MODULATION OF UBIQUITINATION PATHWAY BY *LEGIONELLA* SidE EFFECTOR FAMILY

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

Anil Akturk

May 2018

MODULATION OF UBIQUITINATION PATHWAY BY *LEGIONELLA* SidE EFFECTOR FAMILY

ABSTRACT

Anil Akturk

Cornell University 2018

Legionella pneumophila manipulates a wide array of host cellular processes during infection; one that appears to be highly altered is the ubiquitination pathway. The SidE family of *Legionella* effector proteins has been shown to modify target proteins via a novel ubiquitination pathway, in which phosphoribosyl-ubiquitin is generated and subsequently used to ubiquitinate proteins. To elucidate the biochemical mechanism of this novel modification, we have solved the crystal structure of the catalytic core of SidE family member, SdeA. Using biochemical activity assays, we discovered an unpredicted domain that is crucial for the first step of this novel modification. To elucidate the molecular mechanism, we co-crystallized another SidE family member, SdeD, with both unmodified and modified ubiquitin molecules. These structural studies gave insights into the mechanism of ubiquitin modification during the second step of the reaction. We have also shown that two of the small SidE family members, SdeD and SdeF, remove the SdeA ubiquitin modification in an activity dependent manner, thus functioning as a deubiquitinase towards SdeA target proteins. These deubiquitinase proteins have also been shown to revert the

phenotype exerted by SdeA in host cells. Altogether, this work provides a mechanistic insight into a novel ubiquitination reaction catalyzed by SdeA, in addition to providing a preliminary model for regulation of this enzymatic activity.

BIOGRAPHICAL SKETCH

Anil Akturk grew up in Istanbul,Turkey. He first got interested in the depths of biology during his first year in high school with the “Biotechnology” topic. His interest in science and the promise of progress and prosperity in biological sciences (and his success in the National Entrance Exam) let him pursue a BSc in Biological Sciences and Bioengineering in Sabanci University, Turkey. He joined three different labs to obtain a taste of different methodologies in molecular biology. He did an internship in Karolinska Institute that made him sure he wants to follow an academic life in Biological Sciences. The things he learned during undergraduate studies were not enough, so he decided to earn a Master’s Degree in the same university. He joined Zehra Sayer’s lab and spent quite some time with biochemical studies on *Arabidopsis* G proteins. After successfully completing his Master’s Degree, he then chose to join the BMCB PhD program of Cornell University. He joined Yuxin Mao’s lab and worked on different *Legionella* effector proteins; but mainly on SidE protein family. As a graduate student, he was the president of Turkish Student Association for 2 years. Next, he is going to join Basile Tarchini’s group in Jackson Lab working on inner hair cell development in Bar Harbor, ME.

ACKNOWLEDGEMENTS

I would like to express my gratitude towards my advisor Dr. Yuxin Mao for his mentorship and patience he had for my training. He has guided me for five and a half years and supported my development in academic environment. His feedback and his suggestions, especially during our one-to-one meetings, has improved my ability to tackle important questions and problems in our research. He helped me to improve my communications and organization skills both within the lab and to the science community. I would also like to thank Dr. Fenghua Hu for her help and constructive criticism during our joint lab meetings.

I would like to thank my academic committee members Dr. Bill Brown and Dr. Linda Nicholson. They have helped me improve my scientific skills during our meetings and provided encouragement to continue being a part of the scientific community. I would specially like to thank Dr. Linda Nicholson for giving me an opportunity to teach side by side to her and be a part of her class.

Beyond my committee members, I would also like to thank other mentors that I had the chance to work with. I would specifically like to thank Jim Blankenship for his mentorship while I was a teaching assistant for his course for three semesters. I will always be grateful to be a part of his class and work with him.

I am, and always will be, grateful to Dr. Zehra Sayers for all her help and encouragement she had provided during my Master's studies. She gave me courage and the confidence to continue pursuing a career in biology.

My peers, my friends in my program and in the lab gave me the power to go through tougher times during my studies in Cornell. I owe them a big "Thank You!" for all the great times we had. I will always cherish the memories I had with all of them.

Finally, I would like to thank my wife, Nurten, for all the help and support she had provided me thorough all those years we have been together. Without her, I would not even dare to go abroad and leave my country behind. And last but not least, I would like to thank my family for all their sacrifices, supports and helps that let me be me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADPR	Adenine di phosphate ribosyl
AMP	Adenine mono phosphate
ATP	Adenosine tri phosphate
BEL	Bacterial E3 ligase
Co-IP	Co immunoprecipitation
Dot/Icm	Defective organelle trafficking / intracellular multiplication
DUB	Deubiquitinase
GTPase	Guanisine tri phosphatase
HECT	homologous to the E3AP C terminus
HEK293T	Human Embryonic Kidney 293T
LCV	Legionella Containing vacuole
mART	Mono-ADP-ribosyl transferase
MBP	Maltose Binding Protein
NAD	Nicotinamide adenine dinucleotide
NMR	Nuclear Magnetic Resonance
PAGE	Poly acrylamide gel
PDE	Phosphodiesterase
PI	Phosphoinositide
PR	Phosphoribose
RING	really interesting new gene
SAD	single wavelength anomalous dispersion
SAXS	Small Angle X-Ray Scattering
SDS	Sodium dodecyl sulfate
SidE	Substrate of Icm/Dot protein E
SNARE	Snap Receptor
Stx5	Syntaxin 5
T4BSS	Type 4B secretion system
Ub	Ubiquitin
WT	Wild Type

CHAPTER 1

LEGIONELLA SECRETED SidE EFFECTOR FAMILY and OTHER EFFECTORS DURING INFECTION

1.1 *Legionella pneumophila*

Legionella pneumophila is a Gram-negative, non-sporulated, flagellated bacterium [1]. It is non-capsulated and characterized to be an aerobic bacillus under the coccobacillus shape structure [2]. At least 70 different serovars of Legionella have been identified depending on where they were found and the antigens they have on their surfaces [3]. Free-living amoeba in aquatic environments such as ponds are the natural hosts of these bacteria. Ten amoeba species and two other ciliated species have been shown to be host cells for *Legionella* [4]. Advances in air-conditioning technologies, specifically the newer water-cooled air conditioning systems, introduce the possibility of contamination first with amoeba and then its pathogen *Legionella* [5]. When the *Legionella* contaminated aerosols are dispersed in the artificial atmosphere of automated buildings and inhaled by individuals, Legionnaire's disease can be contracted.

