SNL domain of SdcA with its cognate E2 UbcH5C and a ternary complex consisting of the SNL domain of SidC with the Ub-linked E2 UbcH7. These two structures reveal the molecular determinants governing the Ub transfer cascade catalyzed by SidC. Together, our data support a common mechanism in the Ub transfer cascade in which the donor Ub is immobilized with its C-terminal tail locked in an extended conformation, priming the donor Ub for catalysis.

## 3.2 Introduction

*Legionella pneumophila* is an intracellular opportunistic human pathogen that causes a severe form of pneumonia termed Legionnaires' disease [38, 39]. Following phagocytic uptake,

*Legionella* uses a Dot/Icm (Defective organelle trafficking/Intracellular multiplication) type IV secretion system to secrete over 330 effector proteins into the host cytosol [40, 41]. The concerted action of these effectors leads to the formation of the LCV (*Legionella*-containing vacuole), an organelle that prevents lysosome-mediated degradation of the invading bacteria while also serving as a replicative niche [42-44]. Although the biological functions of most *Legionella* effectors are as yet unknown, a number of effectors have been shown to modulate a rather diverse array of host cellular processes, including membrane trafficking, cellular signaling, lipid metabolism, and in particular, the host ubiquitination pathway [40, 45-47].

Ubiquitination is an essential eukaryotic posttranslational modification involved in myriad cellular activities. Ub is covalently attached to substrates in a three-enzyme cascade [48]. Ub is first activated by an E1 activating enzyme, then transferred to one of approximately 30 E2 conjugating enzymes, and finally ligated to a substrate by one of hundreds of E3 Ub ligases. The E3 ligases, which play a crucial role in the ubiquitination cascade, can be divided into three major classes: Really Interesting New Gene (RING) E3s [49], Ring-Between-Ring (RBR) E3s [50], and Homologous to E6AP C-terminus (HECT) E3s [51, 52]. Each of the three classes of E3s uses a distinct strategy to transfer Ub [6]. The RING E3s promote a "closed conformation" of the E2~Ub by locking the C-terminus of Ub into the active site groove on the E2 to facilitate direct Ub transfer from the E2~Ub to a substrate [53-55]. The HECT E3s catalyze two distinct reactions: a transthiolation reaction, which transfers Ub from E2~Ub to the E3 catalytic cysteine residue; and a subsequent aminolysis reaction, which transfers Ub from E3~Ub to a substrate lysine [51]. The third class, the RBR family of E3 ligases, shares features of both RING and HECT E3s and utilizes a hybrid mechanism for Ub transfer [56, 57]. Besides the three major classes of E3 ligases found in eukaryotes, a large number of bacterial pathogens encode a variety of effectors that mimic either

the RING or HECT type E3 ligases. However, some of these bacterially encoded E3 ligases have little sequence or structural homology to any other E3s [58, 59]. For example, the *Salmonella* protein SopA and the *Escherichia coli* protein NleL possess E3 ligase activity dependent on a catalytic cysteine and interact with E2s that contain a conserved phenylalanine residue yet show little sequence identity to known HECT E3s [35, 60]. How these unique E3 ligases catalyze Ub conjugation remains elusive.

The *Legionella* effector proteins SidC and its paralog SdcA has been shown to be required for the recruitment of ER proteins and ubiquitin signals to the LCV during an infection [61]. Our previous work has further shown that the SidC and SdcA are bona fide Ub E3 ligases [61, 62]. SidC contains an N-terminal Ub Ligase (SNL) domain, a P4C domain that specifically binds phosphatidylinositol-4-phosphate [62, 63] and a C-terminal portion of unknown function. Despite the lack of sequence or structural homology of SidC to any known eukaryotic E3s, the SNL domain of SidC uses a conserved cysteine as its catalytic residue, which is reminiscent of the eukaryotic HECT family of E3s. Our previous structural studies revealed that the SNL domain of SidC contains two lobes, a larger main lobe harboring the catalytic cysteine and a smaller E2-binding lobe connected to main lobe via two flexible linker peptides [61, 62]. However, the molecular mechanism of this unique family of bacterial Ub E3 ligases is largely unknown.

To gain insight into the molecular basis of the ubiquitination reaction catalyzed by SidC/SdcA, we determined crystal structures of two SidC/SdcA complexes: a binary complex of the SNL domain of SdcA with the E2 UbcH5C and a ternary complex of the SNL domain of SidC with the E2 UbcH7 covalently linked to ubiquitin (UbcH7~Ub). The complex structures reveal that upon binding of E2~Ub, the INS lobe undergoes a drastic swiveling to close a nearly 80 Å gap between the two catalytic cysteines of the E2 and E3, respectively. As a consequence, the

donor Ub is clamped between the INS and main lobes and makes an extensive network of contacts with the SNL domain. In particular, the C-terminal tail of the donor Ub adopts an extended β conformation and zip-pairs with a short β-strand upstream of the catalytic cysteine of the SNL domain. Mutations that interfere with the interactions between the donor Ub and the E3 ligase impede the ligase activity. Thus, our data not only reveal the catalytic mechanism of a unique family of bacterial E3 ligases, but also underpin a general concept of the ubiquitination reaction that efficient Ub transfer requires the donor Ub to be placed in a stationary position with the C-terminal tail of the donor Ub locked at the catalytic site.

## 3.3 Results

### 3.3.1 Overall structure of the SdcA-UbcH5C complex

As the first step in deciphering the mechanism of the SidC family of Ub E3 ligases, we sought to investigate how SidC recognizes E2. Our previous studies showed that SidC and SdcA preferentially carry out ubiquitination when either UbcH5C (UBE2D3) or UbcH7 (UBE2L3) serves as the E2 [61] . We screened all combinations of the SNL domain of SidC and SdcA with either UbcH5C or UbcH7 in crystallization trials and successfully obtained crystals formed by the SNL domain of SdcA (a.a. 1-538) and UbcH5C. The crystals diffracted to 3.0 Å and the structure of the binary complex was solved using molecular replacement using our previously determined SidC (a.a.1-542) structure as the search model. The final structural model was refined to an R-factor of 19.4 and a free R-factor of 26.1 (Table 3.1). The SNL domain of SdcA, which shares about 72% sequence homology to SidC, has a bilobed structure containing a large main lobe where the catalytic cysteine (C45) resides and a smaller inserted lobe (INS lobe) (Figure 3.1). We have

previously shown that the SNL domain of wild type SidC forms a complex with E2~Ub, but SidC lacking the INS lobe fails to do so [62]. In agreement with this observation, the binary complex structure reveals that the INS lobe indeed mediates direct binding with the E2, UbcH5C. The overall structure of the complex has an elongated arch-like shape. Strikingly, the catalytic cysteine of UbcH5C (C85) is about 80 Å away from the catalytic C45 of the SNL domain (Figure 3.1B-E). This structural feature indicates that a drastic conformational rearrangement is required for the INS lobe to bring the two catalytic cysteines within attacking distance for efficient Ub transfer between the E2 and the E3.

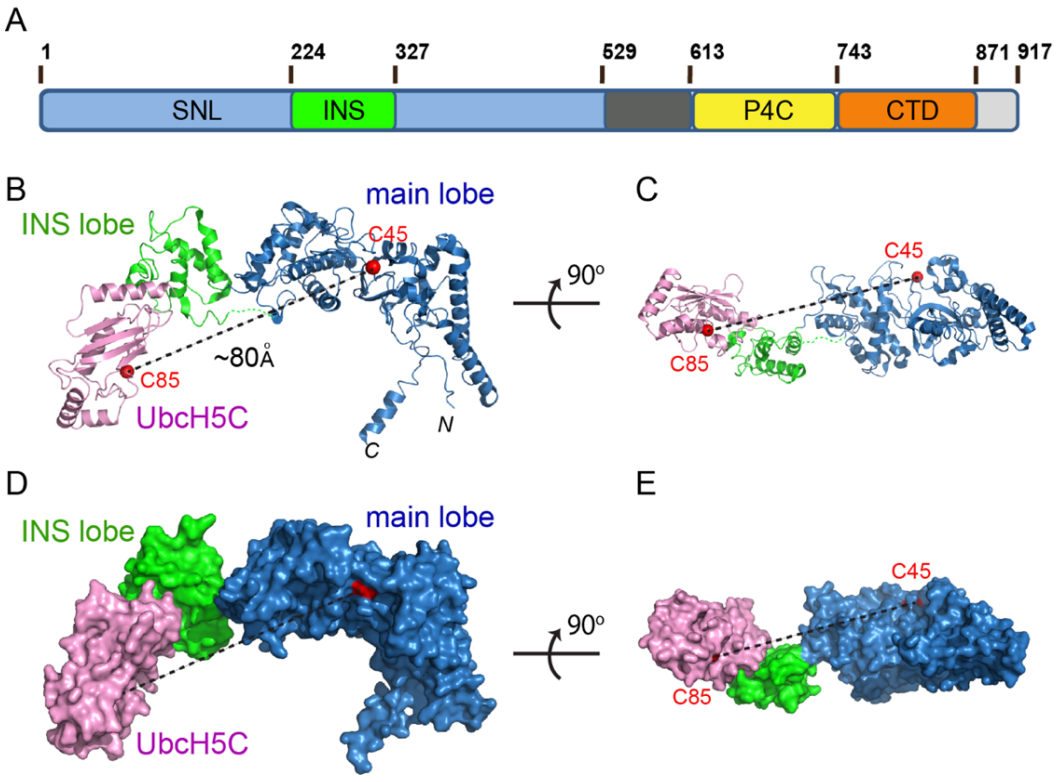


**Figure 3.1: Crystal structure of the SdcA-UbcH5C binary complex.** (A) Domain architecture of a member of the SidC family of Ub E3 ligase. (SNL: SidC N-terminal Ub E3 ligase; INS: insertion lobe; P4C: PI(4)P binding of SidC; CTD: C-terminal domain.) (B-C) Two orthogonal views of a ribbon diagram of the SNL domain of SdcA (blue: main lobe; green: INS lobe domain) in complex with UbcH5C (pink). The catalytic cysteines of both SdcA (C42) and UbcH5C are shown as red spheres. Note that the distance between the two cysteines is about ~80 Å. (D-E) Two orthogonal views of the SdcA-UbcH5C complex presented in surface.

**Table 3.1: SidC/SdcA data collection and refinement statistics.**

| | UbcH7~Ub-SidC SNL (PDB ID: 6CP2) | UbcH5C-SdcA SNL (PDB ID: 6CP0) |
|---|---|---|
| **Data collection** | | |
| Space group | $P6_522$ | $C222_1$ |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 101.522, 101.522, 352.302 | 135.550, 142.202, 118.333 |
| $\alpha, \beta, \gamma$ (°) | 90.0, 90.0, 120.0 | 90.0, 90.0, 90.0 |
| Resolution (Å)[a] | 50.0-2.90 (2.95-2.90) | 50.00-2.90 (2.95-2.90) |
| $R_{sym}$[b] (%) | 14.0 (80.2) | 12.5 (95.3) |
| $I/\sigma(I)$ | 30.3 (16.6) | 8.8 (2.7) |
| Completeness (%) | 99.9 (99.9) | 96.4 (91.5) |
| Redundancy | 13.8 (14.6) | 4.0 (3.7) |
| **Refinement** | | |
| Resolution (Å) | 2.9 | 3.0 |
| No. reflections | 27,387 | 23,102 |
| $R_{work}$ / $R_{free}$[c] (%) | 21.0/28.8 | 19.4/26.1 |
| R.m.s. deviations | | |
| Bond lengths (Å) | 0.0113 | 0.0106 |
| Bond angles (°) | 1.4734 | 1.4619 |
| Ramachandran Plot | | |
| Preferred (%) | 96.84 | 96.94 |
| Allowed (%) | 3.16 | 3.06 |
| Disallowed (%) | 0 | 0 |

[a] Values in parentheses are for highest-resolution shell.
[b] $R_{sym} = \Sigma_h\Sigma_i|I_I(h) - \langle I(h)|/\Sigma_h\Sigma_i I_I(h)$.
[c] $R_{crys} = \Sigma(||F_{obs}|-k|F_{cal}||)/\Sigma|F_{obs}|$. $R_{free}$ was calculated for 5% of reflections randomly excluded from the refinement.

### 3.3.2 Overall structure of the SidC-UbcH7~Ub complex

To gain insight into how this conformational rearrangement occurs during catalysis, we sought to determine the structure of the SNL domain in complex with a Ub-charged E2. To overcome the labile nature of a thioester bond, the catalytic cysteines of both UbcH5C and UbcH7 were mutated to lysine to form an E2~Ub complex linked by a stable isopeptide bond. In addition, to enhance the stability of the E3s, the catalytic cysteines of SidC (C46) and SdcA (C45) were mutated to alanine. We then performed crystallization screens for all four possible combinations of SidC or SdcA with UbcH5C~Ub or UbcH7~Ub, and obtained crystals of the SNL domain of SidC with UbcH7~Ub when mixed together in a 1:1.2 molar ratio. Although the crystals diffracted

poorly with the conventional flash-frozen cryoprotection method, the crystals diffracted to 2.9 Å when cryo-cooled under high-pressure [64]. The structure of the ternary complex was also solved via molecular replacement using the main lobe of SidC as the search model, and the final structure was refined to an R-factor of 21.0 and a free R-factor of 28.8 (Table 3.1). The SidC-UbcH7~Ub complex has a more globular, compact structure (Figure 3.2). In the ternary complex, the SNL domain of SidC maintains an arch-like shape, although with a shortened span compared to the SNL domain of SdcA in the binary complex. UbcH7 interacts with both the main and INS lobes of the SNL domain. The surface area of the INS lobe that mediates the binding with UbcH7 is nearly identical to that observed in the SdcA-UbcH5C binary complex. The Ub moiety is tightly clamped between the INS and main lobes, filling the space under the "arch". Remarkably, the distance between the two Cα atoms of the catalytic residues of the E2 (C86K) and the E3 (C46A) is reduced to ~9 Å (from ~80 Å in the binary complex) and the C-terminal carbonyl group of the donor Ub is within attacking distance from the modeled sulfhydryl group of C46A on SidC (Figure 3.2A-D). This dramatic structural rearrangement is mainly caused by a swinging motion made by the E2-binding INS lobe.

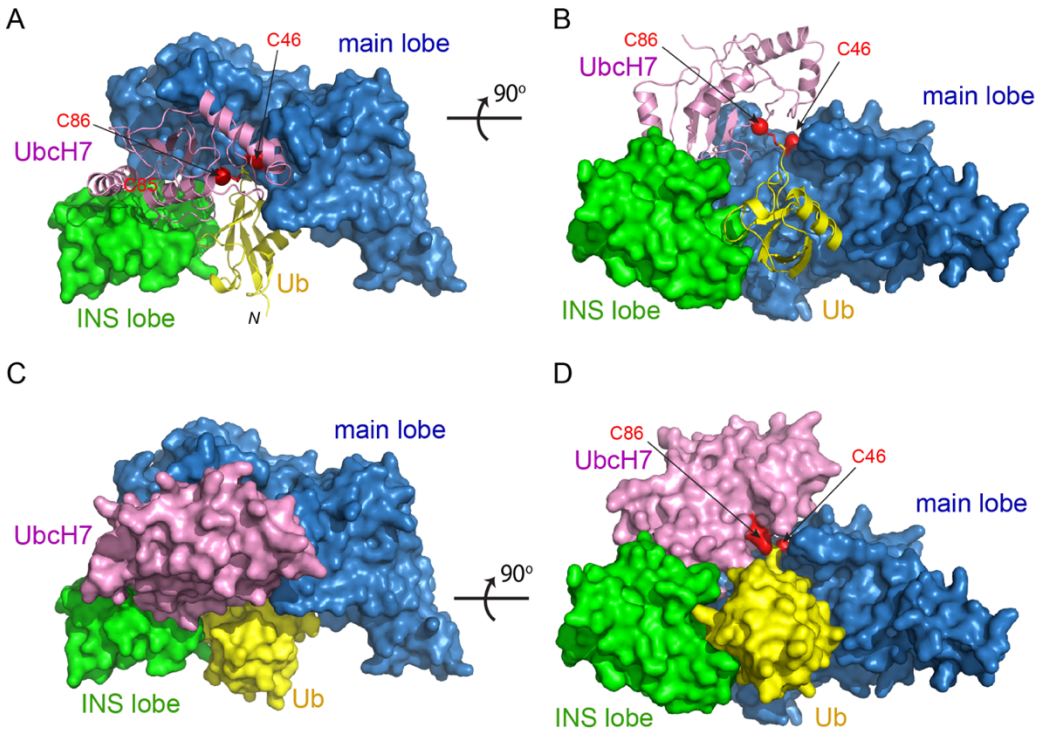


**Figure 3.2: Crystal structure of the SidC-UbcH7~Ub ternary complex. (**A-B) Two orthogonal views of the SidC-UbcH7~Ub complex structure. The SNL domain of SidC is shown in surface with the main lobe colored in blue and INS lobe in green. UbcH7 (pink) and the donor Ub (yellow) are shown in ribbon. The catalytic cysteines of both SidC (C46) and UbcH7 (C86) are shown in red. (C-D) Two orthogonal views of the SidC-UbcH7~Ub ternary complex in surface representation.

### 3.3.3 Conformational flexibility of the E2-binding INS lobe

The INS lobe is hinge-anchored to the main lobe through two flexible loops (Figure 3.1B). A structural comparison of the SNL domains from the two complexes and two previously reported SidC structures reveals that the INS lobe swings within a range of $90^0$ between the most extended conformation observed in the SdcA-UbcH5C complex and the most compact conformation found in the SidC-UbcH7~Ub complex, with the previously solved structures of SidC apo-enzymes falling in the middle (Figure 3.3A-B). Large conformational changes have commonly been observed in HECT-type E3 ligases. The classic HECT domain comprises a larger E2-binding N-lobe and a smaller C-lobe that harbors the catalytic cysteine. Structural studies have shown that a flexible hinge-like linker between the two lobes to allows the C-lobe cysteine to approach the

E2~Ub thioester for the transthiolation reaction, or move in close proximity to a substrate lysine for the aminolysis reaction in the second step of catalysis [36, 37, 60, 65] (Figure 3.3C). Similar conformational changes were also observed in bacterial HECT-like E3 ligases [35] (Figure 3.3D). In all known HECT or HECT-like E3s, the catalytic cysteine residing at the edge of the smaller C-lobe cycles between interacting with the E2~Ub thioester and the substrate lysine, owing to the rotational motion of the C-lobe. However, SidC differs in that the larger main lobe containing the catalytic cysteine likely harbors the substrate binding site and thus considered as the stationary lobe relative to the potential substrate binding site, while the E2-binding INS lobe recruits E2~Ub and delivers Ub to the catalytic cysteine through a large swiveling conformational shift. This variation suggests that SidC might use a different mechanism for Ub transfer (discussed below).

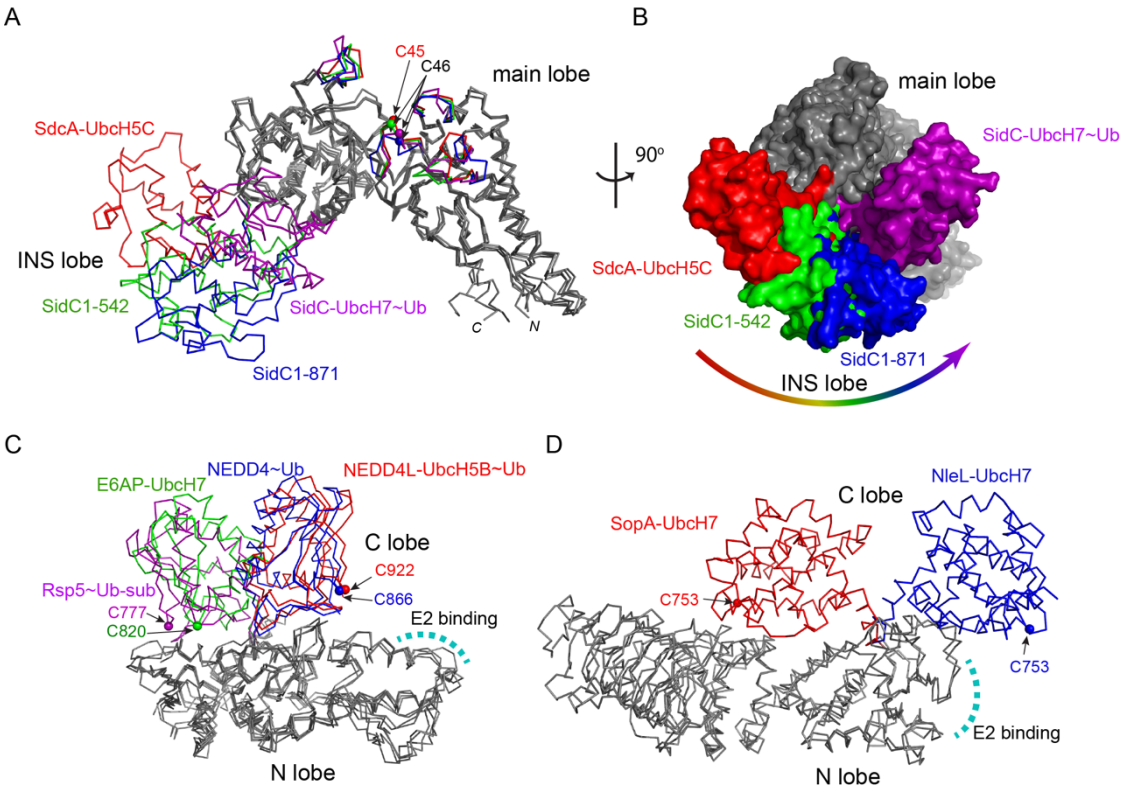


**Figure 3.3: Conformational flexibility of the E2-binding INS lobe.** (A) Overlay of structures of the SNL domains of SidC/SdcA. Four structures were superimposed based on the Cα atoms of the main lobe. The INS lobes adopt a wide range of conformations relative to the main lobe and are colored in red (SdcA-UbcH5C complex), green (SidC 1-542, PDB ID 4TRG), blue (SidC 1-871, PDB ID 4ZUZ), and purple (SidC-UbcH7~Ub), respectively. (B) An orthogonal view of the four superimposed structures shown in surface. The INS lobes are colored with the same scheme as in (A). (C) Structural overlay of HECT domains with their N-lobes superimposed. The C-lobes are colored red (NEDD4L-UbcH5B~Ub, PDB 3JW0), blue (NEDD4~Ub, PDB 4BBN), green (E6AP-UbcH7, PDB 1D5F), and purple (Rsp5~Ub-sub, PDB 4LCD), respectively. The E2 binding site is indicated by a curved dashed line. (D) Structural overlay of two HECT-like bacterial E3s. The E2-binding N-lobe is superimposed. The mobile C-lobe is colored in red (SopA-UbcH7, PDB 3SY2) and blue (NleL-UbcH7, PDB 3SQV), respectively.

### 3.3.4 Structural determinants for E2 recognition by SidC

A comparison of the SdcA-UbcH5C interface with that observed in the SidC-UbcH7~Ub complex reveals common structural determinants for E2 recognition. Both SdcA and SidC apply a similar surface area of about 950 Å$^2$ on the INS lobe to bind a region around the conserved E2 loop4 phenylalanine (UbcH5C F62 and UbcH7 F63) (Figure 3.4A), which is the pivotal residue mediating the binding with HECT E3s [66]. Further structural analysis reveals two key

determinants governing the binding between the INS lobes with E2s. First, the conserved E2 loop4 phenylalanine is placed into a hydrophobic pocket on the INS lobe in both SdcA and SidC (Figure 3.5). In SdcA, this hydrophobic pocket is lined with residues I285, A286, and Y273 (Figure 3.4B). Likewise, the hydrophobic pocket on the SidC INS lobe is formed by residues Y292, L297, M298, L299, and A316. Similar to HECT E3s, these residues are not conserved although the hydrophobic nature of the pocket is maintained. Second, the interaction between E2 and SdcA/SidC is also mediated by complementary electrostatic interactions. The interface surrounding the loop4 phenylalanine on both UbcH5C and UbcH7 has a positive electrostatic potential, while the corresponding interface on the INS lobe of both SdcA and SidC is negatively charged (Figure 3.6). In the SdcA-UbcH5C complex, residues K4 and K63 of UbcH5C form salt bridges with SdcA residues D287 and D305. Moreover, the R5 side chain of UbcH5C H-bonds with the main chain carbonyl group of A286 of SdcA (Figure 3.4B). Similarly, in the SidC-UbcH7~Ub complex, K9 of UbcH7 salt bridges with SidC D301 and R6 and K100 of UbcH7 form H-bonds with the carbonyl group of SidC L297 and L299, respectively (Figure 3.4C).

To validate the importance of the two types of interactions at the E2-E3 interface, we selectively mutagenized residues at the SidC-UbcH7 interface and analyzed the enzymatic activities of those mutants. Owing to the rapid transthiolation reaction catalyzed by SidC/SdcA, we used the ester-linked UbcH7~Ub (UbcH7 C86S was precharged with Ub by E1) in experiments to measure the single-turnover rate of UbcH7~Ub. The activity of the F63A mutant was substantially reduced compared to wild type (Figure 3.4D), suggesting that the hydrophobic loop4 phenylalanine is indispensable for E2 recognition by SidC/SdcA. Moreover, the R6A mutant, designed to disrupt electrostatic and H-bonding interactions with SidC, also showed a severe impairment of activity. On the other hand, mutations of the SNL domain of SidC also displayed

variable effects on the enzymatic activity. A disruption of the hydrophobic pocket by the L297D/L299D double mutant substantially impaired the activity while the single M298S mutant showed no obvious effect. Interestingly, the A316Y mutant, which was expected to maintain the hydrophobic nature of the pocket but with a reduced pocket size due to its bulky side chain, displayed a moderate effect (Figure 3.4E). The hydrophobic interaction appears to be conserved between SdcA and its cognate E2 UbcH5C. The Ub transfer activity was substantially reduced when the loop4 phenylalanine F62 in UbcH5C was mutated to alanine or when the hydrophobic pocket in the INS lobe of SdcA was disrupted by A286Y mutant (Figure 3.7). Together, our data support that both the hydrophobic and electrostatic/H-bond interactions at the E2-INS interface are crucial for E2 recognition and thus the catalytic function of SidC/SdcA.

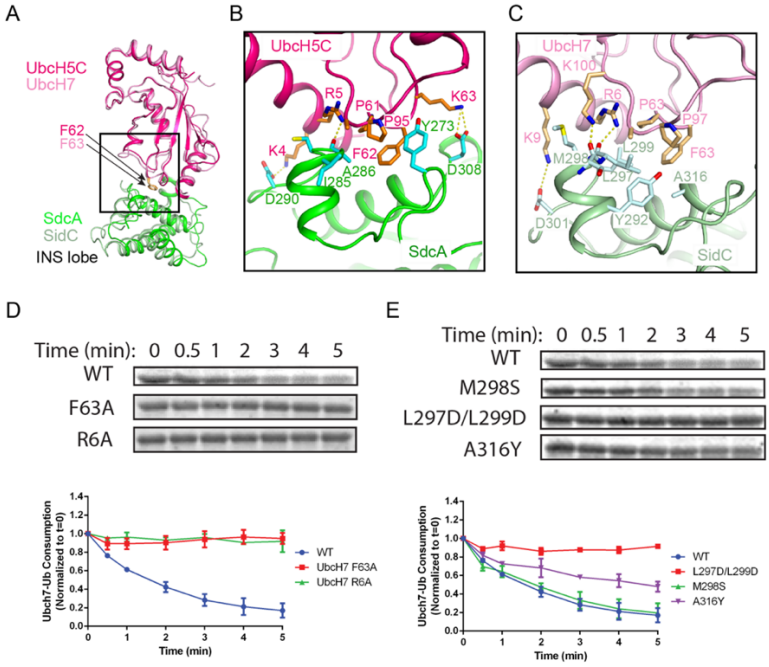


**Figure 3.4: E2 recognition by the INS lobe of SidC/SdcA.** (A) Structural superposition of the INS lobe of SdcA (green) and its bound E2 UbcH5C (pink) with the INS lobe of SidC (light green) and its bound UbcH7 (light pink). (B) Zoomed-in view of SdcA/UbcH5C interface. (C) Detailed interactions at the interface of SidC/UbcH7. (D) Single turnover activity assays of UbcH7 mutants at the SidC/UbcH7 interface. Top panel: representative SDS-gel of UbcH7~Ub complex remaining at the indicated time points. Bottom panel: Quantified intensity of UbcH7~Ub at the indicated time points. The error bar represents the standard deviation of three independent experiments. (E) Single turnover activity assays of SidC mutants at its E2 binding interface.

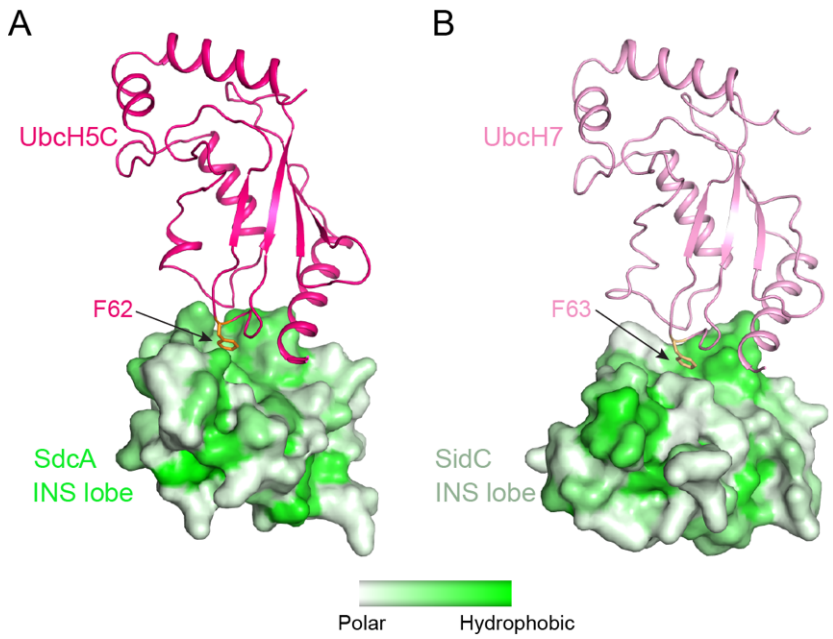


**Figure 3.5: Surface hydrophobicity analysis of the INS lobes.** (A) The interaction between UbcH5C (shown in pink ribbon) and the INS lobe (shown in surface) of SdcA. The INS lobe is colored based on the hydrophobicity of surface residues with the most polar residues in white and most hydrophobic residues in green. Note that the canonical E3-interacting F62 of UbcH5C is accommodated in a hydrophobic pocket on the INS domain. (B) The interaction between UbcH7 (shown in light pink ribbon) and the INS lobe (shown in surface) of SidC.

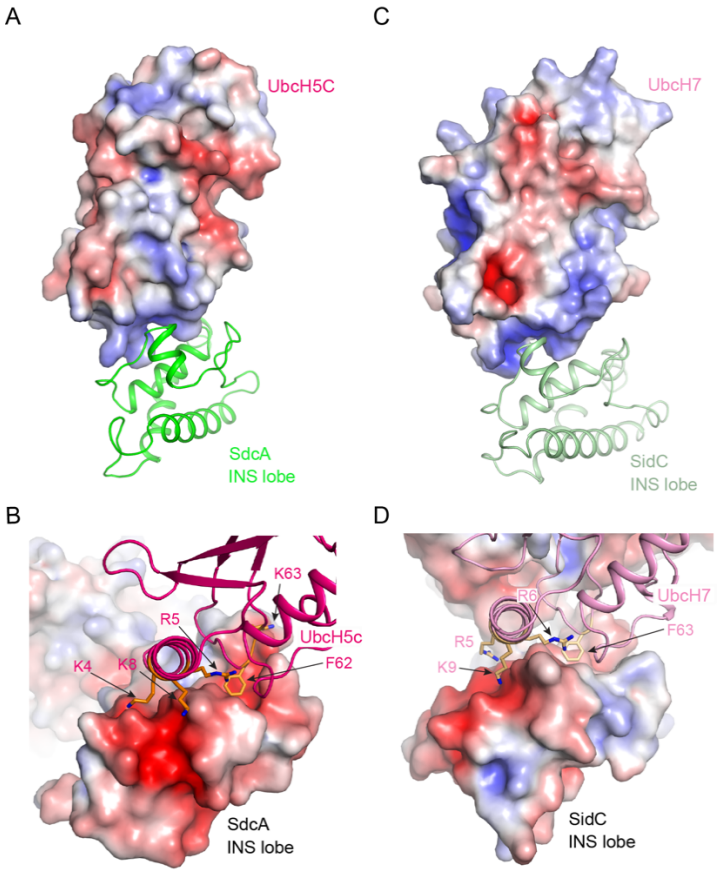


**Figure 3.6: Analysis of the surface electrostatic potential at the interface between the E2 and the INS lobe.** (A) The interaction between UbcH5C (shown in surface) and the INS lobe (shown in ribbon) of SdcA. Note that the surface area on UbcH5C mediated the binding with the SdcA INS lobe is positively charged. (B) Same interface as shown in (A). UbcH5C is shown in ribbon, while the INS lobe is shown in surface. Note that the E2-binding area is highly negatively charged. (C) and (D) Surface and ribbon representations of the interface between UbcH7 and the INS lobe of SidC. Note that similar to SdcA-UbcH5C, the interface on UbcH7 is positively charged (C) while the interface on the INS lobe of SidC is negatively charged (D). The surface is colored based on electrostatic potential with positively charged regions in blue (+5 kcal/electron) and negatively charged surface in red (-5 kcal/electron).

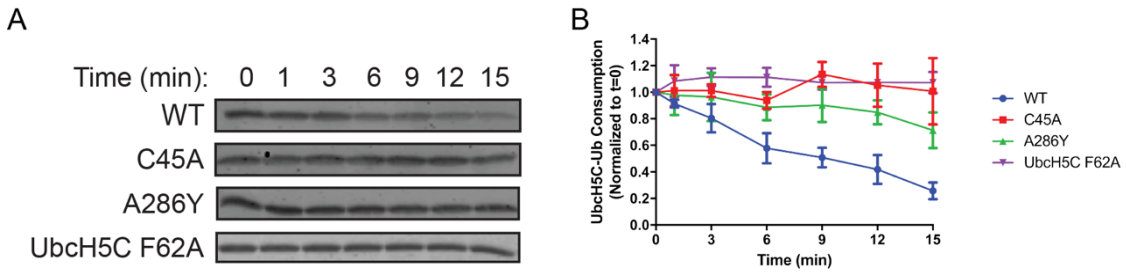


**Figure 3.7: Single turnover activity assays of SdcA and UbcH5C mutants at the SdcA/UbcH5C interface.** (A) Representative SDS-gel of ester bond-linked UbcH5C~Ub complex remaining at the indicated time points. (B) Quantified intensity of UbcH5C~Ub at the indicated time points. The error bar represents the standard deviation of three independent experiments.

### 3.3.5 The C-terminal tail of the donor Ub is locked at the E3 catalytic site

The SidC-UbcH7~Ub structure further reveals how the donor Ub is positioned on the SidC SNL domain and primed for transfer. The donor Ub, which is covalently attached to UbcH7, extends away from the E2 and is trapped between the INS and the main lobes of the SNL domain (Figure 3.8). Notably, the C-terminal tail of the donor Ub adopts a β-strand conformation and pairs in parallel with a short β-strand upstream of the catalytic cysteine. This β-sheet augmentation interaction was also observed in the donor Ub that was covalently attached to the HECT domain of Nedd4 [67]. The interaction between the donor Ub C-terminal tail and the SNL domain is further reinforced by sidechain-mediated H-bonds and salt bridges. The hydroxyl group of SidC T45 H-bonds with the amino group of Ub G75. Strikingly, R74 of the donor Ub forms a bidentate salt bridge with the highly conserved SidC D43 and a salt bridge with another conserved aspartic residue D51. Disruption of these interactions substantially hampers SidC activity as evidenced by the loss of activity of the D43R mutant, which disrupts the bidentate salt bridge, as well as the T45V mutant, which eliminates the hydrogen bond (Figure 3.8C). These results suggest that locking the C-terminal tail of the donor Ub at the E3 catalytic site is crucial for the reaction.

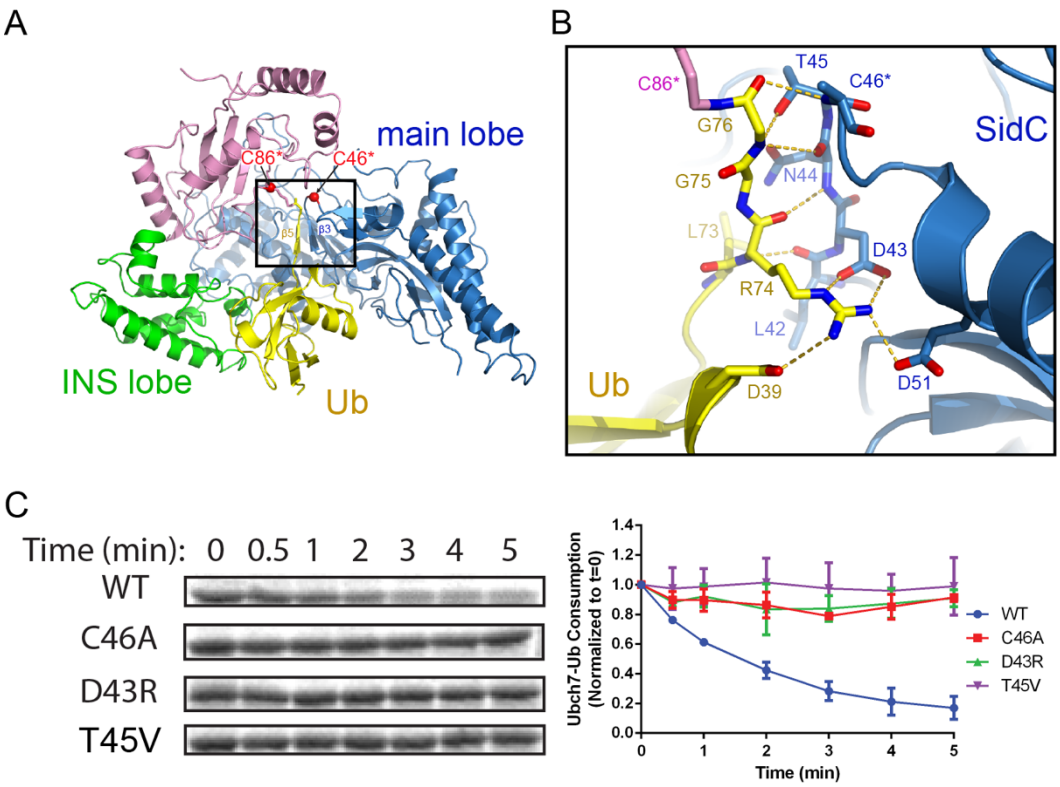


**Figure 3.8: The C-terminal tail of the donor Ub is locked on the main lobe.** (A) Overall view of the interaction of the donor Ub with the SNL domain of SidC. (B) Detailed interactions of the C-terminal tail of the donor Ub at the main lobe active site. (C) Single turnover activity assays of SidC mutants of residues that interacts with the Ub C-terminal tail.

### 3.3.6 The donor Ub is held stationary between the INS and main lobes during catalysis

Besides the locked C-terminal tail, the donor Ub also makes extensive contacts with both the INS lobe and main lobe of the SNL domain by burying ~2900 Å$^2$ of surface area (Figure 3.9A). Surface complementarity plays a significant role at the interface between the Ub and the INS lobe. The Ub loop containing K48 protrudes into a shallow groove on the INS domain with the hydrophobic I44 patch embedded at the interface. Surrounding the K48 loop, a network of H-bonds and electrostatic interactions are formed at the interface (Figure 3.9B). In particular, the carbonyl and ε-amino group of K48 H-bonds with the main chain amino and carbonyl groups on the INS lobe, respectively. Furthermore, K6 and R42 form electrostatic interactions with D271

and D319 of the INS lobe. The donor Ub also makes extensive contacts with the main lobe (Figure 3.9A). In addition to the C-terminal tail, the carboxyl end of the first α-helix of Ub forms hydrogen bonds and salt bridges with H504 and K505 on the main lobe (Figure 3.9B). To assess the importance of contacts between the SNL domain and the donor Ub, several residues on the SNL domain that make electrostatic interactions with Ub were mutated to residues carrying opposite charges. Both the D271R and the D319R mutants on the INS lobe largely impair SidC activity, as does the H504D/K505D double mutant on the main lobe. A sequence and structural comparison of SidC with SdcA indicates that residues involved in Ub binding are conserved in SdcA (Figure 3.10). Indeed, the Ub transfer activity was also substantially impaired when mutations of these conserved Ub-binding residues were introduced in SdcA (Figure 3.11). Taken together, these results indicate that the tight binding of the donor Ub by both SNL domain lobes is required for efficient catalysis and may also help prevent the reverse transthiolation reaction (handing the donor Ub back to the E2).

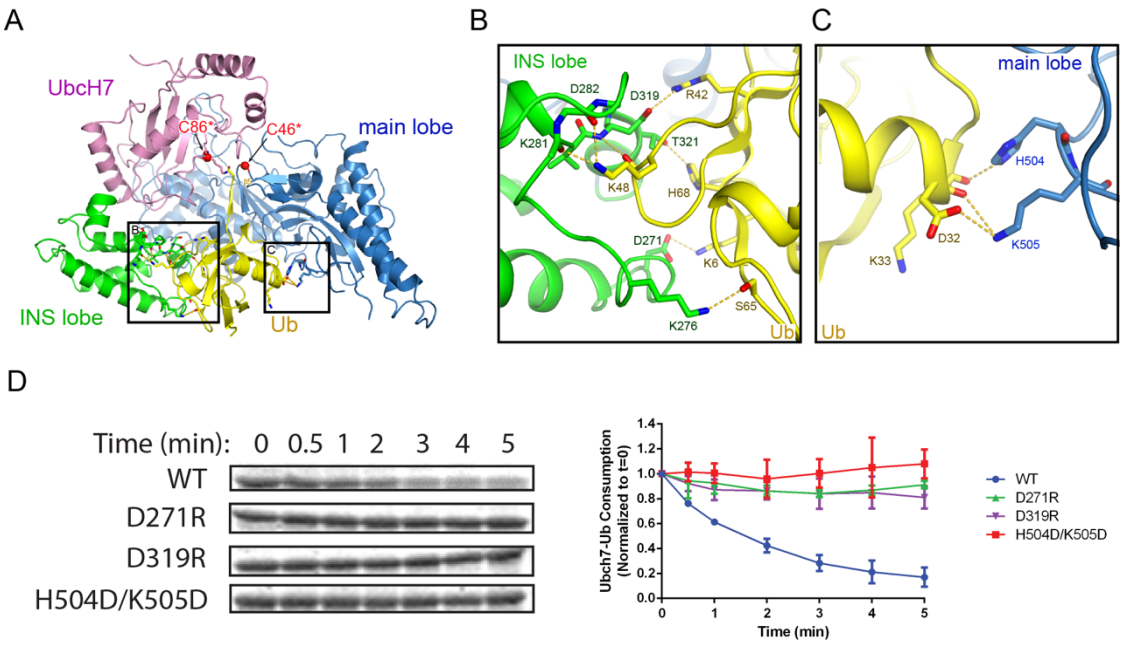


**Figure 3.9: The donor Ub is tightly bound between the INS and main lobes of the SNL domain.** (A) Overall view of the donor Ub bound on the SNL domain of SidC. The SNL domain is shown in surface with its main lobe colored in blue and the INS lobe in green. Residues on the main lobe that contact the donor Ub are colored in light blue and the contacting residues on the INS lobe are colored in light green. (B) Detailed network of interactions between the donor Ub and the INS lobe. H-bonds are illustrated with dashed lines. (C) Zoomed-in view of the interaction between the main lobe and the carboxyl end of the first α-helix of Ub. (D) Single turnover activity assays of SidC mutants of selected residues at the interface between the donor Ub and the SNL domain.