Legionnaire's disease is a result of Legionellosis. The other outcome of this Legionellosis can be Pontiac fever which is a mild flu-like disease [6]. Legionnaire's disease is a severe form of nonzoonotic atypical pneumonia that was first observed in a Legionnaire's Convention in Philadelphia in 1976 [7]. Even though the cure for the disease is as easy as using a cocktail of

antimicrobial drugs (quinolones, azithromycin, rifampin, tigecycline and TMP-SMX) [8], case fatality rate for the disease is around 10%. The elderly, children, and people with lung pathologies (i.e. smokers) have a higher chance of morbidity [9]. Legionnaire's Disease is a public health problem in developing countries but still up to 18 cases/million people can be encountered in developed countries such as the United States, Canada and the European Union [10]. Recent studies also show heart infection in the form of pericarditis, myocarditis and valvular infection by *Legionella* spp [11].

1.2 *Legionella* infection cycle

At the cellular level, *Legionella* infection in alveolar macrophages and amoebas happen in a similar fashion (Fig.1-1). The whole cycle starts with the attachment of *Legionella* to the plasma membrane of the host cell. RtxA and PilEL proteins are involved in the attachment and entry of the bacterium. PilEL is reported to be similar to pili forming proteins from other Gram-negative bacterium species, but the function of RtxA remains unknown [12]. The surface protein that *Legionella* uses for attachment differs according to the host cell species: in amoeba cells, the bacterium mainly exploits N-acetylglucosamine containing lectin [13]; in contrast, on the surface of alveolar macrophages, the attachment of *Legionella* is mediated by complement receptors CR1 (CD35) and CR3 (CD18/CD11b) [14]. When it attaches to the surface, the host cell internalizes the bacterium by phagocytosis. Pseudopods are used by both macrophages and amoebas to engulf *Legionella*. Because of the extensive movement of the host membrane around the bacteria, the cytoskeletal system plays a major role [15].

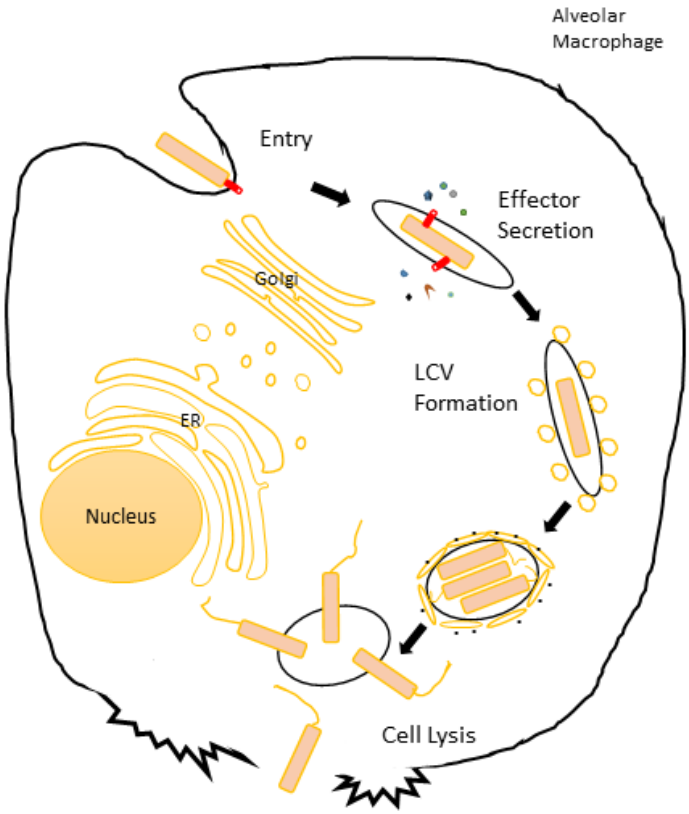


Figure 1-1: *Legionella* lifecycle inside the host cell. The whole process can be divided into two parts: pre- and post- *Legionella* containing vacuole (LCV) formation. Host cell pathway manipulations differ according to the presence of an LCV within the host cell.

The most crucial aspect of *Legionella* infection is the secretion of effector proteins by the Dot/Icm type IVB secretion system (T4BSS). When the bacteria lack the secretory complex (which is generally denoted as Δ*dotA* for the dotA gene deletion), infectability is completely diminished in both macrophages and amoeba[16]. The complex is found in Gram-negative bacteria and in some archaea species. T4B systems are classified into two groups: 4A and 4B. The most studied example of the Type IVA system is found in *Agrobacterium tumefaciens* and is used by the organism to transfer DNA into the host plant cells. The Type IVB system is found in *Coxiella* and *Legionella* species mainly for translocation of effector proteins [17]. There are numerous studies to understand the architecture of the secretion system and the most prominent results have been obtained by the Cryo-EM structures. According to recent findings, the system consists of a Hat

region interacting with the outer membrane layer; a Stem complex to bridge the Stalk and Hat domains; a Stalk domain to bridge the intracellular space to the cytosolic space; Wings that help the localization of Stalk on the membrane; and Rods that help the effector proteins' translocation from the Stalk domain. The whole complex consists of 25 different subunits and more than 100 molecules [18]. More than 330 effector proteins have been reported to have been secreted or identified as substrates for the secretion system [19]. These effector proteins hijack different host pathways, mainly for creating a suitable environment for Legionella containing vacuole (LCV) formation. This will be discussed later.