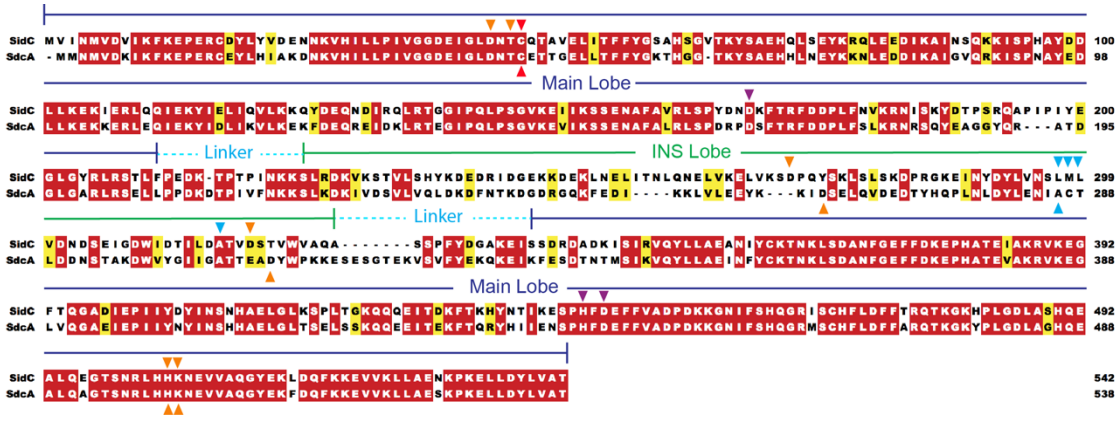


**Figure 3.10: Sequence alignment of the SNL domain region of SidC and SdcA.** Sequence alignment is calculated by Clustal Omega and colored by the Multiple Align Show online server. The INS and main lobes and the two linker peptides between the two lobes are labeled on the top of the sequences. Residues in either SidC or SdcA selected for mutagenesis and used in the enzymatic assays are labeled with red (catalytic residue), blue (at the E2 interface), gold (involved in Ub binding), or purple (forms the acidic patch near the catalytic site) triangles. Entrez database accession numbers are as follows: SidC, GI: 40792743; SdcA, GI: 52629829.

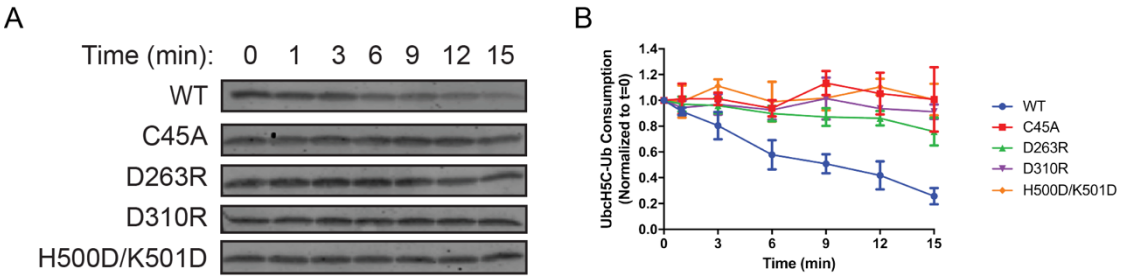


**Figure 3.11: Single turnover activity assays of SdcA mutants of selected residues at the interface between the donor Ub and the SNL domain.** (A) Representative SDS-gel of ester bond-linked UbcH5C~Ub complex remaining at the indicated time points. (B) Quantified intensity of UbcH5C~Ub at the indicated time points. The error bar represents the standard deviation of three independent experiments.

### 3.3.7 Acidic residues near the catalytic cysteine are crucial for E3 activity

A structure-based conservation analysis of the SNL domain with the ConSurf server [68] revealed a surface patch around the catalytic site enriched with conserved residues (Figure 3.12A). Besides the residues (N44, T45, and C46) comprising the catalytic motif, two invariable aspartic residues (D168, D446) and a histidine residue (H444) constitute a major part of the conserved patch (Figure 3.12B). Although in the SidC-UbcH7~Ub ternary complex the distance of these residues from the catalytic cysteine ranges between 10-14 Å, due to the flexible nature of the catalytic loop, these acidic residues can be in close proximity, about 5 Å away from the catalytic cysteine as observed in the SdcA-UbcH5C binary complex (Figure 3.13). It has frequently been observed that acidic residues are present in the vicinity the catalytic cysteine in HECT or HECT-like E3s. These acidic residues were proposed to either guide the substrate lysine into the active site and/or to deprotonate the ε-amino group of the approaching substrate lysine in Ub ligation [35, 60]. In particular, residue H887 in the RBR family ligase, HOIP, acts as a general base to activate the incoming α-amino group for linear Ub chain formation [69]. Similar acidic residues, such as D117 in UbcH5B and D127 in Ubc9, were also found in the vicinity of the catalytic cysteine in E2s [53, 54, 70, 71]. In the case of SidC, D168 may function as a guide or directly as a general base for the deprotonation of the attacking lysine. Although slightly further away, D446 forms a hydrogen bond with H444 and thus may indirectly activate the attacking lysine through the deprotonation of H444. Indeed, single amino acid substitutions of either of these three conserved residues markedly impaired the ligase activity of the SNL domain (Figure 3.12C and D). Together, our data reveal an important role for a cluster of invariable acidic residues at the active site in SidC catalyzed ubiquitination.

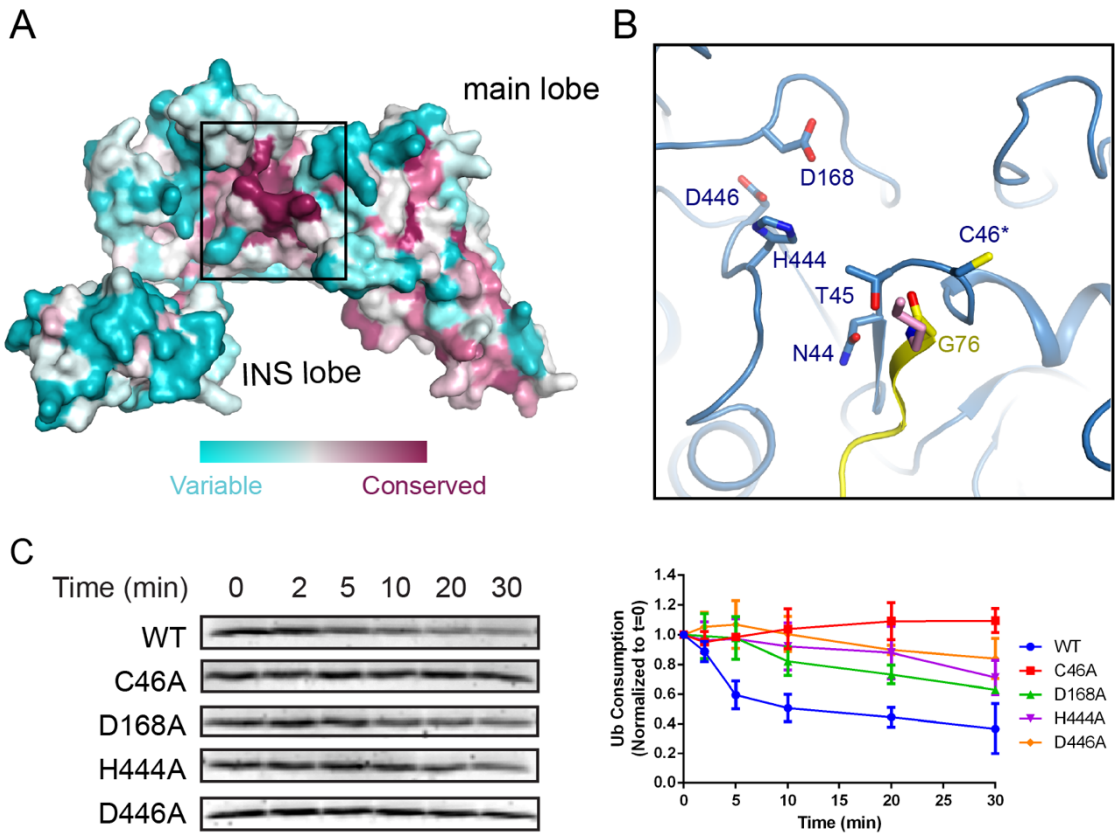


**Figure 3.12: Conserved acidic residues near the SNL domain catalytic site.** (A) Surface conservation analysis of the SNL domain. The conservation was calculated from all SidC homologous sequences from all *Legionella* species with available genomic sequences using the ConSurf server with the most conserved residues colored in purple and the least conserved residues in cyan. Note that the catalytic site of the SNL domain is concentrated with the most conserved residues. (B) A zoomed-in view of the most conserved residues at the catalytic site, including H444 and two acidic residues, D168 and D446. (C) Multiple turnover ubiquitination activity assays of SidC mutants of the conserved residues near the SidC active site monitored by the consumption of Ub. Left panel: representative SDS-gel of Ub remaining by the ubiquitination reaction at the indicated time points. Right panel: Quantified intensity of the remaining Ub at the indicated time points. The error bar represents the standard deviation of three independent experiments.

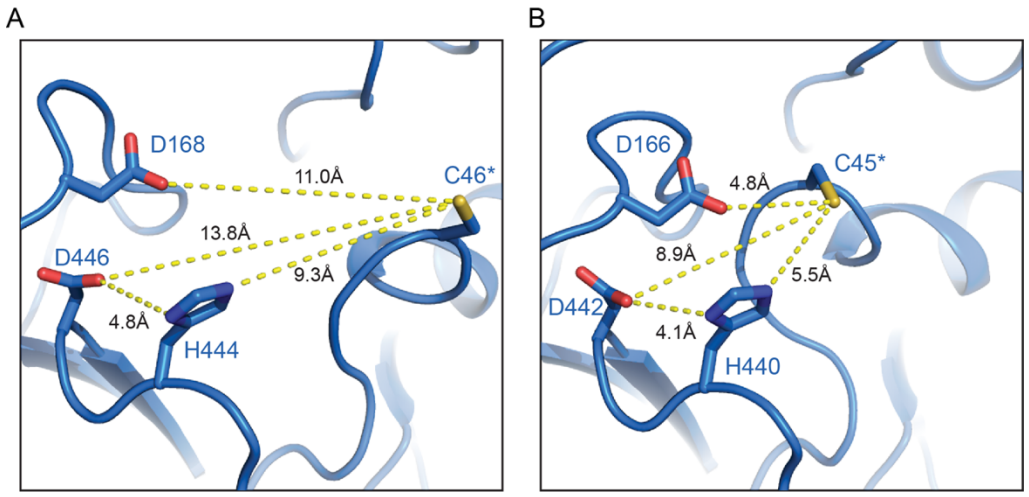


**Figure 3.13: Conformational flexibility of the SidC and SdcA active sites.** (A) Active site of SidC in the SidC-UbcH7~Ub ternary complex. (B) Active site of SdcA in SdcA-UbcH5C binary complex. The flexibility of the loop harboring the catalytic cysteine allows the conserved the acidic patch in a close proximity to the catalytic cysteine.

### 3.4 Discussion

In this study we present the crystal structures of two homologous *Legionella* Ub E3 ligases SidC/SdcA in complex with a human E2 enzyme alone and an E2 covalently conjugated to ubiquitin. These two structures capture SidC in the stage of E2 recognition, and the following stage at which the donor Ub is positioned for transfer from the E2 to the E3. These two catalytic intermediate snapshots allow us to postulate the mechanism of SidC-catalyzed ubiquitination. Although SidC has no sequence or structural homology to any other E3s, the catalytic SNL domain of SidC contains two lobes, a smaller INS lobe that mediates the binding with E2 and a larger main lobe that carries the catalytic cysteine. The E2-binding INS lobe undergoes a large swinging movement around two hinge peptides that connect the INS lobe to the main lobe. Through the swiveling motion of the INS lobe, E2~Ub is brought into the vicinity of the catalytic site on the main lobe. The donor Ub makes extensive contacts with both the INS and the main lobe and is tightly "trapped" on the SNL domain with its C-terminal tail engaged in a β-strand augmentation

interaction with a β-strand upstream of the catalytic cysteine. Since the catalytic cysteine is at a fixed a position at the center of the catalytic site, it is likely that the donor Ub will maintain the same tightly bound state after the transthiolation reaction, ready for ligation with a substrate lysine during the second step. Several acidic residues neighboring to the catalytic site may function to guide and/or activate attacking lysine residues to facilitate Ub ligation.

The structural organization of SidC is reminiscent of the N- and C-lobes in HECT E3s. However, notable differences exist between SidC and HECT E3s. First, in contrast to SidC, the E2-binding N-lobe of HECT E3s is relatively stationary in reference to the substrate binding site, while the catalytic cysteine-bearing C-lobe rotates around the linking peptide between the two lobes to relay the donor Ub from the E2 to a substrate. Second, the donor Ub on HECT E3s makes contacts nearly exclusively with the C-lobe while the donor Ub is clamped between the INS and main lobes and buries much larger surface areas on both the donor Ub and the SNL domain. The third prominent difference is that unlike the stationary donor Ub on SidC, the donor Ub loaded on the C-lobe of HECT E3s undergoes another drastic rotational movement to approach the substrate lysine after the E3~Ub intermediate is formed [60].

Despite these notable mechanistic differences, our studies of this unique family of bacterial E3s also provide exemplary insights into the general theme of ubiquitination reactions. It has been commonly observed that immobilizing the Ub thioester can enhance the efficiency of Ub transfer [53, 54, 67]. In the structure of the dimeric RING domain of RNF4 in complex with UbcH5A~Ub, the donor Ub is stabilized by interactions with the both RING protomers and the carboxyl terminus of the donor Ub is locked into an active site groove on the E2 [54]. In the structure that mimics the HECT~Ub thioester intermediate [67] and the structure of the HECT domain of Rsp5 crosslinked with a donor Ub and a peptide substrate [60], Both of the donor Ubs bind to the C-

lobe in a similar fashion with their C-terminal tails locked in an extended conformation. The structure of SidC-UbcH7~Ub provides an outstanding example of how the donor Ub is tightly locked on the E3 catalytic domain throughout the two-step ubiquitination reaction. Our results support a unifying concept that a reduction of the inherent conformational flexibility of the donor Ub thioester on both E2 and E3 is essential for efficient Ub transfer in the ubiquitination cascade.

## 3.5 Materials and Methods

**3.5.1 Cloning and mutagenesis.** SidC and SdcA were prepared as previously described [61]. In vitro site-directed mutagenesis was used to generate single or multiple point mutations in UbcH5C, UbcH7, SdcA538, and Sid542 using complimentary oligonucleotide primer pairs containing the requisite base substitutions. His-TEV-Ub was generated by PCR amplification of Ub with flanking BamH1 and Xho1 sites. The PCR product was digested and inserted into a pET-His-TEV vector containing an N-terminal 6x-His tag upstream of a TEV cleavage site.

**3.5.2 Protein expression and purification.** *E. coli* Rosetta strains harboring the appropriate expression plasmids were grown in Luria-Bertani medium supplemented with 34 μg/ml chloramphenicol and either 50 μg/ml kanamycin or 100 μg/ml ampicillin and grown to mid-log phase. Protein expression was induced overnight at 18°C with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Harvested cells were resuspended in buffer containing 150 mM NaCl, 50 mM Tris-HCl pH 8.0, and 1 mM PMSF, and lysed by sonication. The lysate was clarified by centrifugation at 31,000 x g for 45 min at 4°C, and incubated with cobalt resin (Gold-Bio) for 1.5 hrs at 4°C. Bound proteins were washed extensively with lysis buffer. UbcH5C, UbcH7,

SdcA538, and SidC542 were incubated with the SUMO-specific protease Ulp1, to release the desired protein and leave the His-SUMO tag conjugated to the resin. The eluted protein samples were further purified using size-exclusion chromatography (HiLoad Superdex 16/600 S200, GE Lifesciences), and assessed for purity via SDS-PAGE. The fractions corresponding to purified protein were collected, pooled, and concentrated in 150 mM NaCl, 20 mM Tris-HCl, pH 7.5. His-TEV-Ub was bound to cobalt resin, washed extensively, and eluted in 300 mM imidazole, 150 mM NaCl, and 50 mM Tris-HCl pH 8.0, before proceeding to size exclusion chromatography.

Human E1 was purified by reacting with ubiquitin-conjugated to Affi-Gel 10 (Bio-Rad) resin. Briefly, His-TEV-Ub K6 (containing only Lys6, with all other lysines mutated to arginine) was conjugated to the Affi-Gel resin via the primary amine of K6 in 150 mM NaCl, 50 mM HEPES pH 8.0, and then equilibrated in 50 mM Tris-HCl pH 8.0. After lysate clarification, 200 mM ATP, 6 mM creatine phosphate, and 50 U creatine phosphokinase was added to the hE1 supernatant, which was incubated with the Ub-Affi-Gel resin and nutated for 1 hr at 4°C. The mixture was poured into a column and after flow-through collection, the column was washed with 20 column volumes of 0.5 M KCl, 50 mM Tris-HCl pH 8.0. hE1 was eluted in 10 mM DTT, 50 mM Tris-HCl pH 8.0, and a buffer exchange into 150 mM NaCl, 50 mM Tris-HCl pH 8.0 was performed using a centrifugal filter, before proceeding to further size exclusion chromatography purification.

**3.5.3 In vitro ubiquitination assays.** For single turnover assays, UbcH7 C86S was used to generate stable, ester-linked UbcH7~Ub. 2 μM hE1, 34 μM UbcH7 C86S, and 203 μM ubiquitin were incubated in 50 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$, 1.8 U inorganic pyrophosphatase, 1.8 U creatine phosphokinase, and 21 mM creatine phosphate at 37°C for 8 h. After the reaction, a centrifugal buffer-exchange was performed (to remove $Mg^{+2}$ and ATP) and a final concentration

of 2 mM EDTA was added to chelate any remaining $Mg^{+2}$. The single turnover reaction was performed at room temperature in a total volume of 100 µl. In the reacrtion, 8 µM of ester-linked UbcH7~Ub was mixed with 0.9 µM SidC and 80 µM UbcH5C was added to the reaction, to act as a substrate proxy for SidC activity. The reaction mixture was sampled at the indicated time point and quenched with SDS loading buffer containing 5mM β-mercaptoethanol. The samples were separated by SDS-PAGE and stained with Coomassie blue. A LI-COR Odyssey scanner was used to image the Coomassie-stained gels, and the ImageStudio software package (LI-COR) was used to quantify the bands. The results were averaged from three independent experiments and plotted using Graphpad Prism (Graphpad).

Since there are no confirmed substrates that can be specifically ubiquitinated by SidC, in our single turnover assays, we used UbcH5C as a substrate proxy. In fact, in our in vitro ubiquitination reaction, E2s were observed to be ubiquitinated by SidC. The time-dependent accumulation of mono-ubiquitinated UbcH7 migrated at the same size as the ester-linked UbcH7~Ub on an SDS gel. Thus, the turnover of ester-linked UbcH7~Ub cannot be distinguished on an SDS gel. By adding an excessive amount of UbcH5C as a substrate proxy, it can prevent/compete SidC mediated ubiquitination of discharged UbcH7. Moreover, since the mono-ubiquitinated UbcH5C migrated at a different size compared to the ester-linked UbcH7~Ub, the consumption of ester-linked UbcH7~Ub can be visualized by SDS-PAGE in a time-dependent manner. The single turnover assay for SdcA was carried out similarly as described above. However, ester-linked UbcH5C~Ub was used in the assay and UbcH7 was used as the substrate proxy. The ubiquitin consumption assays were performed using 2 µM hE1, 3.4 µM UbcH7, 37.5 µM Ub, and 5.1 µM SidC542 (WT or catalytic site mutant) in Tris-HCl pH 8.0, 5 mM $MgCl_2$, 1.8 U inorganic pyrophosphatase (Sigma-Aldrich), 1.8 U creatine phosphokinase (Sigma-Aldrich),

and 21 mM creatine phosphate (Sigma-Aldrich) at 37°C for the indicated timepoints. Samples were analyzed by SDS-PAGE and quantification of the reaction was carried out similarly as described above.

**3.5.4 SdcA/UbcH5C and SidC/UbcH7~Ub complex formation.** To form the binary complex, SdcA and UbcH5C were mixed in a 1:1.2 molar ratio, incubated for 1 hr at 4°C, and concentrated to a final SdcA concentration of 10 mg/ml. To form the SidC-UbcH7~Ub ternary complex, UbcH7~Ub was first generated using an in vitro ubiquitination reaction. 2 μM hE1, 34 μM UbcH7 C86K, and 203 μM His-TEV-ubiquitin were incubated in 50 mM Tris-HCl pH 8.0, 5 mM $MgCl_2$, 1.8 U inorganic pyrophosphatase, 1.8 U creatine phosphokinase, and 21 mM creatine phosphate at 37°C for 24 h. The reaction was mixed with cobalt resin (Gold-Bio) equilibrated in 150 mM NaCl, 50 mM Tris-HCl pH 8.0 and rotated for 2 hrs at 4°C. After extensive washing with 150 mM NaCl, 50 mM Tris-HCl pH 8.0 (to remove unreacted UbcH7), bound proteins were eluted in 300 mM imidazole, 150 mM NaCl, 50 mM Tris-HCl pH 8.0. 0.5 mM EDTA and 1 mM DTT was added to the elution, and the His-TEV tag of ubiquitin was removed via TEV protease cleavage at room temperature for 4 hrs. Following cleavage, a buffer exchange into 150 mM NaCl, 50 mM Tris-HCl pH 8.0 was performed using a 10 kDa centrifugal filter (Amicon). The protein solution was incubated with equilibrated cobalt resin again for 2 hrs at 4°C, and UbcH7~Ub was collected in the flow-through fraction. Another centrifugal buffer exchange was performed into 20 mM NaCl, 20 mM Tris-HCl pH 7.0. The protein was passed through a HiTrap QP HP anion exchange column (GE Lifesciences) equilibrated in 20 mM NaCl, 20 mM Tris-HCl pH 7.0, and loaded onto a HiTrap SP HP cation exchange column (GE Lifesciences) in the same buffer. UbcH7~Ub was eluted using a continuous gradient of 500 mM NaCl, 20 mM Tris-HCl pH 7.0. Fractions

containing UbcH7~Ub were concentrated and further purified using size-exclusion chromatography (HiLoad Superdex 16/600 S200, GE Lifesciences). The SidC-UbcH7~Ub complex was formed by incubation of SidC with UbcH7~Ub in a 1:1.2 molar ratio for 1 hr at 4°C and the complex was concentrated to a final SidC concentration of 10 mg/ml.

**3.5.5 Crystallization, data collection, and processing.** Initial crystallization screens were carried out using a Crystal Phoenix liquid handling robot (Art Robbins Instruments) at room temperature. Crystals were grown at room temperature by hanging drop vapor diffusion by mixing 1.5 µl of protein with an equal volume of mother liquor. Small plate-like SdcA-UbcH5C crystals were formed after 2 days in 1.5 M ammonium sulfate, 3% glycerol, 0.1 M HEPES pH 7.5. The crystals were soaked in cryoprotectant solution containing the crystallization condition supplemented with 20% glycerol, and flash frozen in a stream of liquid nitrogen. Hexagonal rod-shaped SidC-UbcH7~Ub crystals formed after 4 days at room temperature in 16% PEG 3000, 0.1 M Tris pH 9.0. These crystals failed to diffract to high resolution under conventional cryo-cooling conditions. Therefore, a high-pressure cryo-cooling strategy was employed to pressurize the crystals at 380 MPa for 20 minutes using the capillary shielding method before frozen in liquid nitrogen [64, 72]. Diffraction data sets were collected at MacCHESS beamline A1 at the Cornell High Energy Synchrotron Source, and indexed, integrated, and scaled with HKL-2000 [73]. The SdcA-UbcH5C crystal belonged to the space group $C222_1$ with unit cell dimensions a = 135.55 Å, b = 142.20 Å, c = 118.33 Å, $\alpha = \beta = \gamma = 90.0°$. The SidC-UbcH7~Ub crystal belonged to the space group $P6_522$ with unit cell dimensions a = 101.52 Å, b = 101.52 Å, c = 352.302 Å, $\alpha = \beta = 90.0°$, $\gamma = 120.0°$. There is one complex molecule in the asymmetric unit of each of the crystals.

**3.5.6 Structure determination and refinement.** Both structures were solved by molecular replacement using the SNL domain main lobe of the SidC542 structure (PDB ID: 4RTH) as the search model with the AMoRe program [74] of the CCP4 suite [75]. Iterative cycles of model building and refinement were performed using COOT [76] and refmac5 [77] of the CCP4 suite.

**3.5.7 Computational analysis and graphic presentation of protein structure.** Protein surface conservation was calculated by the online ConSurf server (http://consurf.tau.ac.il) [68]. Surface hydrophobicity coloring was based on the defined hydrophobic scale [78]. Surface electrostatic potential was calculated with the APBS [79] plugin in PyMOL. All structural figures were generated using PyMOL (The PyMOL Molecular Graphics System, Version 1.8.X, Schrödinger, LLC). The sequences of SidC and SdcA were aligned using Clustal Omega [80] and colored by the Multiple Align Show online server (http://www.bioinformatics.org/sms/index.html).

CHAPTER 4

# THE MECHANISM OF PHOSPHORIBOSYL-UBIQUITINATION MEDIATED BY A SINGLE LEGIONELLA EFFECTOR

Akturk, A.*, Wasilko, D. J.*, Wu, X., Liu, Y., Zhang, Y., Qiu, J., Luo, ZQ, Reiter, K.H., Brzovic, P.S., Klevit, R.E., Mao, Y. (2018). Mechanism of phosphoribosyl-ubiquitination mediated by a single Legionella effector. *Nature*, 557, 729-733.

* - *Contributed equally to the work*

## 4.1 Abstract

Ubiquitination is a post-translational modification that regulates a myriad of cellular processes in eukaryotes [81-84]. The conventional ubiquitination cascade culminates in a covalent linkage between the C-terminus of ubiquitin (Ub) and a target protein, most often on a lysine sidechain [81, 85]. Recent studies of the *Legionella pneumophila* SidE family of effector proteins revealed a novel mode of ubiquitination in which a phosphoribosylated ubiquitin (PR-Ub) is conjugated to a serine residue on substrates via a phosphodiester bond [26, 27, 86]. To uncover the molecular mechanism of this unique post-translational modification, we determined the crystal structure of a fragment of the SidE family member SdeA that retains ubiquitination activity. The structure reveals that the catalytic module contains two distinct functional units: a phosphodiesterase domain (PDE) and a mono-ADP-ribosyltransferase (mART) domain. Biochemical analysis shows that the mART domain-mediated conversion of Ub to ADP-ribosylated Ub (ADPR-Ub) and the PDE domain-mediated ligation of PR-Ub to substrates are two independent activities of SdeA. Furthermore, we present two crystal structures of a homologous PDE domain from the SidE family member SdeD [87] in complex with Ub or ADPR-Ub. The

structures suggest an intriguing mechanism for how SdeA processes ADPR-Ub to PR-Ub plus AMP and conjugates PR-Ub to a serine residue in substrates. Our study establishes the molecular mechanism of phosphoribosyl-ubiquitination (PR-ubiquitination) and paves the way for future studies of this unusual type of ubiquitination in eukaryotes.

## 4.2 Introduction

A variety of microbial pathogens exploit the eukaryotic ubiquitination pathway during their respective infections [47, 88]. The intracellular pathogen *L. pneumophila* injects more than 300 effectors into host cells during its infection, including at least 10 proteins that are involved in ubiquitin manipulation [40]. These effectors include HECT-like [61, 62] and F- or U-box-containing Ub ligases [89-92] as well as novel Ub ligases of the SidE family, such as SdeA, that act independently of canonical E1 and E2 enzymes [26, 27, 86]. SdeA first uses its mART activity to catalyze the transfer of ADP-ribose from $NAD^+$ to the sidechain of R42 on Ub to generate ADPR-Ub. Subsequently, SdeA uses its PDE activity to catalyze the conjugation of ADPR-Ub to a serine residue on substrates to generate a protein~phosphoribosyl-Ub (Protein~PR-Ub) product. Alternatively, in the absence of substrates, the SdeA PDE domain will catalyze the hydrolysis of ADPR-Ub to generate PR-Ub and AMP (Figure 4.1A and Figure 4.2). The molecular mechanism of this unique ubiquitination pathway has yet to be determined.

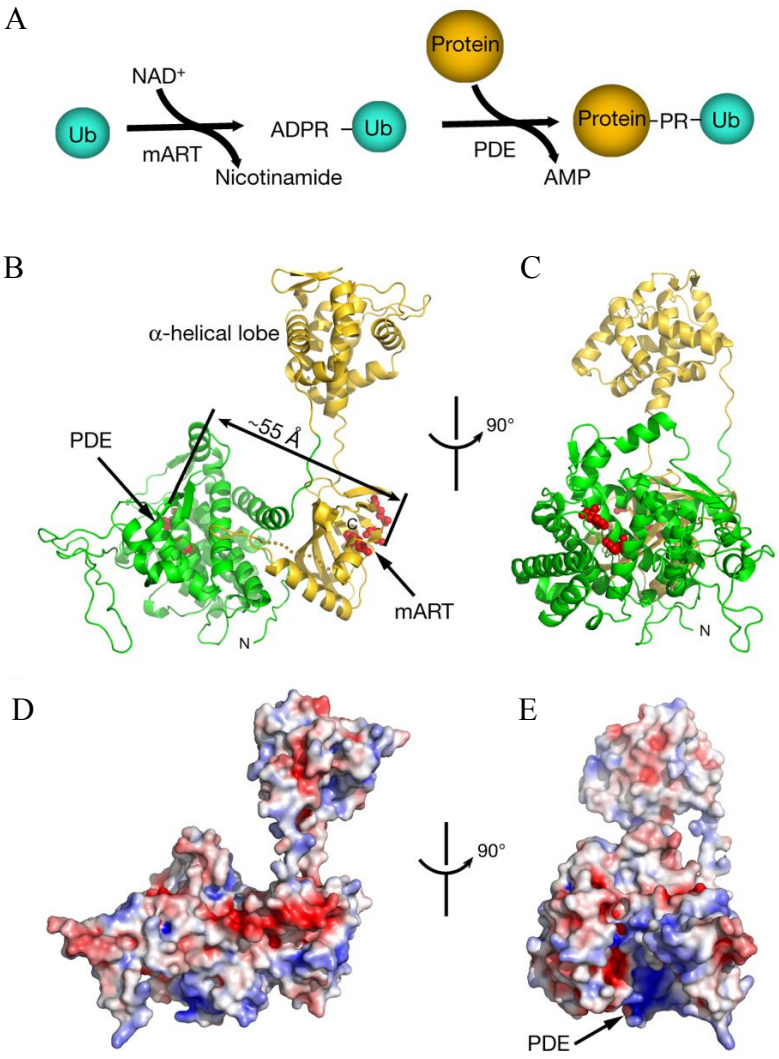


**Figure 4.1: Overall structure of SdeA.** (A) Schematic of the phosphoribosyl-ubiquitination reaction. (B) Overall structure of SdeA-core in ribbon representation. This portion of SdeA has two distinct domains: the PDE (green) and mART (gold) domains. The active site residues of both the mART and PDE domains are shown as red spheres. The linear distance between these two active sites is approximately 55 Å. (C) An orthogonal view of (B). (D) Molecular surface model of SdeA. The surface is coloured on the basis of electrostatic potential with positively charged regions in blue and negatively charged surfaces in red. The orientation of the molecule is the same as shown in (B). (E) An orthogonal view of (D). (Dr. Anil Akturk and Xiaochun Wu crystallized and determined this structure)

**Table 4.1: SdeA/SdeD Data collection, phasing, and structural refinement statistics.**

| | SdeA | SdeD | SdeD-Ub | SdeD-ADPRUB-Ub |
|---|---|---|---|---|
| **Data collection** | | | | |
| Synchrotron beam lines | MCCHESS F1 | MCCHESS A1 | MCCHESS F1 | MCCHESS F1 |
| Wavelength (Å) | 0.9789 | 0.68 | 0.9789 | 0.9789 |
| Space group | $P2_1$ | R3 | $P2_1$ | $P2_1$ |
| Cell dimensions | | | | |
| *a, b, c* (Å) | 69.8, 80.6, 85.6 | 154.4, 154.4, 89.6 | 64.8, 58.6, 74.1 | 64.7, 58.8, 75.1 |
| $\alpha, \beta, \gamma$ (°) | 90, 109.8, 90 | 90, 120, 90 | 90, 114.6, 90 | 90, 114.2, 90 |
| Maximum resolution (Å) | 2.2 | 1.51 | 1.73 | 1.88 |
| Observed reflections | 61,395 | 634,900 | 363,307 | 281,813 |
| Unique reflections | 18,728 | 124,885 | 108,100 | 43,941 |
| Completeness (%)* | 99.3 | 99.5 | 99.4 | 100 |
| Redundancy* | 3.4(3.3) | 5.1(2.9) | 3.4(2.2) | 6.4(5.9) |
| $\langle I \rangle / \langle \sigma I \rangle$* | 7.98 (0.87) | 29.2 (1.52) | 25.4 (1.54) | 19.28 (1.18) |
| $R_{sym}$ (%)* | 0.122(0.759) | 0.07(0.622) | 0.078(0.798) | 0.093(1.105) |
| **Refinement** | | | | |
| Resolution (Å)* | 80.51(2.20) | 77.174(1.51) | 67.36(1.70) | 68.93(1.85) |
| $R_{crys}$ / $R_{free}$ (%)* | 0.195/0.241 | 0.168/0.194 | 0.160/0.196 | 0.159/0.201 |
| No. atoms | | | | |
| Protein | 5338 | 4932 | 3721 | 3765 |
| Ligand/ion | 10 | -- | -- | -- |
| Water | 170 | 448 | 228 | 211 |
| *B*-factors | | | | |
| Protein | 49.728 | 24.998 | 28.036 | 33.327 |
| Ligand/ion | 61.563 | -- | -- | -- |
| Water | 46.001 | 31.136 | 30.525 | 34.580 |
| R.m.s deviations | | | | |
| Bond length (Å) | 0.023 | 0.027 | 0.03 | 0.028 |
| Bond angles (°) | 2.29 | 2.47 | 2.47 | 2.43 |

* Values in parentheses are for highest-resolution shell.

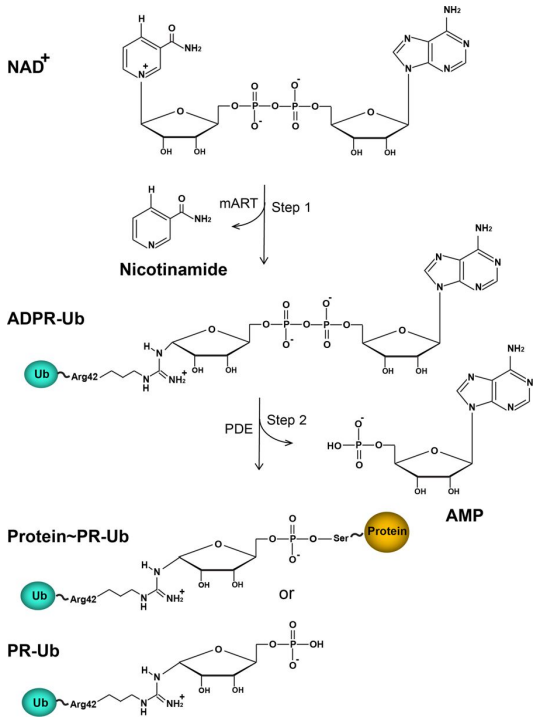


**Figure 4.2: Chemical structure of phosphoribosyl-linked ubiquitination catalysed by SdeA.** Phosphoribosyl-linked ubiquitination catalysed by SdeA involves two enzymatic activities of SdeA. First, using its mART activity, SdeA catalyses the ADP-ribosylation of Ub to generate ADPR-Ub by consuming an $NAD^+$ molecule. Second, SdeA catalyses the conjugation of ADPR-Ub to a serine residue of substrate proteins via its PDE activity to generate protein–PR-Ub and AMP. In the absence of substrate proteins, the PDE domain of SdeA can simply hydrolyse ADPR-Ub to PR-Ub and AMP using a water molecule.

## 4.3 Results

### 4.3.1 Crystal structure of SdeA-Core

To decipher the mechanism of PR-ubiquitination, we determined the crystal structure of a portion of SdeA (SdeA-Core, a.a. 211-910; Table 4.1). The structure is composed of two distinct domains, the PDE and mART domains (Figure 4.1B and C). A calculation of surface electrostatic potential revealed no significantly charged areas on the surface of SdeA other than a deep, highly positively charged groove on the PDE domain (Figure 4.1D and E). Analogous to other PDEs [93], the active site is likely harbored in this deep groove (Figure 4.3A-C). Indeed, a sequence alignment

of PDE domains showed that most of the conserved residues reside in this groove, consistent with their forming the PDE active site (Figures 4.2D and 4.4). The mART domain is composed of two lobes, an N-terminal α-helical lobe (a.a. 592-758) and a main lobe (a.a. 759-911). The main lobe contains a β-sandwich core and harbors the three catalytic motifs: the (F/Y)-(R/H), STS, and EXE motifs (Figures 4.5A-F and 4.6) conserved in other mART proteins, such as the *Pseudomonas syringae* effector HopU1 and the *Clostridium perfringens* toxin Iota-toxin [94-96]. A structural comparison of the α-helical lobe with its counterpart in other mARTs revealed that although the total number and the length of α-helices are variable, three α-helices form a structural core that is conserved in most mART proteins (Figure 4.5G-I). Surprisingly, while it packs in close contact with the main lobe in other mARTs, the α-helical lobe is extended away from the main lobe in our SdeA-Core crystal structure (Figure 4.7A and B). The extended conformation observed in our crystal structure is consistent with the conformation in solution as detected by small angle X-ray scattering (SAXS) and does not change in the presence of $NAD^+$ (Figure 4.7 C-F). However, the α-helical lobe adopts a closed conformation and mediates contact with $NAD^+$ in a structure of Iota-toxin [95]. Moreover, the α-helical lobe is enriched with highly conserved residues (including N723, Q727, and R729) that form a cluster on its surface, as revealed by an analysis of surface residue conservation with the ConSurf server [68] (Figures 4.6 and 4.8A). Thus, we hypothesized that the α-helical lobe may play a similar role in catalysis by SdeA. Indeed, SdeA carrying an α-helical lobe deletion (SdeA-Δα-lobe), as well as SdeA carrying any of the N723A, Q727A, or R729A point mutations in the α-helical lobe completely lacked Ub ADP-ribosylation activity (Figure 4.8B and C). The F719A mutation, in a residue that is not conserved but is close to the conserved surface patch, yielded a substantial impairment of activity, while the D622A mutation of a conserved residue that is away from the patch yielded an activity level comparable to wild

type SdeA. Taken together, our data show that the α-helical lobe is crucial for Ub ADP-ribosylation, and a surface patch composed of highly conserved residues may mediate the binding of $NAD^+$ during catalysis. These observations further suggest that the closed conformation of the α-helical lobe is required for the mART activity of SdeA. Indeed, an accompanying paper describing the crystal structure of a longer construct of SdeA in complex with both $NAD^+$ and Ub reports that the α-helical lobe is observed in a closed conformation [29].

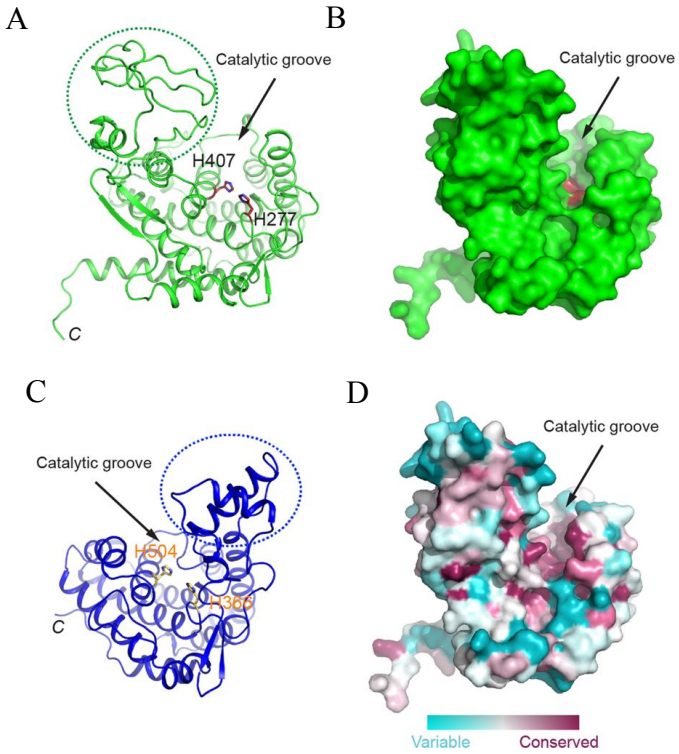


**Figure 4.3: Structure of the PDE domain of SdeA.** (A) Model of the PDE domain of SdeA in ribbon representation. Two invariable histidine residues (H277 and H407) are shown in stick representation and labelled. (B) Surface representation of the PDE domain. The two invariable histidine residues (shown in red) are situated at the bottom of a deep groove. (C) The PDE domain from a Legionella effector (lpg1496). Notably the all α-helical structural core of the PDE domains is easy to superimpose onto that of SdeA with a root mean square deviation (r.m.s.d.) of 1.9 Å over 225 aligned Cα atoms. A prominent difference between the two PDE domains is that some loops (indicated by dashed outlines) connecting the α-helices vary both in primary sequence and in length (Figure 4.4). (D) Surface residue conservation analysis of the PDE domain. The conservation is calculated using the ConSurf server with the most conserved residues coloured in purple and the least conserved residues in cyan. Note that the catalytic groove is enriched with the most conserved residues.