LCV is the continuation of the phagosome formed by the engulfment of *Legionella*. Even though it starts as a phagosome, studies showed that the internal pH is kept in a neutral range during formation and returns to an acidic level right before egress [20]. It is thought that this process is achieved by targeting the V-ATPase proton pumps located on the LCV with the effector protein SidK inhibiting activity of the pump [21]. The main purpose of *Legionella* to form LCVs is to evade the host phagolysosomal degradation pathway. Once a phagosome is formed by macrophages (or by amoeba cells in a similar fashion), the vesicle is destined to fuse with the lysosome and the content of the internalized material gets degraded. With LCV formation, the phagosome can evade the fusion and "camouflage" itself as a part of the host cell [22]. The evasion by camouflage is achieved by recruiting small vesicles that are destined from Endoplasmic reticulum to the Golgi around the LCV. Rab1 and Arl1 molecules have been shown to be collected around LCV by the effectors SidM and RalF, respectively, that act as a Guanine Exchange Factor (GEF) for the GTPases [23, 24]. The activation of GTPases help recruit small vesicles around the LCV, although this notion has been challenged by recent studies that show how phosphoribose(PR)-ubiquitination of ER resident proteins by SdeA cause a long tubular extension

from the ER to the LCV instead of individual small vesicles [25]. At this stage, LCV resembles smooth ER. Recruitment of ribosomes, whose mechanism remains elusive, creates the "illusion" of hiding the LCV as a rough ER [26]. As expected, the division of the bacteria and the development phases requires a great deal of energy. This need is met by the recruitment of mitochondria around the LCV [27]. Effector proteins sneak in as SNARE proteins, such as LncP, LegS2 and LegS7, and induce the binding of mitochondria to the LCV [28, 29]. Host cells deficient in mitochondria formation show almost no LCV formation and *Legionella* growth [30]. When the replication cycle of *Legionella* has been achieved and development of *Legionella* bacteria is finalized (the formation of flagellum [31]), the bacteria gets ready to egress from the host cell. Even though the whole process of egress has not been elucidated, formation of an egress pore has been observed that is essential for the lysis of the host cell [32].

1.3 Pathways modified by *Legionella*

During infection, *Legionella* modifies different host cellular pathways to promote the formation and development of LCV and replicate within it. The effector proteins secreted are involved in manipulating these pathways within a very tightly regulated scheme of functionality [33] (Table1-1). Most of these functions are essential from the initial stage of phagocytosis. Even though there is a huge redundancy within the effector protein repertoire, deletion of certain effector genes causes the loss of infection ability for *Legionella* [34].

Small GTPases

Effector secretion system T4BSS interacts with the membrane during phagocytosis and phagosome formation. Effector secretion starts as early as the phagosome formation and inclusion to the host cell. The triggering signal for effector secretion remains elusive [35]. As the main goal of the infection cycle is to form and maintain LCV, two main pathways need to be modified almost immediately: modulation of host vesicle trafficking to "steal" ER to Golgi destined anterograde vesicles and derailing from the lysosome fusion pathway [36]. Ras-like Small GTPases, Rab proteins, govern the regulation pathway of vesicle trafficking in cells. These molecules are inactive when a GDP molecule is located within their nucleotide binding pocket and they get activated and go under a conformational change to let their membrane-interacting alpha helical anchor get lipidated and bind to their respective membrane surfaces [37]. Rab1 and Arf1 are two small GTPases that regulate trafficking from ER to Golgi and these two proteins are among the first targets of *Legionella* effectors [38]. SidM/DrrA acts as a GEF for the Rab1 molecule by both its native function of exchanging GDP with GTP and forcing Rab1 to dissociate from its guanosine nucleotide dissociation inhibitor (GDI) [39]. SidM increases the activity of Rab1 by AMPylating its Tyr77 and blocks access to its GTPase Activation Proteins (GAP). This way, Rab1 gets more active and helps to recruit more membrane to the LCV [40]. Rab1 is returned to its normal state after getting deAMPylated by SidD effector protein so the LepB GAP molecule can start the deactivation process of Rab1[41, 42]. Effector AnkX phosphocholinates Rab1 to inhibit its activity further and Lem3 effector acts on Rab1 to de-phosphocholinate it only on LCV and keeps Rab1 active on LCV [43] This whole Rab1 modification by *Legionella* effectors is a great example for the importance of the consecutive secretion of effectors into the host cell and the essentiality of the timing of the activity of these effectors for the continuation of the infection cycle [44].

Vesicle recruitment is not the only branch of membrane trafficking that *Legionella* interferes with. To inhibit the fusion of phagosome that contains the LCV, *Legionella* must evade the lysosomal degradation pathway. Rab35, which is localized on endosomes, has been shown to be phosphocholinated by AnkX to inhibit its activity at the early stages of infection. It was speculated that the inactivation of Rab35 by AnkX might inhibit the cargo sorting, thus inhibiting the endosome maturation process [43, 45]. Like SidM acting on Rab1, another effector, Lpg0393, is shown to act as a general GEF molecule for different Rab molecules including Rab5 and Rab21, which are heavily involved in endosomal maturation; and their modulation affects the transport of LCV to lysosomes [46]. Another study also showed the effector LegG1 to have GEF activity towards RAN GTPase which is involved in promoting microtubule stabilization and LCV mobility, again hypothesized to be inhibiting the interaction of lysosomes with the LCV [47, 48]. Vesicle trafficking is shown to be strongly connected to actin metabolism as well and *Legionella* effector repertoire contains proteins that modulate the actin pathway. VipA molecule associates with actin bundles around multi vesicular bodies (MVBs) and is shown to be an actin nucleator that inhibits the new bundle formation for the motility of the vesicle [49]. LegK2 protein inhibits the activity of ARP2/3 actin nucleator complex by phosphorylating the ARPC1B and ARP3 subunits [50].