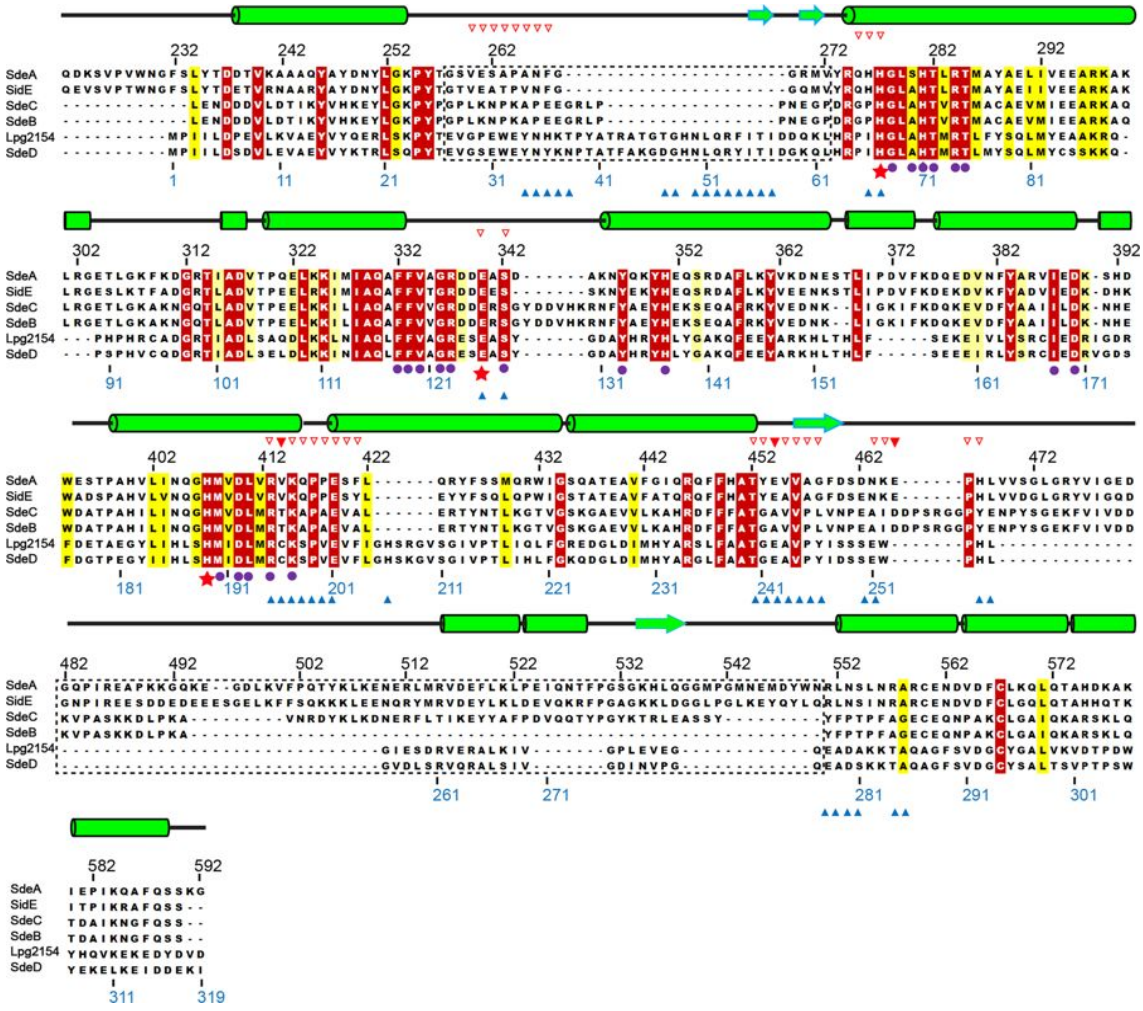


**Figure 4.4: Multiple sequence alignment of selected PDE domains from the SidE family effectors.** Representative sequences corresponding to the PDE domain of SdeA (amino acids 222–502) were aligned using the MultAlin online server (http://www.bioinformatics.org/sms/index.html). Secondary structural elements are drawn above the alignment. The numbering for the SdeA sequence is marked on the top of the alignment and the numbering for the SdeD sequence is marked below. Variable loop regions are outlined with dashed squares. Conserved residues located within the catalytic groove are highlighted with purple dots. In particular, three essential catalytic residues (H277, H407 and E340) are highlighted with red stars below the sequences. SdeD residues that are in close contact with Ub1 (Figure 4.11 (A)) are marked by blue triangles at the bottom of the sequences and the predicted Ub1-interacting residues of the PDE domain of SdeA (Figure 4.11 (E)) are depicted by red triangles on the top of the sequences. Amongst the potential Ub1-interacting residues, V414, E454 and E465 of SdeA used in mutagenesis studies in Figure 4.11 (F,G) are marked with solid red triangles. Entrez database accession numbers are as follows: SdeA, GI: 1064303039; SidE, GI: 52840489; SdeB, GI: 52842367; SdeC, GI: 52842370; lpg2154, GI: 52842368; and SdeD, GI: 52842717.

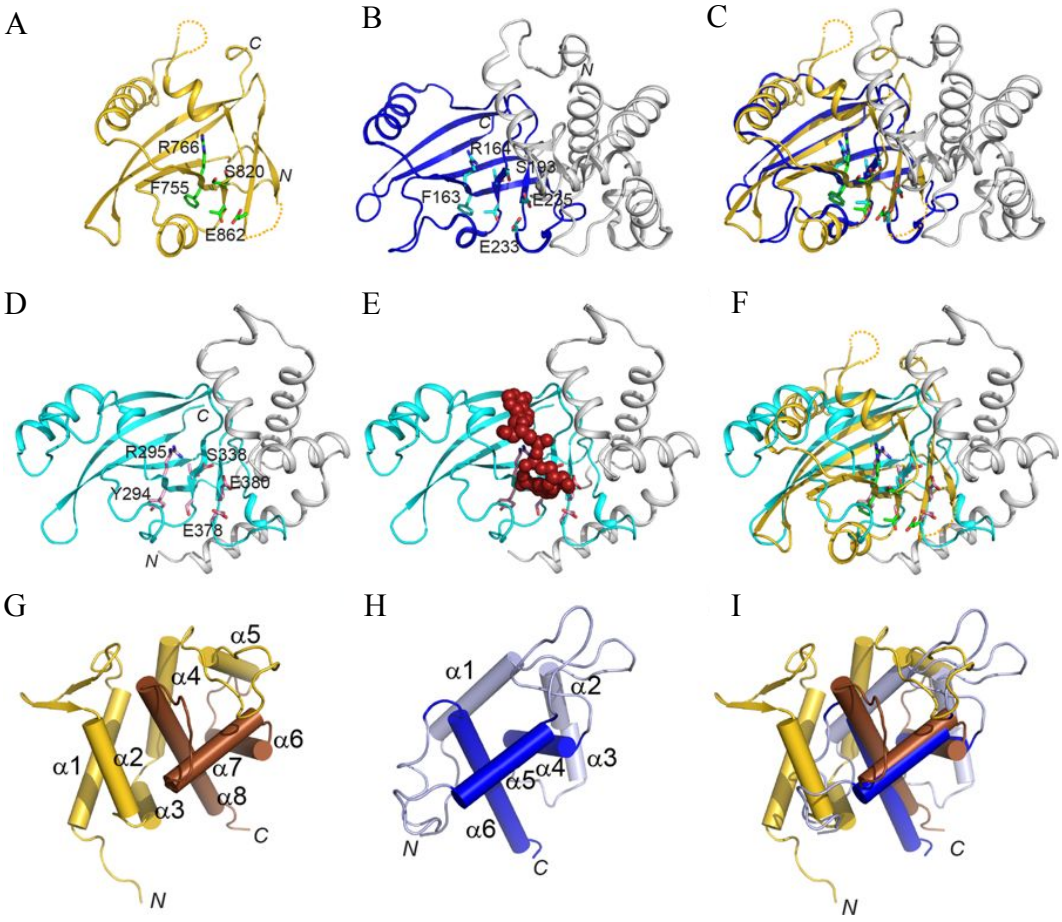


**Figure 4.5: Structural comparison of the SdeA mART domain with other mART domains from bacterial toxins.** (A) Model of the main lobe of the SdeA mART domain in ribbon representation. The main lobe is composed of two nearly perpendicular β-sheets forming a two-layered β-sandwich core. Residues comprising the three mART catalytic signature motifs: (F/Y)-(R/H), STS and EXE motif are shown in sticks. (B) HopU1 from P. syringae (PDB ID: 3U0J) in ribbon representation. (C) Structural superimposition of the mART domains from SdeA (gold) and HopU1 (blue). (D) Iota-toxin from C. perfringens (PDB ID: 4H03). (E) Iota-toxin in complex with $NAD^+$ (red spheres). (F) Structural overlay of the mART domains from SdeA (gold) and iota-toxin (cyan). (G) A cartoon diagram of the α-helical lobe of the SdeA mART domain. The α-helical lobe consists of eight α-helices. Three structurally conserved α-helices (α6–8) are coloured in brown. (H) A cartoon diagram of the α-helical lobe of HopU1, the three equivalent α-helices (α4–6) are highlighted in blue. (I) Structural overlay of the α-helical lobe of SdeA and HopU1.

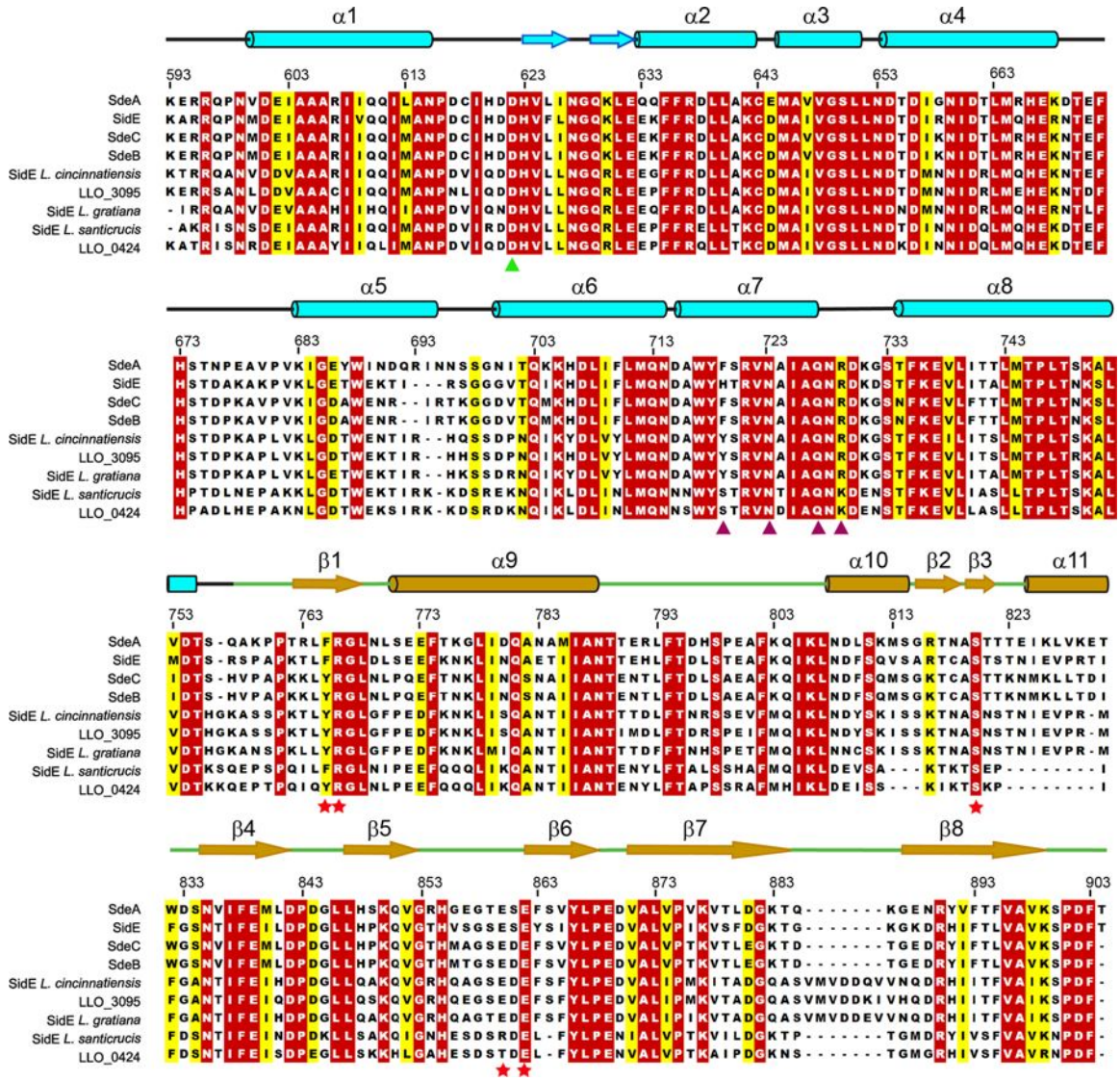


**Figure 4.6: Multiple sequence alignment of the mART domains.** Representative sequences corresponding to the mART domains of SdeA (amino acids 593–904) were aligned using MultAlin. Secondary structural elements (cyan for the α-helical lobe and gold for the main lobe of the mART domain) are drawn above the alignment. The numbering for the SdeA sequence is marked on the top of the alignment. Residues comprising the catalytically important (F/Y)-(R/H), STS and EXE motifs are marked with red stars. Residues in the α-helical lobe, which form—or are close to—the conserved surface patch and are essential for the mART activity (Figure 4.8), are marked with purple triangles. D622, which is conserved but has no effect on the mART activity is marked with a green triangle. Entrez database accession numbers are as follows: SdeA, GI: 1064303039; SidE, GI: 52840489; SdeB, GI: 52842367; SdeC, GI: 52842370; SidE Legionella cincinnatiensis, GI: 966421657; LLO_3095, GI: 489730495; SidE Legionella gratiana, GI: 966468332; SidE Legionella santicrucis, GI: 966496250; LLO_0424, GI: 502743808.

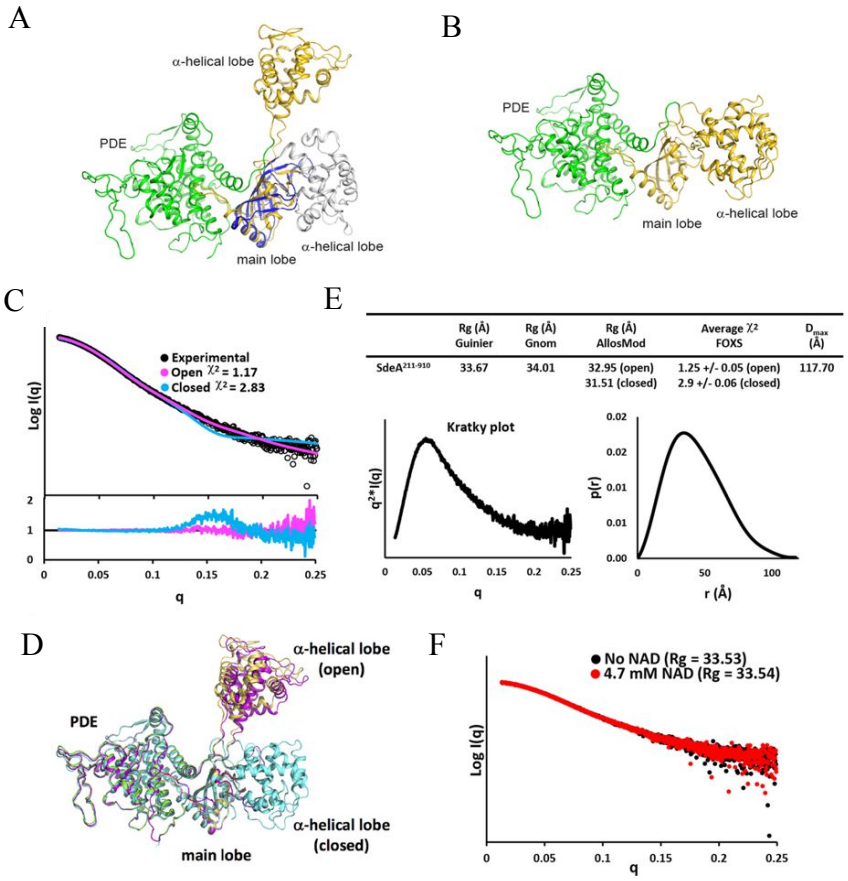


**Figure 4.7: The α-helical lobe of SdeA mART domain has an extended conformation compared to other mART proteins.** (A) Structural superimposition of SdeA onto the HopU1 structure referenced on the main lobe of the mART domain. SdeA is coloured using the same scheme as Figure 4.1 (B). The main lobe of HopU1 is coloured in blue and its α-helical lobe is in grey. The α-helical lobe of the SdeA mART is extended away from the main lobe whereas its counterpart in HopU1 packs in close contact with the main lobe. (B) Structural model of SdeA with the α-helical lobe in a closed conformation. The positioning of the α-helical lobe was based on a structural overlay of the three structurally conserved α helices identified in all mART domains (Figure 4.5 (G-I)). (C) Experimental and theoretical SAXS curves for SdeA-core and the resulting best-fit AllosMod structure for the determined structure (open) and modelled closed conformation, with residual plots shown below. Best fit $\chi^2$ values are indicated. (D) Overlay of the determined SdeA-core structure (PDE, green; mART main lobe and α-helical lobe, yellow) and best-fit AllosMod structures for the open (magenta) and closed (cyan) conformations. (E) Summary of the experimentally derived SAXS parameters for SdeA-core, AllosMod derived best-fit $R_g$ and average FOXS $\chi^2$ for the five best-fitting AllosMod models compared to the experimental SAXS curve. The program Primus was used to calculate the radius of gyration ($R_g$) and maximum linear dimension ($D_{max}$). Kratky plot ($I(q)q^2$ versus q), and distance-distribution plot P(r) obtained from GNOM are shown. (F) Overlay of SdeA-core SAXS curves in the presence of 4.7 mM $NAD^+$(10× protein concentration), with corresponding Guinier $R_g$ values. Data shown in (C), (E) and (F) are representative of two biologically independent experiments. (Work of Dr. Katherine Reiter)

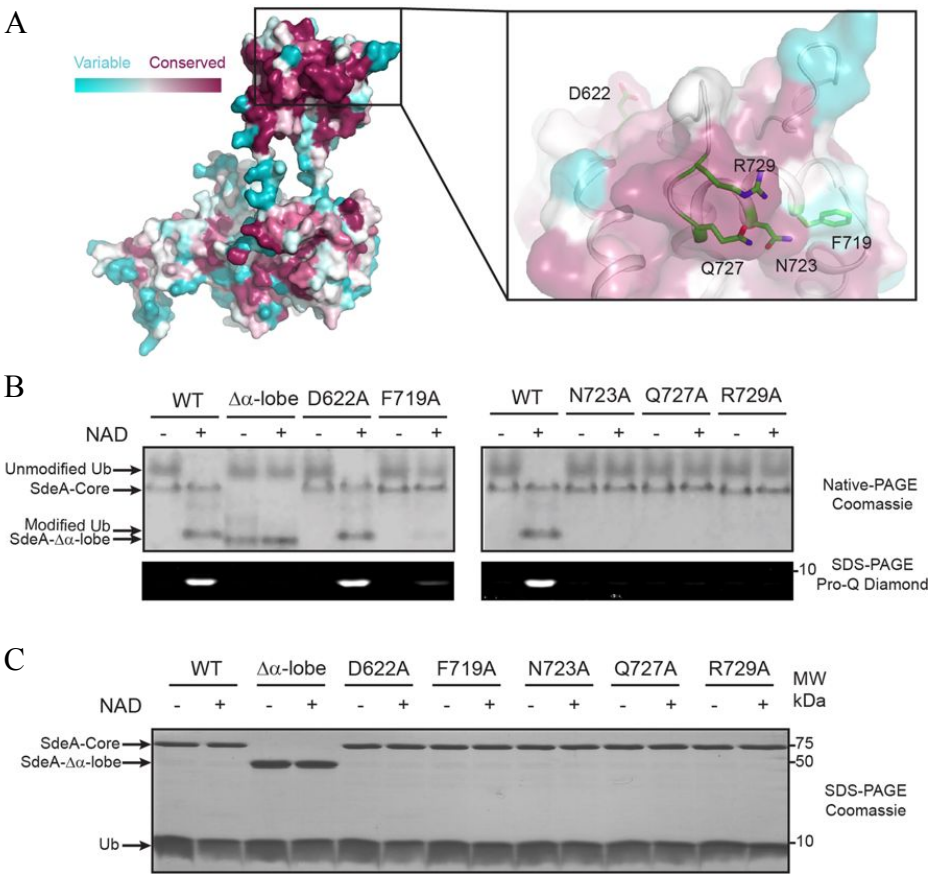


**Figure 4.8: The α-helical lobe of SdeA mART domain is indispensable for Ub ADP-ribosylation.** (A) Surface representation of residue conservation of SdeA (the most conserved residues are shown in purple and the least conserved residues in cyan). Surface residue conservation was calculated using the ConSurf server. An expanded view of a surface cluster that consists of the most conserved residues on the α-helical lobe is shown on the right. (B) Analysis of in vitro ubiquitin-modification assays by SdeA mutants carrying mutations on the α-helical lobe. The reaction products were analysed using native PAGE with Coomassie blue stain (top) and SDS–PAGE with Pro-Q phosphoprotein stain (bottom). (C) SDS–PAGE analysis of the proteins in the reaction mixture. Data shown in (B) and (C) are representative of three independent experiments.

### 4.3.2 ADP-ribosylation and serine PR-ubiquitination are independent activities

The main lobe of the mART domain is packed against the PDE domain in the SdeA structure. The two catalytic sites face in opposite directions and are separated by a distance of over 55 Å (Figure 4.1B), raising the interesting question: How are the activities of the two domains coordinated? To address this question, we performed assays with SdeA fragments that retain only mART or PDE activity (Figure 4.9A). Similar to wild-type SdeA-Core, reactions that contain both

SdeA-PDE and SdeA-mART efficiently generate PR-Ub and ubiquitinate the substrate Rab33b (Figure 4.9B and C). SdeA-Core carrying a PDE active-site residue mutation (H277A) retained the ability to generate ADPR-Ub but failed to process ADPR-Ub to PR-Ub or to ubiquitinate Rab33b. However, the presence of both SdeA-Core (H277A) and SdeA-PDE successfully catalyzed both the production of PR-Ub and the ubiquitination of Rab33b. Moreover, SdeA-PDE alone can PR-ubiquitinate Rab33b when purified ADPR-Ub is supplied (Figure 4.9D). The independence of the two activities was further validated by SdeA-mediated Rab33b ubiquitination when co-expressed in cells (Figure 3.9E). Taken together, these results suggest that ADP-ribosylation of Ub and serine PR-ubiquitination are two mechanistically and spatially-independent activities housed in a single protein.

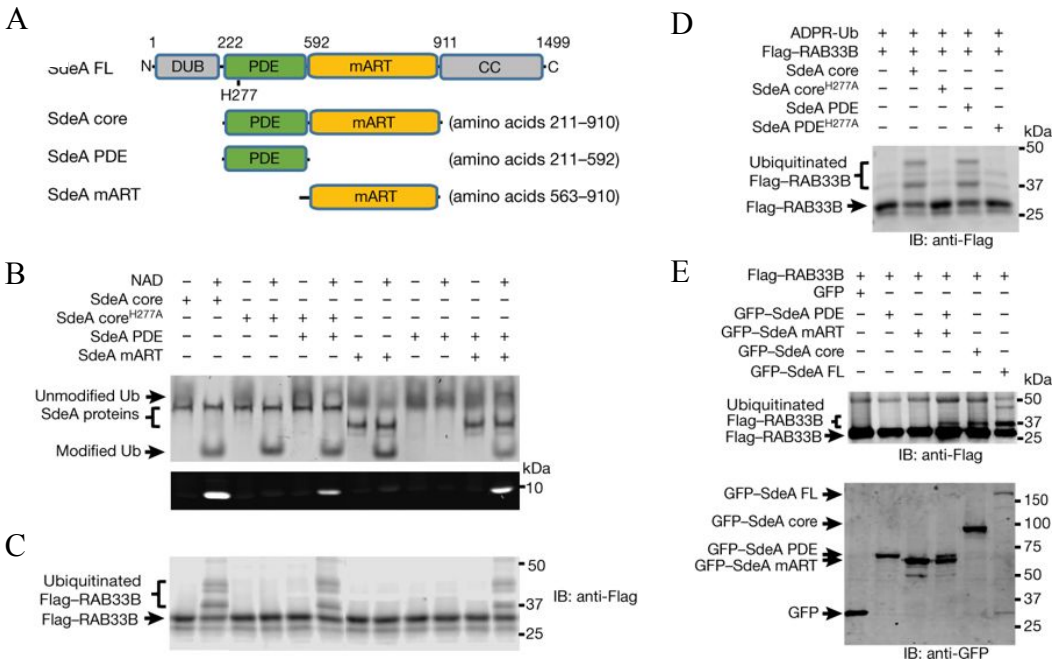


**Figure 4.9: ADP-ribosylation of Ub and phosphoribosyl ubiquitination of serine are two independent activities of SdeA.** (A) Schematic diagram of SdeA constructs. SdeA has an N-terminal deubiquitinase (DUB) domain, followed by PDE, mART, and C-terminal coiled-coil (CC) domains. (B) In vitro Ub-modification assays. The modification of Ub to ADPR-Ub or PR-Ub was monitored by the band-shift of Ub in native PAGE with Comassie staining (top). The production of PR-Ub was visualized by SDS-PAGE and phosphoprotein staining with Pro-Q Diamond (bottom). ADPR-Ub and PR-Ub migrate at the same position on a native gel (labelled as modified Ub), however, only PR-Ub is visible by Pro-Q phosphoprotein stain. (C) In vitro phosphoribosyl-ubiquitination assay of RAB33B by indicated the SdeA proteins. IB, immunoblot. (D) In vitro phosphoribosyl-ubiquitination assay of RAB33B in the presence of purified ADPR-Ub. (E) Intracellular-ubiquitination assays of RAB33B by SdeA. Data shown in (B-D) are representative of four independent experiments. GFP, green fluorescent protein. (E) Similar results were obtained from three independent experiments. (Work of Dr. Anil Akturk and Dr. Zhao-Qing Luo)

### 4.3.3 Ubiquitin recognition by the PDE domain

Despite sharing 23% sequence similarity with a well-characterized cyclic di-3',5'-GMP phosphodiesterase in *Pseudomonas aeruginosa* PA4781 [97], the PDE domain of SdeA uses ADPR-Ub as its substrate and catalyzes the unprecedented serine PR-ubiquitination reaction. This striking difference raised the question of how ADPR-Ub is recognized by the SdeA PDE domain. To address this question, we assessed the interaction of Ub and several homologous PDE domains

from the *Legionella* SidE family effectors using $^{1}H$, $^{15}N$-HSQC-TROSY NMR titration experiments (Figure 4.10A-C). The SdeA PDE domain showed no detectable interaction with Ub in solution, while the PDE domain of another SidE family member, SdeD, exhibited a direct and specific interaction with Ub as evidenced from NMR peak perturbations. We then successfully determined the structures of SdeD, both on its own and in complex with Ub (Figure 4.10D-F). Unexpectedly, two Ub molecules are in contact with a single PDE domain in the crystal. One Ub (Ub2) binds on the opposite side from the catalytic groove, making the physiological significance of this binding mode unclear (Figure 4.10G). The other Ub (Ub1) binds to a flat surface at the opening of the catalytic groove (Figure 4.11A). Similar to the Ub surface area mapped by NMR titration experiments in solution (Figure 4.10C), three regions of Ub1 contact the PDE domain: the loop region around residue T9, the C-terminus, and a region that includes R42 (Figure 4.11A). At the T9 loop region, beside hydrophobic interactions mainly contributed by L8, residue K6 of Ub1 forms electrostatic interactions with E251 on SdeD (Figure 4.11B). At the C-terminus of Ub1, in addition to hydrophobic interactions mediated by L73, R72 of Ub1 forms salt bridges with E242 on SdeD (Figure 4.11C). Notably, the R42 sidechain of Ub1 extends into the catalytic groove and forms hydrogen bonds and electrostatic interactions with the conserved residues Q52 and E126 at the PDE catalytic site (Figure 4.11D). To test whether the PDE domain of SdeA interacts with Ub in a similar mode as observed for SdeD, we modeled Ub binding by the PDE domain of SdeA based on the SdeD-Ub1 complex (Figure 4.11E). The model predicts that E465 and E454 in SdeA would play analogous roles in Ub binding to E251 and E242 in SdeD, respectively (Figure 4.11A and E). Consistent with this prediction, PDE activity was substantially impaired in SdeA E465A and E454A mutants as evidenced by the marked reduction of both the Pro-Q staining signal and Rab33b ubiquitination (Figure 4.11F and G). In further validation of the model, a V414Y mutant

designed to sterically block the access of ADPR-Ub to the catalytic site also largely impaired the PDE activity (Fig 4.11E-G). All three SdeA mutants were able to cause the band-shift of Ub on the native gel (Figure 4.11E top panel) indicating that the mART activity of these mutants remained intact. Together, these data support the notion that the SdeA PDE domain uses a similar strategy for Ub recognition as observed in SdeD, although the interaction is markedly weaker as evidenced by the aforementioned NMR titration analysis.

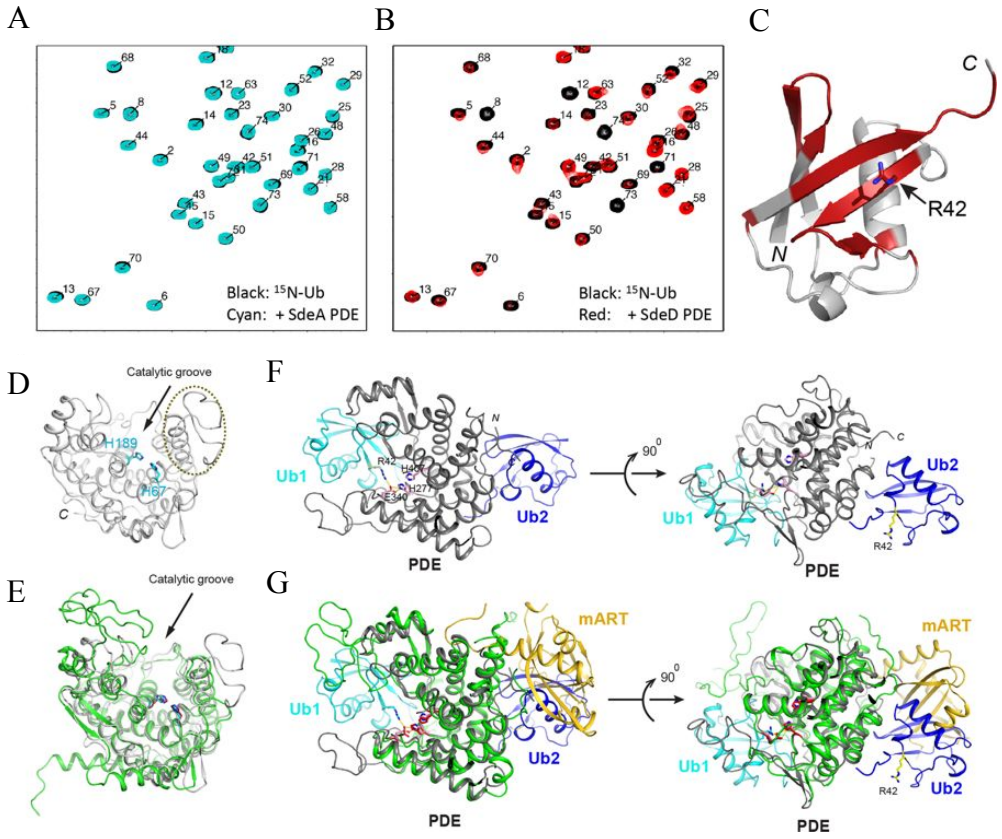


**Figure 4.10: The interaction between Ub and the SdeD PDE domain.** (A) NMR $^{1}H–^{15}N$ HSQC TROSY spectral overlay of 150 μM Ub (black) in the presence or absence of 300 μM SdeA PDE domain (cyan). Ub binds very weakly to SdeA as manifested by minimal changes in $^{15}NH$ peaks of Ub. (B) Spectral overlay of 150 μM Ub (black) with 75 μM SdeD PDE. Ub binds with higher affinity to SdeD as evidenced by peak broadening and/or disappearance of Ub resonances. (C) Residues whose resonances are most affected by the presence of SdeD are mapped in red on a cartoon structure of Ub. (D) PDE domain of SdeD (grey) shown in ribbon representation. Two invariable histidine residues (H67 and H189) are shown in stick representation (cyan). The variable loop unique to SdeD is outlined. (E) Structural overlay of the PDE domain of SdeD (grey) and the PDE domain of SdeA (green). The overall structures of these two PDE domains are very similar with an r.m.s.d. of 1.73 Å over 251 overlaid Cα atoms. (F) Two orthogonal views of the SdeD PDE domain in complex with two Ub molecules in ribbon representation: Ub1 (cyan) and Ub2 (blue). Ub1 binds at the opening of the PDE catalytic groove with its R42 side chain sticking into the groove. Ub2 binds a region on the opposite side of the catalytic groove. (G) Structural superimposition of SdeA onto the SdeD PDE–Ub complex referenced on the PDE domain. The PDE domain of SdeA is shown in green and the mART domain is shown in gold. Note that Ub1 shows no conflicting contacts against the superimposed SdeA molecule whereas the Ub2 binding site largely overlaps with the space occupied by the mART domain in SdeA. This analysis suggests that the binding of the PDE domain of SdeD to Ub1 is probably applicable to the PDE domain of SdeA; however, the second Ub-binding site observed in SdeD might not exist in SdeA. Experiments in (A) and (B) were repeated independently two times. (Work of Dr. Peter Brzovic)

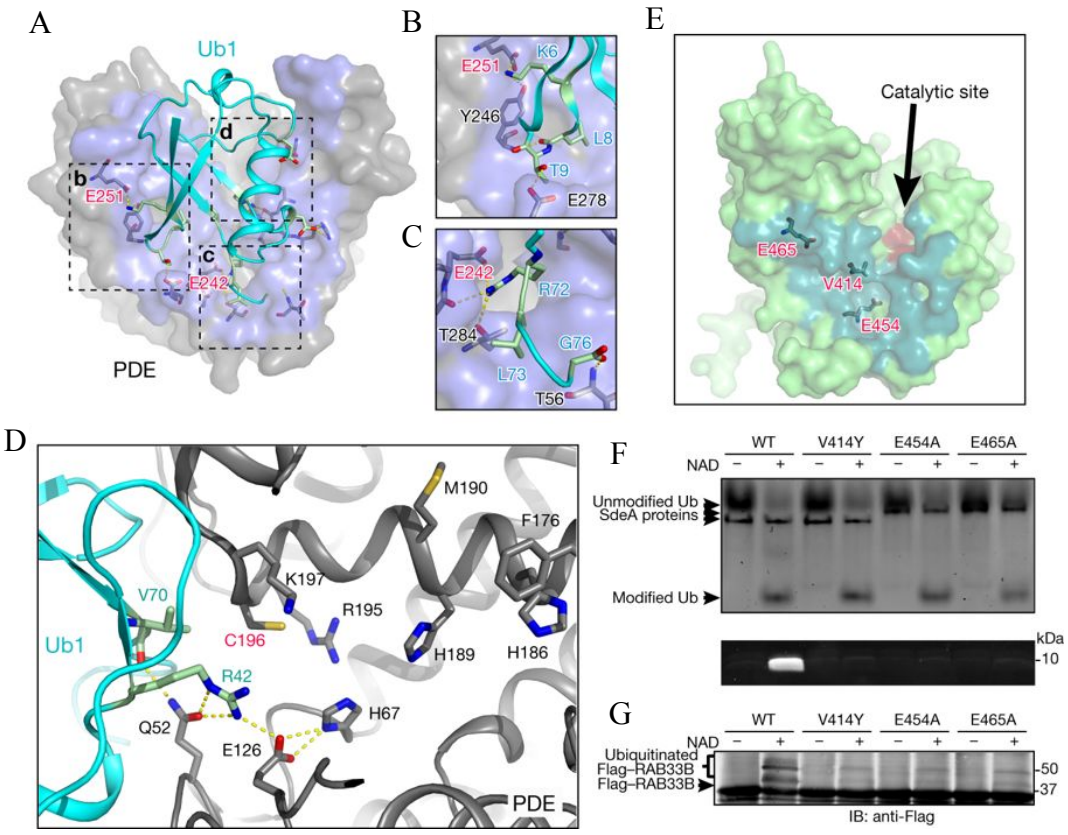


**Figure 4.11: The interaction between Ub and the PDE domains of SdeD and SdeA.** (A) Overall view of the binding of Ub (Ub1) with the PDE domain of SdeD. The PDE domain residues within Van der Waals distance of Ub1 are coloured in light blue. Three interacting regions of Ub1 that contact SdeD are marked by dashed outlines. (B-D) Expanded views of the three Ub1–SdeD interacting regions outlined in (A). (E) Surface representation of the PDE domain of SdeA. Ub-binding was modelled the SdeD–Ub1 complex structure and the potential Ub-interacting surface is highlighted in dark green. Three key residues (E465, E454 and V414) at the potential Ub-interacting interface are shown in stick representation. The PDE active site is shown in red. (F),(G) In vitro Ub-modification (F) and phosphoribosyl-ubiquitination assays (G) of SdeA mutants at the potential Ub interacting interface. The modification of Ub and phosphoribosyl-linked ubiquitination were monitored as described in Figure 4.9 (B,C). Data shown in (F) and (G) are representative of four independent experiments. WT, wild type.

### 4.3.4 ADPR-Ub binding to the active site of the PDE

To further address the question of how the ADPR moiety of ADPR-Ub fits in the active-site groove of the PDE domain, we determined the structure of a complex of a catalytically-inactive SdeD (H67A) mutant and ADPR-Ub. ADPR-Ub uses a similar mode of binding as Ub1 with the

ADPR moiety nestled in the catalytic groove (Figure 4.12A-D). ADPR sits atop several invariant residues, including H67A, H189, and E126, and engages in extensive interactions with a large number of conserved residues within the catalytic groove (Figures 4.13A-C and 4.12E). To test the role of the ADPR-interacting residues within the catalytic groove, we mutated several corresponding residues in SdeA. PDE activity was completely abolished in the H277A, H407A, and E340A mutants, as indicated by the lack of both the Pro-Q staining signal and Rab33b ubiquitination (Figure 4.13C and D). The activity of R413A was substantially impaired, while H281A and W394A showed little or no effect on PDE activity.

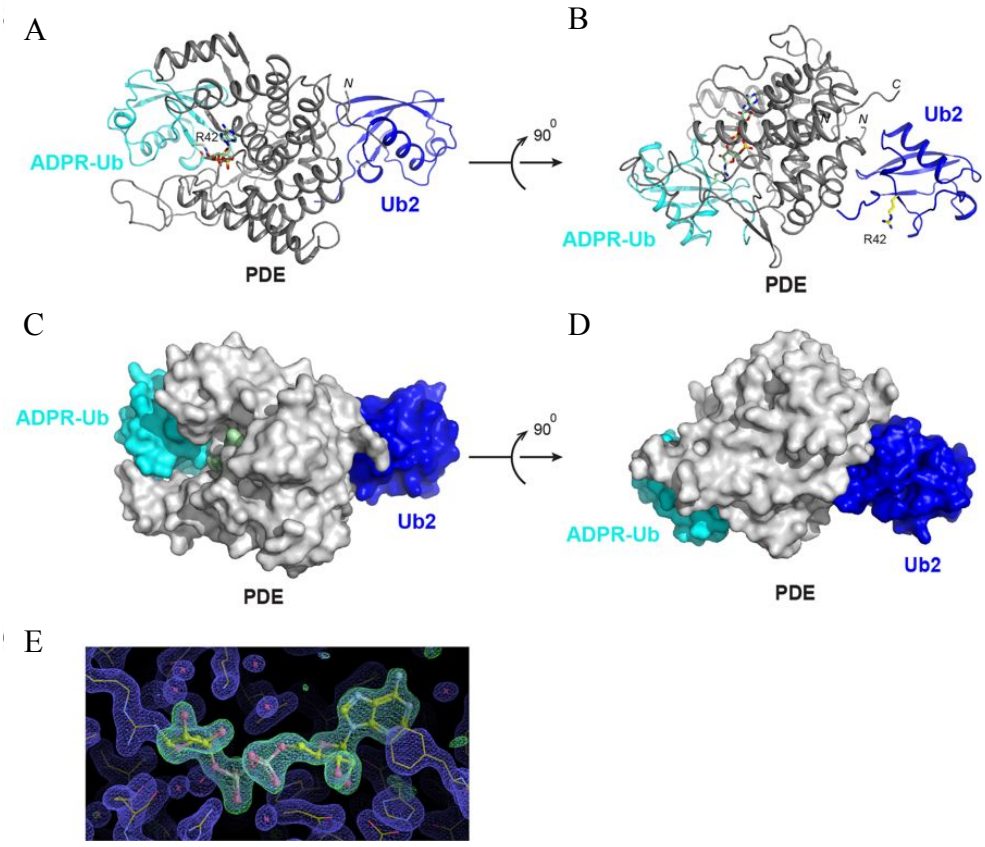


**Figure 4.12: Crystal structure of the PDE domain of SdeD in complex with ADPR-Ub and Ub.** (A) SdeD PDE domain H67A mutant in complex with both ADPR-Ub and unmodified Ub. The crystal was obtained by mixing the SdeD PDE H67A mutant, ADPR-Ub, and Ub in a 1:2:3 molar ratio (see the 'Protein crystallization' section of the Methods for details). The PDE domain is shown in grey, the bound ADPR-Ub is shown in cyan and the unmodified Ub is shown in blue. The unmodified Ub binds a region identical to Ub2 found in the SdeD–Ub complex shown in Figure 4.10 (F). ADPR-Ub binds in a mode that is similar to that of Ub1 in the SdeD–Ub complex with the ADPR moiety fitting into the catalytic groove. (B) An orthogonal view of (A). (C), (D), Two orthogonal views of the complex shown in a in surface representation. Note that the ADPR-moiety shown in light green fits deeply into the catalytic groove. (E), The density was generated by refinement against the structural model without the ADPR portion. The $F_o - F_c$ difference map is shown in green and contoured at 1σ.

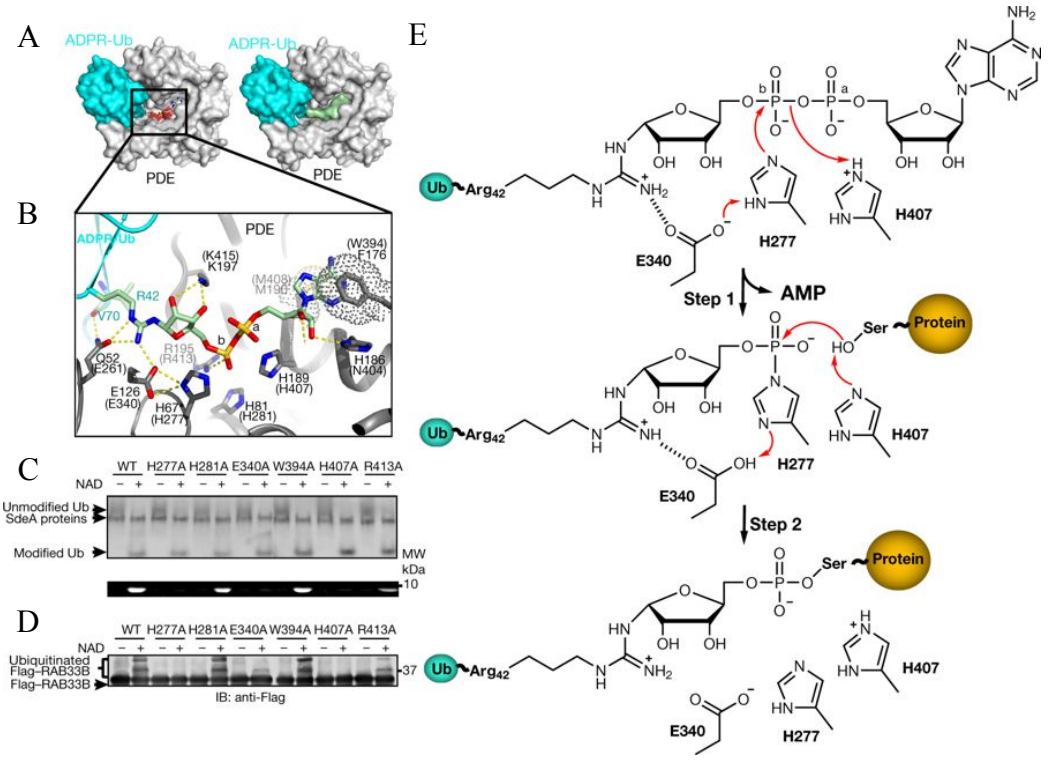


**Figure 4.13: Structure of the complex formed by ADPR-Ub and the PDE domain of SdeD.** (A) Surface representation of ADPR-Ub (cyan) in complex with the SdeD PDE domain (grey). The catalytic site is coloured in red. The ADPR moiety is coloured in light green and shown in stick (left) and surface (right) representation. (B) A detailed interaction of the ADPR moiety with residues of the PDE domain. SdeD residues involved in ADPR-binding are labelled and the corresponding residues in SdeA are labelled in parentheses. In the structure, H67 is substituted with alanine, but is modelled with histidine and labelled as H67*. (C) Enzymatic-activity analysis of SdeA-core with mutations in conserved residues of the catalytic groove. The modification of Ub was monitored as described in Figure 4.9 (B). (D) Phosphoribosyl-ubiquitination assay of RAB33B. (E) A two-step reaction model of phosphoribosyl-linked ubiquitination catalysed by the PDE domain of SdeA. Data shown in (C) and (D) are representative of three independent experiments.