Protein Name	gene_id	Annotation	Reference
AnkX	lpg0695	Phosphocholine transferase	Mukherjee, 2011
Ceg14	lpg0437	Inhibition of actin polymerization	Guo, 2014
ceg15	lpg0439	Phospholipase A2	Zusman, 2007
ceg18	lpg0898	Activates caspase 3	Zhu, 2013
Ceg19	lpg1121	Possible dna binding domain	Burstein, 2009
Ceg2	lpg0059	Possible MCF transporter	Huang, 2010
Ceg22	lpg1484	peptidase (low chance kinase)	Zusman, 2007
Ceg28	lpg2311	Interaptin	Huang, 2010
Ceg4	lpg0096	haloacid dehalogenase-like hydrolase	Huang, 2010
Ceg9	lpg0246	Interaction with Reticulon 4, manipulate secretory trafficking	Haenssler, 2015
GobX	lpg2155	E3 Ub ligase	Lin, 2015
LecA	lpg1692	F-boxdomain-containing protein, manipulates host phospholipid biosynthesis	Viner, 2012
legA7	lpg0403	ankyrin-repeat containing protein	Habyarimana, 2008
LegAS4/RomA	lpg1718	Methyltransferase	Rolando, 2013
LegAU13 /AnkB	lpg2144	Fbox protein, ankyrin repeat, E3 ub ligase	Price, 2011
LegC3	lpg1701	kinectin1 - kinesin receptor, involved in subversion of vesicle trafficking	de Felipe, 2008
LegC5/Lgt3	lpg1488	Glucosyl transferase	Shen, 2009
LegC7(YlfA)	lpg2298	SNARE mimicry	O'Brian, 2015
LegC8/Lgt2	lpg2862	glucosyltransferase	Belyi, 2008
LegG1/PieG	lpg1976	UVB-resistance protein UVR8 (potential GAP)	Simon, 2014
LegK1	lpg1483	Serine/threonine protein kinase	Ge, 2009
LegK2	lpg2137	kinase/ATP interaction	Michard, 2015
LegS2/LpSPL	lpg2176	sphingosine-1 phosphate lyase	Sheedlo, 2015
LegU1	lpg0171	E3 Ubiquitin Ligase	Ensminger, 2010
LegU2/LubX	lpg2830	E3 ub ligase	Kubori, 2010
Lem10	lpg1496	SidE member (contains PDE)	Wong, 2015
Lem12	lpg1625	involved in caspase 3 activation	Zhu, 2013
Lem14	lpg1851	Structural paralog og lpg0634(LpiR1)	Beyrakhova, 2016
Lem3	lpg0696	PC hydrolase	Tan, 2011
Lem7	lpg1145	Possible RNA binding domain	Burstein, 2009
LepB	lpg2490	myosin like domain/Rab1 GAP	Preissler, 2017
Lgt1	lpg1368	Glucosyltransferase	Belyi, 2018
LpdA	lpg1888	Phohspholipase D	Viner, 2012
MavA	lpg1687	Ankyrin repeat	Huang, 2010
MavB	lpg1752	metallophosphoesterase	Huang, 2010
MavJ	lpg2498	Helicase	Huang, 2010
MavK	lpg2525	Fbox and WD40 repeat containing	Huang, 2010
MavN/IroT	lpg2815	Iron Transport	Isaac, 2015
MavQ	lpg2975	ribosomal RNA large subunit methyltransferase	Huang, 2010
MavR	lpg0209	Serine kinase	Zusman, 2007
MavU	lpg1798	slit-robo rho GTPase activating	Huang, 2010
MavW	lpg2907	Ubiquitin-like protease	Catic, 2007

PieA/LirC	lpg1963	Alteration of lysosome morphology	Ninio, 2009
PieE	lpg1969	Binds Rab1, 5, 6, 7, 10	Mousnier, 2014
pkn5/Ceg6	lpg0208	ser/thr/ kinase	Havey, 2015
PlcC/CegC1	lpg0012	phosphatidylcholine-hydrolyzing phospholipase	Aurass, 2013
PpiB	lpg1962	Peptidyl-prolyl cis-trans isomerase (ppiA)	Soderberg, 2007
RalF	lpg1950	Arf1 GEF	Nagai, 2002
RavJ	lpg0944	Accumulates F-actin	Liu, 2017
RavK(SidK ortholog)	lpg0969	Metalloprotease	Liu, 2017
RavL	lpg1108	Putative lipase	Huang, 2010
RavN	lpg1111	acyl transferase	Huang, 2010
RavO	lpg1129	Metaeffector for lpg0208	Havey, 2015
RavP	lpg1152	3,4-dihydroxy-2-butanone-4-phosphate synthase	Huang, 2010
RavR	lpg1166	Coiled coil protein	Huang, 2010
RavS	lpg1183	Chaperone	Huang, 2010
RavX	lpg1489	protein synthesis inhibition	Copenhaver, 2015
RavZ	lpg1683	ATG8 Cysteine protease	Choy, 2012
rvfA	lpg1797	coiled coil protein	Huang, 2010
sdbA	lpg0275	alpha/beta hydrolase	Bartfield, 2009
sdbC	lpg2391	alpha/beta hydrolase	Huang, 2010
sdcA	lpg2510	E3 Ub Ligase	Hsu, 2014
sdeA	lpg2157	SidE member	Bhogaraju, 2016
sdeB	lpg2156	SidE member	Bhogaraju, 2016
sdeC	lpg2153	SidE member	Bhogaraju, 2016
sdeD	lpg2509	SidE member (contains PDE)	Bhogaraju, 2016
sdeF	lpg2154	SidE member (contains PDE)	Bhogaraju, 2016
sdjA	lpg2155	DUB of PR-Ub	Qiu, 2017
SdmB	lpg2603	contains DrrA P4M domain	Huang, 2010
SetA	lpg1978	Glucosyl-transferase	Heidtman, 2009
SidC	lpg2511	E3 Ub Ligase	Hsu, 2014
SidD	lpg2465	Adenosine monophosphate-protein hydrolase	Neunuebel. 2011
SidE	lpg0234	mART, PDE, DUB	Bhogaraju, 2016
SidF	lpg2584	PI phosphatase	Hsu, 2012
SidJ	lpg2508	PR-Ub DUB	
SidK	lpg0968	Coiled coil protein (binds V-ATPase)	Zhao, 2017
SidM member	lpg1101	PI4P binding	Hubber, 2014
SidM/DrrA	lpg2464	Rab1 Ampylation	Brombacher, 2009
SidP	lpg0130	PI phosphatase	Toulabi, 2013
StuC	lpg2146	Sensor histidine kinase	Loza-Correa, 2014
VipA	lpg0390	Actin Nucleator	Franco, 2012
VipD	lpg2831	Patatin-like phospholipase	Gaspar, 2014
VipE	lpg2813	Patatin-like phospholipase	Shohdy, 2005

VipF	lpg0103	N-acetyltransferase activity	Young, 2016
VipF	lpg1227	Patatin-like phospholipase	Shohdy, 2005
VpdA	lpg2410	Vipd ortholog (patatin like phospholipase)	Shohdy, 2005
wipB	lpg0642	Ser/Thr Phosphatase	Prevost, 2017

Table 1-1: List of *Legionella* effector proteins with reported functions. Various host cellular pathways are manipulated by the effector proteins of *Legionella pneumophila*. Approximately 30% of effectors (among 330) have either known function or a role during infection.

Phospholipids

Host proteins are not the only targets of *Legionella* effectors. Small molecules like phospholipids are heavily modified during infection. Phosphoinositides (PIs), a subclass of phospholipids that have an inositol ring as a head group, act as organelle identifiers. Different proteins can bind to different phosphorylated states of the inositol ring (phosphorylation permutation on third, fourth, and fifth carbon atoms) [51]. *Legionella* effectors have been reported to either bind to these PI molecules or directly modify the phosphorylation state. One of the first evidences of PI re-modulation by *Legionella* was shown on the LCV: Regular phagosomes possesses PI(4,5)P_2 on the surface, because it is the major PIP on the plasma membrane; then the PIPs are converted to PI(3)P on early endosome and PI(3,5)P_2 on late endosome and lysosomes [52]; but the LCV surface was reported to have PI(4)P on the membrane, which is an indicator for the Golgi or plasma membrane, not the phagosome [53]. The exact mechanism of PI(4)P enrichment is not elucidated, but two effector proteins are acting directly on the PIPs on LCV: PI3 phosphatases SidF and SidP. SidF removes P3 from PI(3,4,5)P_3 and PI(3,4)P_2 and SidP removes from PI(3)P directly [54, 55]. Effector LpnE indirectly helps with the PI(4)P accumulation by recruiting OCRL1, a 5-phosphatase that acts on PI(4,5)P_2, to the LCV [56]. Phosphatases and kinases are not the only enzymes acting on PIPs by *Legionella*. VipD is a phospholipase A that acts on PI(3)P to a lysophopholipid and a fatty acyl chain. The protein is recruited by Rab5 and

Rab22 so it can be located on LCV during the early endosomal stage and stay on LCV during infection [57]. LpSpl and its homolog LegS2 are both sphingosine-1-phosphate lyases that act on sphingomyelins and modulate autophagy [58]. Some effectors use PIPs only for localizing at certain locations in the cell. PI(3)P and PI(4)P are main PIPs that effectors interact with during the infection [53]. SidM, Rab1 GEF, binds to PI(4)P on the LCV [23]; SidC, an E3 Ubiquitin ligase that will be discussed further, again binds to PI(4)P on the LCV [59]; LidA, which binds to Rab1-GTP to keep it activated, performs the same binding to PI(4)P as well [60]; and finally altough the functions of Lpg1101 and Lpg2603 are not known, they are shown to bind PI(4)P, too [61]. PI(3)P binding effectors spread throughout the cell during infection. RidL, retrograde trafficking inhibitor that binds to retromer complex directly, binds to PI(3)P on early endosomes [62]; SetA, a glucosyl-transferase, has a PI(3)P binding domain that causes bridging of vesicles [63]; LtpD, effector with no reported function, binds to the remaining PI(3)P pool on LCV and early endosomes [64].

1.4 Modulation of Ubiquitination

Ubiquitin is a small protein made up of 76 amino acids and is one of the most studied molecules in molecular biology. The molecule is used in a post-translational modification called ubiquitination in almost all eukaryotic cells [65]. The modification has been linked to various cellular pathways including protein degradation, inflammation, endocytosis, Histone regulation, DNA repair, and cell cycle [66-68]. Canonical ubiquitination involves three types of enzyme: a ubiquitin activation enzyme E1, a ubiquitin conjugating enzyme E2, and a ubiquitin ligase E3 (Figure 1-2) [69]. The cascade starts with the ubiquitin being activated by E1 by forming a thioester linkage between the G76 of ubiquitin and the cysteine residue in the active site of E1. The energy needed to form the bond is obtained from ATP. The “activated” ubiquitin is transferred to an E2 molecule and then an E2-Ub thioester intermediate state forms. Next, ubiquitin is transferred from

the E2 to the surface-accessible lysine residue of a substrate with the help of the E3 ligase enzyme. E3 ubiquitin ligase enzyme transfers ubiquitin molecules to its target proteins, which is detailed below [70]. Ubiquitin itself has seven lysine residues and the activating glycine residue that helps ubiquitination machinery to create complex ubiquitination branches or just polyubiquitin chains on substrates. The lysine residue used for ubiquitination also determines the fate of the substrate. One of the most studied lysine resides, K48, destines the proteins to proteolysis. This specific ubiquitination is recognized by the 26S subunit of proteasome [71]; another example is K63 linked polyubiquitin chains that act as a secondary messenger involved in inflammatory responses [72]. This whole ubiquitination process is reversible like most of the other post-translational modifications. There are enzymes called deubiquitinases (DUBs) that cleave the isopeptide bond between ubiquitin and the substrate [73]. The enzymatic activity of E1 and E2 enzymes are not distinct, but there are two main E3 ligases that recognize their substrates in different ways. HECT (homologous to the E3AP C-terminus) type E3 ligases form an intermediate with the activated ubiquitin molecule through a thioester bond with the cysteine residue in the active site and transfer it to the substrate in a two-step reaction [74]; whereas the RING (really interesting new gene) or

U-box type E3 ligases modulate the transfer of ubiquitin from the activated E2 molecule directly to the substrate protein. No intermediate is formed by RING domain ligation [75].