### 4.3.5 Mechanism of substrate PR-Ubiquitination

Based on our results, we propose a two-step reaction mechanism for the transfer of PR-Ub to a substrate (Figure 4.13E). In the first step, negatively-charged E340 helps to position both R42 of ADPR-Ub and H277. This interaction could enhance the nucleophilicity of H277 through induction. H277 attacks the β-phosphate of ADPR to form a transient phosphoramidate bond with PR-Ub. The presence of this transient intermediate is supported by biochemical evidence reported

in an accompanying paper by Dikic and colleagues [98]. The nearby H407 functions as a general acid to donate a proton to the α-phosphate of the releasing AMP molecule. The underlying mechanism of this step is similar to that of histidine protein kinases [99], [100]. In the second step, H407 deprotonates the hydroxyl of a serine residue of the approaching substrate. The activated hydroxyl then attacks the phosphoryl group to form a stable phosphoserine linkage between the substrate protein and PR-Ub. The protonated E340 then functions as a general acid to protonate H277, thereby regenerating the enzyme to its initial state. Alternatively, if a water molecule serves as the Ub acceptor in the second step, the reaction results in the cleavage of ADPR-Ub to PR-Ub.

## 4.4 Discussion

To date, modification of Ub to yield PR-Ub has not been reported in (non-infected) eukaryotes. However, many *Legionella* effector proteins have eukaryotic origins evolutionarily [101], raising the possibility that eukaryotes also harbor an equivalent machinery that may be encoded in multiple polypeptides, since the mART and PDE activities are functionally independent. Future elucidation of such a eukaryotic enzyme system is of fundamental importance to our understanding of the versatile Ub code.

## 4.5 Materials and Methods

**4.5.1 Cloning and mutagenesis.** DNA fragments encoding the SdeA-Core (a.a. 211-910) and SdeD (a.a. 1-341) were amplified from *L. pneumophila* genomic DNA. The PCR products were digested with BamHI and XhoI restriction enzymes and inserted into a pET28a-based vector in frame with an N-terminal 6xHis-SUMO tag for protein overexpression in bacteria cells. Amino

acid substitutions of SdeA and SdeD were introduced by site-directed mutagenesis using oligonucleotide primer pairs containing the appropriate base changes. The Ub gene was subcloned into a pET21a vector. All constructs were confirmed by DNA sequencing.

**4.5.2 Protein expression and purification.** Relevant plasmids (containing *Legionella* protein constructs or Rab33b) were transformed into E. coli BL21(DE3) cells. Cultures derived from single colonies were grown in Luria-Bertani medium supplemented with 50 μg/ml kanamycin or 100 μg/ml ampicillin to mid-log phase. Protein expression was induced with 0.1 mM isopropyl-B-D-thiogalactopyranoside (IPTG) for 12 h at 18°C. Harvested cells were resuspended in a lysis buffer containing 20 mM Tris-HCl (pH 8.0) and 150 mM NaCl and were lysed by sonication. Insoluble cellular debris was pelleted by centrifugation at 31,000xg for 30 min at 4°C, and the clarified lysate was incubated with cobalt resin (Gold-Bio) for 1.5 h at 4°C. Proteins bound to the resin were extensively washed with lysis buffer. The SUMO-specific protease Ulp1 was then added to the resin slurry to release the expressed protein from the His-SUMO tag. Eluted protein samples were further purified by FPLC size exclusion chromatography (Superdex 16/60, GE Lifesciences) in 150 mM NaCl, 20 mM Tris pH 7.5. Peak fractions were collected, pooled, and concentrated. Protocols for Ub expression and purification were adapted from the published literature [102]. Briefly, harvested cells were resuspended in 20 mM ammonium acetate, pH 5.1. Cells were lysed by sonication and cell lysate was clarified by centrifugation (31,000xg for 30 min). The pH of the clarified lysate was lowered to 4.8 using glacial acetic acid. The decrease in pH caused the lysate to turn milky-white (a result of precipitated proteins), and the solution was again centrifuged at 31,000xg for 30 min at 4°C to remove the precipitated protein fraction. The pH of the remaining soluble fraction was adjusted to 5.1 by the addition of NaOH. The soluble

fraction was then loaded onto a HiTrap SP cation exchange column (GE Healthcare) in 20 mM ammonium acetate pH 5.1 and eluted in a continuous gradient of 500 mM ammonium acetate pH 5.1. Fractions containing the ubiquitin peak were pooled and further purified with a size exclusion chromatography in 150 mM NaCl, 20 mM Tris pH 7.5. Ubiquitin-containing fractions were pooled and concentrated.

To generate ADPR-Ub for both biochemical assays and crystallographic trials, 1 μM of SdeA-Core (211-910) H277A (which lacks PDE activity) was incubated with 25 μM Ub and 1 mM $NAD^+$ for 1 h at 37°C. ADPR-Ub was purified by size exclusion chromatography in 150 mM NaCl, 20 mM Tris (pH 7.5).

**4.5.3 Protein crystallization.** Generally, all protein crystallization screens were performed with a Crystal Phoenix liquid handling robot (Art Robbins Instruments) at room temperature. The crystallization conditions, which yielded the initial crystals from the screen, were further optimized by the hanging-drop vapor diffusion method by mixing 1.5 μl of protein with an equal volume of reservoir solution.

Specifically, for SdeA-Core crystallization, SdeA-Core protein was concentrated to 12 mg/ml and crystallized in 100 mM HEPES pH 7.9, 12% PEG 8000. Thin-plate shaped crystals appeared in about two weeks. For SdeD crystallization, SdeD was concentrated to 14 mg/ml and crystallized in 200 mM $CaCl_2$, 100 mM MES pH 5.5, 18% PEG 6000, and 100 mM DTT. Cube shaped crystals formed within 2-3 days. To generate the SdeD-Ub crystals, SdeD (1-341) was mixed with WT-Ub at a 1:5 molar ratio, with a final SdeD concentration of 8 mg/ml. Rod-shaped crystals formed in 200 mM NaCl, 100 mM imidazole pH 7.0, and 24% PEG 8000.

We also obtained crystals of a catalytically inactive SdeD H67A mutant with purified ADPR-Ub. However, those crystals diffracted poorly (likely due to conflicting crystal packing contacts mediated by the ADPR moiety at the Ub2 site). We therefore attempted to crystallize the SdeD PDE domain with a mixture of ADPR-Ub and unmodified Ub in a 1:2:3 molar ratio, with a final SdeD concentration of 12 mg/ml. We expected ADPR-Ub to have a higher affinity for binding at the Ub1 site, allowing unmodified Ub to bind to the Ub2 site to satisfy crystal packing constraints. Rod-shaped crystals appeared in one day in a solution containing 100 mM sodium cacodylate pH 6.7 and 21% PEG 8000. This strategy yielded diffraction quality crystals in which ADPR-Ub is bound at the Ub1 site and unmodified Ub bound at the Ub2 site.

**4.5.4 X-ray diffraction data collection and processing.** Diffraction data sets for SdeA-Core, the SdeD-Ub complex, and the SdeD/Ub/ADPR-Ub complex were collected at Cornell synchrotron light source MacCHESS beamline F1, while data sets for SdeD crystals were collected at the A1 beamline. Before data collection, all crystals were soaked in cryoprotectant solutions containing their respective crystallization condition buffer supplemented with 20% glycerol and flash frozen in stream of liquid nitrogen. All data sets were indexed, integrated and scaled with HKL-2000 [73].

**4.5.5 Structure determination and refinement.** The structure of SdeA-Core was solved by the single wavelength anomalous dispersion (SAD) method. Before data collection, SdeA-Core crystals were soaked in cryoprotectant (0.1 M HEPES pH 7.9, 12% PEG 8000, and 25% (v/v) glycerol) with the addition of 10 mM ethylmercury chloride for 5 min at room temperature. Heavy atom sites were determined and the initial phase was calculated using the program HKL2MAP

[103]. The structure of the PDE domain of SdeD was solved by SAD phasing with selenomethionine-incorporated SdeD crystals. The structures of the SdeD-Ub and SdeD/Ub/ADPR-Ub complexes were solved by molecular replacement with the AMoRe program [74] of the CCP4 suite [75], using the apo SdeD structure as the search model. For all data sets, iterative cycles of model building and refinement were carried out with Coot[76] and refmac5 [77] of the CCP4 suite.

**4.5.6 NMR titration analysis.** All NMR spectra were collected on a Bruker 500 MHz DMX at 25°C. Data were processed using NMRPipe [104] and analyzed using NMRViewJ [105]. NMR samples were prepared in 25mM NaPi, 150mM NaCl buffer at pH 7.0 with 5% (v/v) D2O. For all NMR experiments, the concentration of 15N-Ub or Ub-ADPR, in which only the Ub-subunit was isotopically labeled, was maintained at 150 μM. Concentrations of other protein components varied from 35-300 μM.

**4.5.7 SAXS data collection.** SAXS experiments were performed on beamline 4-2 at the Stanford Synchrotron Radiation Lightsource (SSRL) [106]. Concentrated SdeA-Core (a.a. 211-910) protein samples were buffer exchanged into 20 mM HEPES pH 7.5, 150 mM NaCl, and stored at 4 °C before data collection. Fifty microliters of SdeA-Core (7 mg/mL) were injected onto a Superdex 200 Increase PC 3.2/30 (GE) column in buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM DTT, 0.02% $NaN_3$, with a flow rate of 0.05 mL/min for online SEC-SAXS. Data were collected using a Pilatus3 X 1M detector with a 2.5 m sample-to-detector distance and X-ray beam energy of 12.4 keV (wavelength, $\lambda$ = 1Å), with 1 sec exposures collected every 5 sec. The first 100 images were averaged as buffer scattering data and subtracted from the corresponding protein

scattering data. SAXS patterns, the radius of gyration ($R_g$), the maximal particle dimension ($D_{max}$), and the pairwise distance distribution histogram [*P(r)* plot] and Kratky plot were analyzed by using the ATSAS software suite [107]. The AllosMod-FOXS server was used for the comparison of solution and X-ray structure conformations [108, 109]. The X-ray determined 'open' structure and modeled 'closed' conformation were used as input structures. AllosMod generated one hundred static structures, using MODELLER[110], which were similar to the input X-ray determined (open) or modeled (closed) structures of SdeA-Core [109]. Theoretical SAXS profiles were calculated and compared against the raw SAXS data using FOXS rigid-body modeling with a maximal q value of 0.25 [108]. The average and standard deviation in $\chi^2$ among the five best-fitting models were examined for fit comparisons.

**4.5.8 Computational analysis and graphic presentation of protein sequence and structure.** Sequences homologous to SdeA were selected from results generated by the BLAST server (NCBI). Edited sequences were aligned with Clustal Omega [80] and colored by the Multiple Align Show online server (http://www.bioinformatics.org/sms/index.html). Protein surface conservation was calculated by the online ConSurf server (http://consurf.tau.ac.il) [68]. All structural figures were generated using PyMOL (The PyMOL Molecular Graphics System, Version 1.8.X, Schrödinger, LLC) except for the difference Fourier electron density map figure (Extended Data Fig. 9e), which was generated in Coot. The electrostatic surface potential is calculated with the program APBS (http://www.poissonboltzmann.org). The surface is colored based on electrostatic potential with positively charged regions in blue (+4 kcal/electron) and negatively charged surfaces in red (-4 kcal/electron).

**4.5.9 Ubiquitin modification and Rab33b ubiquitination assays.** Ub modification reactions were carried out by mixing 1 μM of SdeA-Core (a.a. 211-910) or SdeA-mART (563-910) with 25 μM ubiquitin in a reaction buffer containing 50 mM NaCl and 50 mM Tris pH 7.5, in the presence or absence of 1 mM $NAD^+$. The reactions were incubated for 1h at 37°C and reaction products were assessed by both 8% native PAGE and 12% SDS-PAGE. Native gels were stained with Coomassie, while SDS-PAGE gels were stained with Pro-Q Diamond phosphoprotein stain (Invitrogen) to assay for PDE activity. ADPR-Ub and PR-Ub migrate at the same position on a native gel (labeled as modified Ub), however, only PR-Ub is visible by Pro-Q phosphoprotein stain due to its free phosphoryl group [111]. Rab33b ubiquitination reactions were performed by the addition of 4 μM of recombinant Flag-Rab33b to the Ub modification reaction described above. The reaction products were analyzed by SDS-PAGE followed by Western blot using an anti-Flag antibody (Sigma-Aldrich) at a 1:2500 dilution. To perform the intracellular PR-ubiquitination assay of Rab33b, plasmids expressing Flag-Rab33b, GFP alone or the indicated GFP-tagged SdeA were co-transfected in NIH HEK293T cells. Whole cell lysates were subjected to immunoprecipitation with Flag beads and the products were analyzed by anti-Flag Western blot. The expression of GFP-SdeA constructs was analyzed by anti-GFP Western blot.

**4.5.10 Data availability.** Atomic coordinates and structure factors for the reported structures have been deposited into the Protein Data Bank under the accession codes 6B7Q (Hg-bound SdeA), 6B7P (Se-SdeD), 6B7M (SdeD•Ub), and 6B7O (SdeD•Ub•ADPR-Ub)

# CHAPTER 5

# CONCLUDING REMARKS AND FUTURE RESEARCH DIRECTIONS

## 5.1 Summary

The work presented in this thesis has revealed mechanistic details of two ubiquitination pathways in *Legionella*. The complex structures of SidC/SdcA have given mechanistic insight into how SidC recognizes an E2 and facilitates the transfer of ubiquitin to the catalytic cysteine of the SNL domain. The structural studies of SdeA have elucidated a detailed reaction mechanism for the second step in the PR-ubiquitination reaction. This work has laid a foundation for future studies to uncover more details of these pathways, learn more about host trafficking, and to potentially impact future therapeutics.

## 5.2 SidC ubiquitination

We have demonstrated how SidC recognizes E2s and subsequently brings ubiquitin to the active site, holding it in place and ready for the transfer to a substrate. A major research focus for the future of this project is the identification of SidC substrates. The best approach to this problem is to perform a SILAC experiment of eukaryotic cells infected with either WT (LP02) or ΔsidC/sdcA *Legionella* strains. Since active SidC is anchored to the PI(4)P enriched on the LCV, it is likely that the targets of SidC ubiquitination are also localized to the LCV. Therefore, rather than looking for the difference in ubiquitination across the whole cell (which could disguise any small changes in ubiquitination), differences should be looked for only in LCV-associated proteins. The infected cells would be lysed and LCVs isolated using a sucrose density gradient ultracentrifugation procedure [112, 113]. After trypsin digestion, ubiquitinated peptides would be pulled down using immunobeads that detect the c-terminal glycine attached to a lysine (the K-ε-

GG remnant motif). A potential pitfall in this experiment is that SidC is very promiscuous in vitro and will ubiquitinate a wide range of proteins including E2s, ubiquitin, and GFP. To date, a protocol has not been developed to differentiate between "real" targets and artifacts. Thus, the major challenge is in validating the SILAC results. In the event that no differences are observed in ubiquitination on the LCVs, a more exacting subcellular fractionation (to isolate Golgi and ER membranes) could be carried out. A recent study suggested that prior to the recruitment of ER-derived vesicles to form the LCV, *Legionella* may begin the formation of the vacuole using tubular ER associated with the plasma membrane, before subsequently recruiting ER-derived vesicles [114]. If this proves to be true, performing a subcellular fractionation experiment and looking for changes in ubiquitination of tubular ER early in the infection may provide SidC/SdcA ubiquitination substrates.

If the challenge of substrate validation is overcome and substrates can be confirmed, this line of inquiry could form the basis for a new understanding of eukaryotic membrane trafficking. A remaining question is how ER-derived vesicles are recruited to the LCV. A ΔsidC-sdcA *Legionella* strain shows impaired recruitment of ER-derived vesicles in the early stages of infection, indicating that the substrates of SidC ubiquitination likely play a role in this recruitment [25]. By using *Legionella* as a tool to disrupt host membrane trafficking, it may be possible to learn more about the host pathway itself.

Once a substrate is determined, the next mechanistic question to be answered is how SidC transfers ubiquitin to a substrate. An in vitro reaction using SidC with a C46K mutation generates ubiquitin that is locked at the active site. To this end, a covalent SidC C46K-Ub K6 binary complex can be produced. A substrate could then be mixed with SidC-Ub in an attempt to form a complex for subsequent crystallographic experiments. Such a project would determine the

interface on SidC used to interact with substrates, and the position of the dynamic catalytic histidine and aspartate residues could deliver more insights into how this step of catalysis is carried out.

### 5.3 Phosphoribosyl ubiquitination

Since the discovery of phosphoribosyl ubiquitination in 2016, a quick succession of papers has determined the structure, mechanism, and potential substrates of this unique *Legionella* enzyme family. While much has been reported in the recent publications on this protein family, there are still several unresolved questions that will be of great interest to the field.

Although the DUB, mART, and PDE domains have been characterized both structurally and functionally, the C-terminal coiled-coil domain has remained largely uncharacterized. Despite playing a role in stabilizing the α-helical lobe of the mART domain in close proximity to the conserved active site, it was shown to be inessential for the PR-ubiquitination activity of the SidE family. While it was reported that the coiled-coil domain played a role in the dimerization of SidE, a longer structure of SdeA was found to be monomeric [29, 115]. The coiled-coil domain is involved in binding the adaptor protein complex IcmS-IcmW, two cytosolic components of the Dot/Icm secretion machinery [29, 116]. This suggests that the coiled-coil domain could be involved in the translocation of SdeA into host cells. Determining the function of the C-terminal domain will complete the characterization of this protein family.

Another major question to be addressed is what proteins are targeted by the SidE family for PR-ubiquitination. The Rab small GTPases were first identified as targets almost serendipitously, due to the reasoning that this class of proteins is a common target for *Legionella* effectors [26]. Since then RTN4, an ER-localized protein responsible for generating ER curvature

and regulating ER tubule formation, has also been reported to be a substrate [86]. Aside from the obvious fact that a serine residue must be present to be modified, there has been no identified substrate motif, although it has been suggested that the serine should be surrounded by hydrophobic residues to better interact with the catalytic cleft of the PDE domain [98]. Arguably the most promising method to identify further substrates is to perform SILAC mass spectrometry experiments with wild-type and SidE family deletion strains. Once substrates are identified, a role for the SidE family of effectors in an infection can be defined.

One question of active research is that of how the activity of the SidE family proteins are regulated. The protein SidJ was shown to suppress the toxicity of SidE family members in a yeast model, leading to SidJ to be referred to as a "metaeffector" (an effector that regulates other effectors) [31]. Recently, it was reported that SidJ could remove SidE protein from the LCV in vitro [117]. Also, SidJ was reported to be a deubiquitinase; however, current work in the lab has brought this finding into question [32]. We have preliminary data indicating that *Legionella* encodes for at least two PR-Ub deubiquitinases, SdeD and SdeE. These proteins are small members of the SidE family which only contain PDE domains. Current work is ongoing to determine the structural differences which lend these effectors DUB activity. Clearly, there is still much to be learned about the regulation of the SidE family.

Perhaps the most intriguing question in the field is whether PR-ubiquitination is found in eukaryotes. *Legionella* has generated its effector proteins via horizontal gene transfer from its hosts [118]. Eukaryotes contain arginine-specific mono-ADP-ribosyltransferases, which are involved in many cellular processes, as well as phosphodiesterases, although none shown to perform a ligation reaction involving ubiquitin [119, 120]. To date, bioinformatics analyses have not yielded a eukaryotic homolog to the SidE family. PR-ubiquitination has yet to be detected in

eukaryotes, but if PR-Ub is in low abundance in the cell, its presence could be masked by the overwhelming amount of WT-Ub and a lack of antibodies to detect and pulldown PR-ubiquitin. One approach to detect PR-ubiquitination would be to add ADPR-Ub to eukaryotic cell lysates and look for PDE activity. This approach runs the risk of false-positive discoveries of PR-ubiquitination activity, since adding an excess of ADPR-ubiquitin could force eukaryotic PDEs to carry out a non-natural enzymatic activity. It is possible that PR-ubiquitination could be carried out by a complex of proteins, which would further complicate the discovery of this pathway. Alternatively, such a pathway could be upregulated during a particular condition, such as nutrient starvation or oxidative stress. In this case, identification would require the extensive testing of culture conditions. The discovery of a eukaryotic PR-ubiquitination pathway would lead to a more complete understanding of ubiquitination and could inform therapeutic studies.