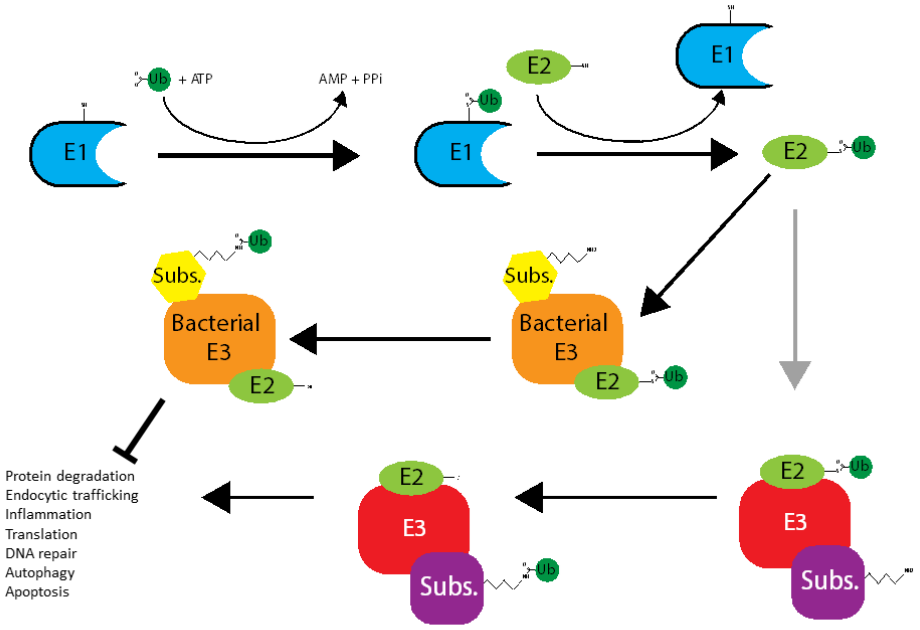


Figure 1-2: Conventional Ubiquitination Pathway and Subversion by Bacterial E3 Ligase Effectors. Conventional ubiquitination requires the enzyme cascade of E1, E2, and E3 for ubiquitination of target proteins. Pathogenic bacteria secrete effectors that behave as E3 ligases and manipulate the ubiquitination pathway.

Modification by ubiquitination is specific to eukaryotic cells. Even though bacteria have a prokaryotic ubiquitin-like protein (Pup) system, this is not as complex and highly regulated as ubiquitination and its machinery [76]. During infection, ubiquitination machinery targets phagosomal membrane proteins to direct the organelle for autophagy and the effector proteins secreted for degradation as an innate immunity response [77]. To use this machinery for infection, through horizontal gene transfer, various bacteria species have "stolen" the ubiquitination enzymes

and added them to their "arsenal" [78]. Evolutionarily, because bacteria either use the ubiquitin directly or divert the ubiquitination machinery, bacterial pathogens can have E3 ligase activity and these proteins are named as bacterial E3 ligase effectors (BELs) [79]. Varying pathogenic bacterium species that use Type III or Type IV secretion systems have BELs secreted through infection. *Salmonella, Shigella, Chlamydia, Pseudomonas,* and *Legionella,* all exploit ubiquitination machinery.

There are two RING domain containing BEL proteins secreted by *Legionella* (Figure 1-2). GobX protein sequence shows similarities to other eukaryotic E3 enzymes and it is shown to localize at the Golgi when over-expressed in mammalian cells. Palmitoylation of GobX allows the enzyme to be localized on the membrane. The exact role of the enzyme during infection is not known, but it is speculated to be involved in the regulation of either vesicle traffic to the LCV or the blocking of the exocytosis pathway [80]. LubX E3 ligase is secreted at a later stage of infection and does not target host proteins. It acts as a meta-effector: a bacterial effector protein that acts on another effector protein from the same bacteria. SidH effector is shown to be ubiquitinated by LubX and sent to proteasome to regulate its function during infection [81]. LubX was also shown to interact with and ubiquitinate Cdc2-like kinase 1 that regulates alternative splicing by phosphorylating SR protein of splicing regulators [82]. The exact reason behind the possible splicesome modulation by *Legionella* remains elusive because the infection can progress normally with *lubX* deletion from *Legionella* genome [83].

F-box motif containing proteins are involved in a slightly modified ubiquitination pattern. The proteins with F-box motif bind to Skp1; and Cul1 and this complex bind to RING-box 1 (Rbx1) protein to form a fully functional E3 ligase [84]. *Legionella* is reported to secrete five effectors that have a predicted F-box motif. LegU1 (AnkB) and LegAU13 (Ceg27) are shown to

physically interact with Skp1 and Cul1 [85]. AnkB was shown to target the HAL-B associated transcript 3 protein Bag6 which is ER-localized and shown to be involved in various pathways including protein degradation, chaperone, and ER stress [86]. The target of LegAU13 protein was not identified, but the effector is shown to interact with other proteins by its ankyrin repeat domain [87] and the essentiality of the protein for infection is not conserved in all species. Interestingly, even though the reason is yet to be determined, the protein is shown to be farnesylated [88]. Another interesting E3 ligase secreted by *Legionella* is SidC (and its paralog SdcA). The protein is a typical BEL to "steal" the ubiquitin from E2s in the host (UbcH5) and transfer it to targets yet-to-be-determined [59]. The protein is shown to have a PI(4)P binding domain (P4C) that targets the membrane on LCV and allosterically regulates the E3 activity by blocking the active site when not bound to the membrane [89].