REFERENCES


1. Komander, D. and M. Rape, *The ubiquitin code*. Annual review of biochemistry, 2012. **81**: p. 203-229.
2. Varshavsky, A., *The ubiquitin system, an immense realm*. Annual review of biochemistry, 2012. **81**: p. 167-176.
3. Pickart, C.M. and M.J. Eddins, *Ubiquitin: structures, functions, mechanisms*. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2004. **1695**(1): p. 55-72.
4. Metzger, M.B., V.A. Hristova, and A.M. Weissman, *HECT and RING finger families of E3 ubiquitin ligases at a glance*. Journal of cell science, 2012. **125**(3): p. 531-537.
5. Rotin, D. and S. Kumar, *Physiological functions of the HECT family of ubiquitin ligases*. Nature reviews Molecular cell biology, 2009. **10**(6): p. 398-409.
6. Berndsen, C.E. and C. Wolberger, *New insights into ubiquitin E3 ligase mechanism*. Nat Struct Mol Biol, 2014. **21**(4): p. 301-7.
7. Huibregtse, J. and J.R. Rohde, *Hell's BELs: B acterial E 3 L igases That Exploit the Eukaryotic Ubiquitin Machinery*. 2014.
8. Alomairi, J., et al., *Alterations of host cell ubiquitination machinery by pathogenic bacteria*. Frontiers in cellular and infection microbiology, 2015. **5**.
9. Hicks, S.W. and J.E. Galán, *Hijacking the host ubiquitin pathway: structural strategies of bacterial E3 ubiquitin ligases*. Current opinion in microbiology, 2010. **13**(1): p. 41-46.
10. Di Paolo, G. and P. De Camilli, *Phosphoinositides in cell regulation and membrane dynamics*. Nature, 2006. **443**(7112): p. 651-657.
11. Balla, T., *Phosphoinositides: tiny lipids with giant impact on cell regulation*. Physiological reviews, 2013. **93**(3): p. 1019-1137.
12. Pizarro-Cerdá, J. and P. Cossart, *Subversion of phosphoinositide metabolism by intracellular bacterial pathogens*. Nature cell biology, 2004. **6**(11): p. 1026-1033.
13. Hilbi, H., S. Weber, and I. Finsel, *Anchors for effectors: subversion of phosphoinositide lipids by Legionella*. Frontiers in microbiology, 2011. **2**.
14. Hilbi, H., *Modulation of phosphoinositide metabolism by pathogenic bacteria*. Cellular microbiology, 2006. **8**(11): p. 1697-1706.

15. Weber, S.S., C. Ragaz, and H. Hilbi, *Pathogen trafficking pathways and host phosphoinositide metabolism*. Molecular microbiology, 2009. **71**(6): p. 1341-1352.
16. Hsu, F., et al., *Structural basis for substrate recognition by a unique Legionella phosphoinositide phosphatase*. Proceedings of the National Academy of Sciences, 2012. **109**(34): p. 13567-13572.
17. Toulabi, L., et al., *Identification and structural characterization of a Legionella phosphoinositide phosphatase*. Journal of Biological Chemistry, 2013. **288**(34): p. 24518-24527.
18. Cunha, B.A., A. Burillo, and E. Bouza, *Legionnaires' disease*. Lancet, 2015.
19. Hubber, A. and C.R. Roy, *Modulation of host cell function by Legionella pneumophila type IV effectors*. Annual review of cell and developmental biology, 2010. **26**: p. 261-283.
20. De Felipe, K.S., et al., *Legionella eukaryotic-like type IV substrates interfere with organelle trafficking*. PLoS Pathog, 2008. **4**(8): p. e1000117.
21. Weber, S.S., et al., *Legionella pneumophila exploits PI (4) P to anchor secreted effector proteins to the replicative vacuole*. PLoS Pathog, 2006. **2**(5): p. e46.
22. Ragaz, C., et al., *The Legionella pneumophila phosphatidylinositol-4 phosphate-binding type IV substrate SidC recruits endoplasmic reticulum vesicles to a replication-permissive vacuole*. Cellular microbiology, 2008. **10**(12): p. 2416-2433.
23. Horenkamp, F.A., et al., *Legionella pneumophila subversion of host vesicular transport by SidC effector proteins*. Traffic, 2014. **15**(5): p. 488-499.
24. Gazdag, E.M., et al., *The structure of the N-terminal domain of the Legionella protein SidC*. Journal of structural biology, 2014. **186**(1): p. 188-194.
25. Hsu, F., et al., *The Legionella effector SidC defines a unique family of ubiquitin ligases important for bacterial phagosomal remodeling*. Proceedings of the National Academy of Sciences, 2014. **111**(29): p. 10538-10543.
26. Qiu, J., et al., *Ubiquitination independent of E1 and E2 enzymes by bacterial effectors*. Nature, 2016. **533**(7601): p. 120-4.
27. Bhogaraju, S., et al., *Phosphoribosylation of Ubiquitin Promotes Serine Ubiquitination and Impairs Conventional Ubiquitination*. Cell, 2016. **167**(6): p. 1636-1649 e13.

28. Sheedlo, M.J., et al., *Structural basis of substrate recognition by a bacterial deubiquitinase important for dynamics of phagosome ubiquitination*. Proc Natl Acad Sci U S A, 2015. **112**(49): p. 15090-5.
29. Dong, Y., et al., *Structual basis of ubiquitin modification by the Legionella effector SdeA*. Nature, 2018. **557**: p. 674-678.
30. Puvar, K., et al., *Ubiquitin Chains Modified by the Bacterial Ligase SdeA Are Protected from Deubiquitinase Hydrolysis*. Biochemistry, 2017. **56**(36): p. 4762-4766.
31. Havey, J.C. and C.R. Roy, *Toxicity and SidJ-Mediated Suppression of Toxicity Require Distinct Regions in the SidE Family of Legionella pneumophila Effectors*. Infect Immun, 2015. **83**(9): p. 3506-14.
32. Qiu, J., et al., *A unique deubiquitinase that deconjugates phosphoribosyl-linked protein ubiquitination*. Cell Res, 2017. **27**(7): p. 865-881.
33. Luo, X., et al., *Structure of the Legionella Virulence Factor, SidC Reveals a Unique PI (4) P-Specific Binding Domain Essential for Its Targeting to the Bacterial Phagosome*. PLoS Pathog, 2015. **11**(6): p. e1004965.
34. Dolinsky, S., et al., *The Legionella longbeachae Icm/Dot substrate SidC selectively binds phosphatidylinositol 4-phosphate with nanomolar affinity and promotes pathogen vacuole-endoplasmic reticulum interactions*. Infection and immunity, 2014. **82**(10): p. 4021-4033.
35. Lin, D.Y., J. Diao, and J. Chen, *Crystal structures of two bacterial HECT-like E3 ligases in complex with a human E2 reveal atomic details of pathogen-host interactions*. Proc Natl Acad Sci U S A, 2012. **109**(6): p. 1925-30.
36. Verdecia, M.A., et al., *Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase*. Mol Cell, 2003. **11**(1): p. 249-59.
37. Kamadurai, H.B., et al., *Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex*. Mol Cell, 2009. **36**(6): p. 1095-102.
38. Rowbotham, T.J., *Preliminary report on the pathogenicity of Legionella pneumophila for freshwater and soil amoebae*. J Clin Pathol, 1980. **33**(12): p. 1179-83.
39. Cunha, B.A., A. Burillo, and E. Bouza, *Legionnaires' disease*. Lancet, 2016. **387**(10016): p. 376-385.

40. Qiu, J. and Z.Q. Luo, *Legionella and Coxiella effectors: strength in diversity and activity*. Nat Rev Microbiol, 2017.
41. Ensminger, A.W., *Legionella pneumophila, armed to the hilt: justifying the largest arsenal of effectors in the bacterial world*. Curr Opin Microbiol, 2016. **29**: p. 74-80.
42. Isberg, R.R., T.J. O'Connor, and M. Heidtman, *The Legionella pneumophila replication vacuole: making a cosy niche inside host cells*. Nat Rev Microbiol, 2009. **7**(1): p. 13-24.
43. Swanson, M.S. and R.R. Isberg, *Association of Legionella pneumophila with the macrophage endoplasmic reticulum*. Infect Immun, 1995. **63**(9): p. 3609-20.
44. Tilney, L.G., et al., *How the parasitic bacterium Legionella pneumophila modifies its phagosome and transforms it into rough ER: implications for conversion of plasma membrane to the ER membrane*. J Cell Sci, 2001. **114**(Pt 24): p. 4637-50.
45. Hubber, A. and C.R. Roy, *Modulation of host cell function by Legionella pneumophila type IV effectors*. Annu Rev Cell Dev Biol, 2010. **26**: p. 261-83.
46. Hubber, A., T. Kubori, and H. Nagai, *Modulation of the Ubiquitination Machinery by Legionella*. Curr Top Microbiol Immunol, 2013.
47. Lin, Y.H. and M.P. Machner, *Exploitation of the host cell ubiquitin machinery by microbial effector proteins*. J Cell Sci, 2017. **130**(12): p. 1985-1996.
48. Hershko, A. and A. Ciechanover, *The ubiquitin system*. Annu Rev Biochem, 1998. **67**: p. 425-79.
49. Deshaies, R.J. and C.A. Joazeiro, *RING domain E3 ubiquitin ligases*. Annu Rev Biochem, 2009. **78**: p. 399-434.
50. Aguilera, M., et al., *Ariadne-1: a vital Drosophila gene is required in development and defines a new conserved family of ring-finger proteins*. Genetics, 2000. **155**(3): p. 1231-44.
51. Huibregtse, J.M., et al., *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): p. 2563-7.
52. Rotin, D. and S. Kumar, *Physiological functions of the HECT family of ubiquitin ligases*. Nat Rev Mol Cell Biol, 2009. **10**(6): p. 398-409.
53. Dou, H., et al., *BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer*. Nat Struct Mol Biol, 2012. **19**(9): p. 876-83.

54. Plechanovova, A., et al., *Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis*. Nature, 2012. **489**(7414): p. 115-20.
55. Pruneda, J.N., et al., *Structure of an E3:E2~Ub complex reveals an allosteric mechanism shared among RING/U-box ligases*. Mol Cell, 2012. **47**(6): p. 933-42.
56. Wenzel, D.M., et al., *UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids*. Nature, 2011. **474**(7349): p. 105-8.
57. Dove, K.K., et al., *Molecular insights into RBR E3 ligase ubiquitin transfer mechanisms*. EMBO Rep, 2016. **17**(8): p. 1221-35.
58. Huibregtse, J. and J.R. Rohde, *Hell's BELs: bacterial E3 ligases that exploit the eukaryotic ubiquitin machinery*. PLoS Pathog, 2014. **10**(8): p. e1004255.
59. Ashida, H., M. Kim, and C. Sasakawa, *Exploitation of the host ubiquitin system by human bacterial pathogens*. Nat Rev Microbiol, 2014. **12**(6): p. 399-413.
60. Kamadurai, H.B., et al., *Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3*. Elife, 2013. **2**: p. e00828.
61. Hsu, F., et al., *The Legionella effector SidC defines a unique family of ubiquitin ligases important for bacterial phagosomal remodeling*. Proc Natl Acad Sci U S A, 2014. **111**(29): p. 10538-43.
62. Luo, X., et al., *Structure of the Legionella Virulence Factor, SidC Reveals a Unique PI(4)P-Specific Binding Domain Essential for Its Targeting to the Bacterial Phagosome*. PLoS Pathog, 2015. **11**(6): p. e1004965.
63. Ragaz, C., et al., *The Legionella pneumophila phosphatidylinositol-4 phosphate-binding type IV substrate SidC recruits endoplasmic reticulum vesicles to a replication-permissive vacuole*. Cell Microbiol, 2008. **10**(12): p. 2416-33.
64. Huang, Q., et al., *Reduction of lattice disorder in protein crystals by high-pressure cryocooling*. J Appl Crystallogr, 2016. **49**(Pt 1): p. 149-157.
65. Huang, L., et al., *Structure of an E6AP-UbcH7 complex: insights into ubiquitination by the E2-E3 enzyme cascade*. Science, 1999. **286**(5443): p. 1321-6.
66. Nuber, U. and M. Scheffner, *Identification of determinants in E2 ubiquitin-conjugating enzymes required for hect E3 ubiquitin-protein ligase interaction*. J Biol Chem, 1999. **274**(11): p. 7576-82.

67. Maspero, E., et al., *Structure of a ubiquitin-loaded HECT ligase reveals the molecular basis for catalytic priming*. Nat Struct Mol Biol, 2013. **20**(6): p. 696-701.
68. Ashkenazy, H., et al., *ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules*. Nucleic Acids Res, 2016. **44**(W1): p. W344-50.
69. Stieglitz, B., et al., *Structural basis for ligase-specific conjugation of linear ubiquitin chains by HOIP*. Nature, 2013. **503**(7476): p. 422-426.
70. Berndsen, C.E., et al., *A conserved asparagine has a structural role in ubiquitin-conjugating enzymes*. Nat Chem Biol, 2013. **9**(3): p. 154-6.
71. Yunus, A.A. and C.D. Lima, *Lysine activation and functional analysis of E2-mediated conjugation in the SUMO pathway*. Nat Struct Mol Biol, 2006. **13**(6): p. 491-9.
72. Kim, C.U., R. Kapfer, and S.M. Gruner, *High-pressure cooling of protein crystals without cryoprotectants*. Acta Crystallogr D Biol Crystallogr, 2005. **61**(Pt 7): p. 881-90.
73. Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode*. Methods Enzymol., 1997. **276**: p. 307-326.
74. Trapani, S. and J. Navaza, *AMoRe: classical and modern*. Acta Crystallogr D Biol Crystallogr, 2008. **64**(Pt 1): p. 11-6.
75. Collaborative Computational Project, N., *The CCP4 suite: programs for protein crystallography*. Acta Cryst., 1994. **D**(50): p. 760-763.
76. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
77. Murshudov, G.N., A.A. Vagin, and E.J. Dodson, *Refinement of macromolecular structures by the maximum-likelihood method*. Acta Crystallogr D Biol Crystallogr, 1997. **53**(Pt 3): p. 240-55.
78. Eisenberg, D., et al., *Analysis of membrane and surface protein sequences with the hydrophobic moment plot*. J Mol Biol, 1984. **179**(1): p. 125-42.
79. Baker, N.A., et al., *Electrostatics of nanosystems: application to microtubules and the ribosome*. Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10037-41.
80. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega*. Mol Syst Biol, 2011. **7**: p. 539.

81. Hershko, A., A. Ciechanover, and A. Varshavsky, *Basic Medical Research Award. The ubiquitin system*. Nat Med, 2000. **6**(10): p. 1073-81.
82. Chen, Z.J. and L.J. Sun, *Nonproteolytic functions of ubiquitin in cell signaling*. Mol Cell, 2009. **33**(3): p. 275-86.
83. Hurley, J.H. and H. Stenmark, *Molecular mechanisms of ubiquitin-dependent membrane traffic*. Annu Rev Biophys, 2011. **40**: p. 119-42.
84. Haglund, K. and I. Dikic, *The role of ubiquitylation in receptor endocytosis and endosomal sorting*. J Cell Sci, 2012. **125**(Pt 2): p. 265-75.
85. Komander, D. and M. Rape, *The ubiquitin code*. Annu Rev Biochem, 2012. **81**: p. 203-29.
86. Kotewicz, K.M., et al., *A Single Legionella Effector Catalyzes a Multistep Ubiquitination Pathway to Rearrange Tubular Endoplasmic Reticulum for Replication*. Cell Host Microbe, 2017. **21**(2): p. 169-181.
87. Luo, Z.Q. and R.R. Isberg, *Multiple substrates of the Legionella pneumophila Dot/Icm system identified by interbacterial protein transfer*. Proc Natl Acad Sci U S A, 2004. **101**(3): p. 841-6.
88. Zhou, Y. and Y. Zhu, *Diversity of bacterial manipulation of the host ubiquitin pathways*. Cell Microbiol, 2015. **17**(1): p. 26-34.
89. Price, C.T., et al., *Molecular mimicry by an F-box effector of Legionella pneumophila hijacks a conserved polyubiquitination machinery within macrophages and protozoa*. PLoS Pathog, 2009. **5**(12): p. e1000704.
90. Kubori, T., A. Hyakutake, and H. Nagai, *Legionella translocates an E3 ubiquitin ligase that has multiple U-boxes with distinct functions*. Mol Microbiol, 2008. **67**(6): p. 1307-19.
91. Kubori, T., et al., *Legionella metaeffector exploits host proteasome to temporally regulate cognate effector*. PLoS Pathog, 2010. **6**(12): p. e1001216.
92. Ensminger, A.W. and R.R. Isberg, *E3 ubiquitin ligase activity and targeting of BAT3 by multiple Legionella pneumophila translocated substrates*. Infect Immun, 2010. **78**(9): p. 3905-19.
93. Wong, K., et al., *Structure of the Legionella Effector, lpg1496, Suggests a Role in Nucleotide Metabolism*. J Biol Chem, 2015. **290**(41): p. 24727-37.

94. Jeong, B.R., et al., *Structure function analysis of an ADP-ribosyltransferase type III effector and its RNA-binding target in plant immunity*. J Biol Chem, 2011. **286**(50): p. 43272-81.

95. Tsurumura, T., et al., *Arginine ADP-ribosylation mechanism based on structural snapshots of iota-toxin and actin complex*. Proc Natl Acad Sci U S A, 2013. **110**(11): p. 4267-72.

96. Simon, N.C., K. Aktories, and J.T. Barbieri, *Novel bacterial ADP-ribosylating toxins: structure and function*. Nat Rev Microbiol, 2014. **12**(9): p. 599-611.

97. Rinaldo, S., et al., *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, 2015. **197**(8): p. 1525-35.

98. Kalayil, S., et al., *Insights into catalysis and function of phosphoribosyl-linked serine ubiquitination*. Nature, 2018. **557**: p. 734-738.

99. Klumpp, S. and J. Krieglstein, *Phosphorylation and dephosphorylation of histidine residues in proteins*. Eur J Biochem, 2002. **269**(4): p. 1067-71.

100. Stock, A.M., V.L. Robinson, and P.N. Goudreau, *Two-component signal transduction*. Annu Rev Biochem, 2000. **69**: p. 183-215.

101. Burstein, D., et al., *Genomic analysis of 38 Legionella species identifies large and diverse effector repertoires*. Nat Genet, 2016. **48**(2): p. 167-175.

102. Raasi, S. and C.M. Pickart, *Ubiquitin chain synthesis*. Methods Mol Biol, 2005. **301**: p. 47-55.

103. Pape, T. and T.R. Schneider, *HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs*. J. Appl. Cryst., 2004. **37**: p. 843-844.

104. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J Biomol NMR, 1995. **6**(3): p. 277-93.

105. Johnson, B.A. and R.A. Blevins, *NMR View: A computer program for the visualization and analysis of NMR data*. J Biomol NMR, 1994. **4**(5): p. 603-14.

106. Smolsky, I.L., et al., *Biological small-angle x-ray scattering facility at the Stanford synchrotron radiation laboratory*. Journal of Applied Crystallography, 2007. **40**: p. S453-S458.

107. Franke, D., et al., *ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions*. J Appl Crystallogr, 2017. **50**(Pt 4): p. 1212-1225.

108. Schneidman-Duhovny, D., M. Hammel, and A. Sali, *FoXS: a web server for rapid computation and fitting of SAXS profiles*. Nucleic Acids Res, 2010. **38**(Web Server issue): p. W540-4.

109. Weinkam, P., J. Pons, and A. Sali, *Structure-based model of allostery predicts coupling between distant sites*. Proc Natl Acad Sci U S A, 2012. **109**(13): p. 4875-80.

110. Marti-Renom, M.A., et al., *Comparative protein structure modeling of genes and genomes*. Annu Rev Biophys Biomol Struct, 2000. **29**: p. 291-325.

111. Daniels, C.M., S.E. Ong, and A.K. Leung, *Phosphoproteomic approach to characterize protein mono- and poly(ADP-ribosyl)ation sites from cells*. J Proteome Res, 2014. **13**(8): p. 3510-22.

112. Hoffmann, C., I. Finsel, and H. Hilbi, *Purification of pathogen vacuoles from Legionella-infected phagocytes*. J Vis Exp, 2012(64).

113. Bruckert, W.M. and Y. Abu Kwaik, *Complete and ubiquitinated proteome of the Legionella-containing vacuole within human macrophages*. J Proteome Res, 2015. **14**(1): p. 236-48.

114. Haenssler, E., et al., *Endoplasmic Reticulum Tubule Protein Reticulon 4 Associates with the Legionella pneumophila Vacuole and with Translocated Substrate Ceg9*. Infect Immun, 2015. **83**(9): p. 3479-89.

115. Wang, Y., et al., *Structural Insights into Non-canonical Ubiquitination Catalyzed by SidE*. Cell, 2018. **173**(5): p. 1231-1243 e16.

116. Ninio, S., et al., *The Legionella IcmS-IcmW protein complex is important for Dot/Icm-mediated protein translocation*. Mol Microbiol, 2005. **55**(3): p. 912-26.

117. Jeong, K.C., J.A. Sexton, and J.P. Vogel, *Spatiotemporal regulation of a Legionella pneumophila T4SS substrate by the metaeffector SidJ*. PLoS Pathog, 2015. **11**(3): p. e1004695.

118. Gomez-Valero, L., et al., *Extensive recombination events and horizontal gene transfer shaped the Legionella pneumophila genomes*. BMC Genomics, 2011. **12**: p. 536.

119. Stevens, L.A. and J. Moss, *Mono-ADP-Ribosylation Catalyzed by Arginine-Specific ADP-Ribosyltransferases*. Methods Mol Biol, 2018. **1813**: p. 149-165.

120. Conti, M. and J. Beavo, *Biochemistry and physiology of cyclic nucleotide phosphodiesterases: essential components in cyclic nucleotide signaling*. Annu Rev Biochem, 2007. **76**: p. 481-511.