1.5 SidE Effector Protein Family

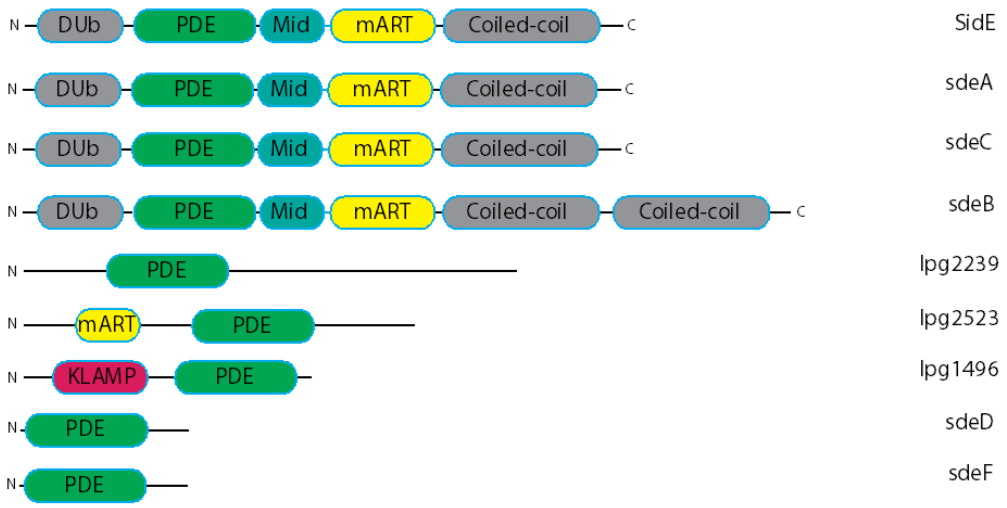


Figure 1-3: SidE Protein Family. *Legionella* effectors SidE protein family consists of nine members. Large members, SidE, SdeA-C, contain a DUB domain, a PDE domain, a mART domain linked to a Mid domain and an uncharacterized C'-coiled coil domain. The small members share the PDE domain.

Legionella effector proteins are highly regulated during and after the host cell takeover. These effectors are either regulated by other effectors named as meta-effectors or by localizing to different parts of the cell during infection. For the latter strategy, *Legionella* has protein paralogs that are very similar but differ slightly in functionality. SidC and SdcA, as an example, are two E3 ligases that share a high level of homology but can function with two different E2 enzymes [59]. There are also protein families that share domains among other functional domains. SidE protein family was first found to be expressed in the early stages of the infection and localized at the tip of the LCV [90] and even was predicted to be secreted before the phagocytosis of the bacterium [91]. The protein family consists of nine members: four large proteins and five smaller proteins. All these proteins share a phosphodiesterase (PDE) domain with previously unknown function (Figure 1-3). The first characterization work was for the N terminal region of the large SidE members. It was found that the N terminal region acts as a DUB enzyme and can cleave ubiquitin modification by K11, K48 or K63 linkages [92]. The actual interesting story about these proteins was reported by the following domains that the large members of the SidE family share. It was shown that these proteins can ubiquitinate targets independent of ubiquitin activation by an E1 and E2 enzyme-cascade. It was revealed that by using an NAD molecule as a cofactor, the mono ADP ribosyl-transferase (mART) domain ADP ribosylates ubiquitin molecule from its R42 residue [93]. The domain upstream of mART contains the shared PDE domain by the SidE members. PDE ligates the ADP-ribosylated ubiquitin to the target protein by cleaving the ADP-ribose moiety from the β-phosphate and linking the phosphate to the serine residue of the target protein and releasing an AMP molecule. This creates a phosphoribosyl ubiquitin moiety on the target protein (Figure 1-4) [94].

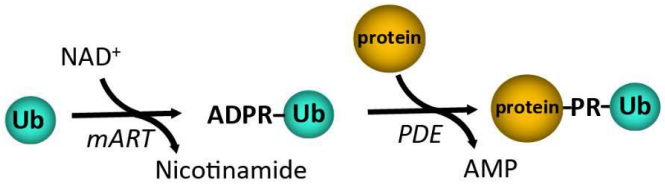


Figure 1-4: Phosphoribosyl-ubiquitination by SdeA. Schematic diagram of the phosphoribosyl ubiquitination reaction catalyzed by SdeA. First reaction is carried out by its mART domain and the second by its PDE domain.

In two distinct studies, two proteins were shown to be targeted by SidE family: Rab33b [93] and Reticulon 4 [25]. The outcome of phosphoribosyl-ubiquitination of Rab33b is only speculated: the Rab GTPase is localized on Golgi and is involved in Golgi to ER trafficking and progression of autophagy by helping the localization of Atg12 protein on autophagosome membrane [95]. With this information, the authors claim the PR-ubiquitination of Rab33b might affect either Golgi to ER retrograde trafficking, or it might affect the autophagy pathway. Both pathways might help with the formation of LCV [93], but the fact that PR-ubiquitination of Rab33 does not affect its GTPase activity raises some questions [94]. The other target protein, Reticulon 4, is an ER membrane protein [96] and is shown to generate ER membrane curvature by inserting hydrophobic hairpin structures into the cytoplasmic site of ER [97]. Reticulon 4 is shown to be localized to LCV during infection and after the PR-ubiquitination of this protein, it is shown to "pull" ER tubules to cover the LCV [25]. The only other study about this protein family shows lpg1496 to have an additional KLAMP domain that interacts with cyclic nucleotides [98].

1.6 Conclusion

Legionella pneumophila is a pathogen that invades alveolar macrophages and causes Legionnaires' disease. The completion of the host cell invasion relies on the formation of LCV by the bacteria. LCV formation requires a highly regulated secretion and functioning of *Legionella* effector proteins. These proteins affect host pathways including protein expression, lipid metabolism, protein-lipid signaling, protein degradation, and autophagy. Like other pathogens, *Legionella* secretes proteins that contain eukaryotic protein domains and can alter host pathways that prokaryotes normally do not have. One of the eukaryotic pathways that *Legionella* disrupts is ubiquitination. There are E3 ligases secreted to steal ubiquitin from the corresponding E2 enzymes that help with the formation of the LCV. Recently, it was shown that SidE proteins modify ubiquitin in a unique phosphoribosylation mechanism and labels target proteins with this unusual phosphoribosyl-ubiquitin moiety.

The significance of SidE family proteins lies in the phosphoribosylation reaction. The PDE domain of SidE family is significant for being the first documented PDE domain to act on proteins instead of nucleic acids or lipids. The coordination of an mART domain with a PDE domain is documented for the first time as well. For other ADP Ribosylation domains, the ligated moiety is hydrolyzed from the protein directly without cleaving any part of the ADP moiety in a reversible way, just as in the example of Poly (ADP-ribose) polymerase enzymes (PARPs) [99]. These proteins are involved in DNA repair and synthesis processes and can poly or mono-ADP-ribosylate proteins. [100]. The PDE domain of SidE family differs by cleaving the AMP from the ADP-ribose moiety to irreversibly modify ubiquitin before ligating it to a target protein. There are currently no documented eukaryotic PDE domains that can cleave AMP from an ADP-ribose moiety. How *Legionella* acquired this specific PDE activity remains elusive. Rab33b and Reticulon 4 proteins

are shown to be targeted by SidE proteins. The modification during infection is shown by different studies, but why these proteins are targeted during infection also remains elusive. It is important to know why *Legionella* is targeting host proteins in a unique way, especially because the modification of Rab33b doesn't inhibit or increase its GTPase activity [101].

There are currently 324 known effector proteins secreted during infection and approximately one-third of these proteins have documented functions or roles during infection. The arsenal of *Legionella* remains elusive and new pathways that these proteins are involved in need to be solved to completely understand and map the infection submersion by *Legionella*. This knowledge will help to understand the infection of other pathogens as well.

References

1. 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.

2. Bruggemann, H., C. Cazalet, and C. Buchrieser, *Adaptation of Legionella pneumophila to the host environment: role of protein secretion, effectors and eukaryotic-like proteins.* Curr Opin Microbiol, 2006. **9**(1): p. 86-94.

3. Zhang, Q., et al., *Legionnaires' disease caused by Legionella pneumophila serogroups 5 and 10, China.* Emerg Infect Dis, 2014. **20**(7): p. 1242-3.

4. Solomon, J.M. and R.R. Isberg, *Growth of Legionella pneumophila in Dictyostelium discoideum: a novel system for genetic analysis of host-pathogen interactions.* Trends Microbiol, 2000. **8**(10): p. 478-80.

5. Lau, H.Y. and N.J. Ashbolt, *The role of biofilms and protozoa in Legionella pathogenesis: implications for drinking water.* J Appl Microbiol, 2009. **107**(2): p. 368-78.

6. Jarraud, S., et al., *Identification of legionella in clinical samples.* Methods Mol Biol, 2013. **954**: p. 27-56.

7. Honigsbaum, M., *Legionnaires' disease: revisiting the puzzle of the century.* Lancet, 2016. **388**(10043): p. 456-7.

8. Cunha, C.B. and B.A. Cunha, *Legionnaire's Disease Since Philadelphia: Lessons Learned and Continued Progress.* Infect Dis Clin North Am, 2017. **31**(1): p. 1-5.

9. Burillo, A., M.L. Pedro-Botet, and E. Bouza, *Microbiology and Epidemiology of Legionnaire's Disease.* Infect Dis Clin North Am, 2017. **31**(1): p. 7-27.

10. Magira, E.E. and S. Zakynthinos, *Legionnaire's Disease and Influenza.* Infect Dis Clin North Am, 2017. **31**(1): p. 137-153.

11. Brusch, J.L., *Legionnaire's Disease: Cardiac Manifestations.* Infect Dis Clin North Am, 2017. **31**(1): p. 69-80.

12. Samrakandi, M.M., et al., *Entry into host cells by Legionella.* Front Biosci, 2002. **7**: p. d1-11.

13. Venkataraman, C., et al., *Identification of a Gal/GalNAc lectin in the protozoan Hartmannella vermiformis as a potential receptor for attachment and invasion by the Legionnaires' disease bacterium.* J Exp Med, 1997. **186**(4): p. 537-47.

14. Payne, N.R. and M.A. Horwitz, *Phagocytosis of Legionella pneumophila is mediated by human monocyte complement receptors.* J Exp Med, 1987. **166**(5): p. 1377-89.

15. Charpentier, X., et al., *Chemical genetics reveals bacterial and host cell functions critical for type IV effector translocation by Legionella pneumophila.* PLoS Pathog, 2009. **5**(7): p. e1000501.

16. Berger, K.H. and R.R. Isberg, *Intracellular survival by Legionella.* Methods Cell Biol, 1994. **45**: p. 247-59.

17. Chandran Darbari, V. and G. Waksman, *Structural Biology of Bacterial Type IV Secretion Systems.* Annu Rev Biochem, 2015. **84**: p. 603-29.

18. Ghosal, D., et al., *In situ structure of the Legionella Dot/Icm type IV secretion system by electron cryotomography.* EMBO Rep, 2017. **18**(5): p. 726-732.

19. 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.

20. Sturgill-Koszycki, S. and M.S. Swanson, *Legionella pneumophila replication vacuoles mature into acidic, endocytic organelles.* J Exp Med, 2000. **192**(9): p. 1261-72.

21. Xu, L., et al., *Inhibition of host vacuolar H+-ATPase activity by a Legionella pneumophila effector.* PLoS Pathog, 2010. **6**(3): p. e1000822.

22. Horwitz, M.A., *Formation of a novel phagosome by the Legionnaires' disease bacterium (Legionella pneumophila) in human monocytes.* J Exp Med, 1983. **158**(4): p. 1319-31.

23. Brombacher, E., et al., *Rab1 guanine nucleotide exchange factor SidM is a major phosphatidylinositol 4-phosphate-binding effector protein of Legionella pneumophila.* J Biol Chem, 2009. **284**(8): p. 4846-56.

24. Folly-Klan, M., et al., *A novel membrane sensor controls the localization and ArfGEF activity of bacterial RalF.* PLoS Pathog, 2013. **9**(11): p. e1003747.

25. 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.

26. 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.

27. Chong, A., et al., *The purified and recombinant Legionella pneumophila chaperonin alters mitochondrial trafficking and microfilament organization.* Infect Immun, 2009. **77**(11): p. 4724-39.

28. Degtyar, E., et al., *A Legionella effector acquired from protozoa is involved in sphingolipids metabolism and is targeted to the host cell mitochondria.* Cell Microbiol, 2009. **11**(8): p. 1219-35.

29. Dolezal, P., et al., *Legionella pneumophila secretes a mitochondrial carrier protein during infection.* PLoS Pathog, 2012. **8**(1): p. e1002459.

30. Francione, L., et al., *Legionella pneumophila multiplication is enhanced by chronic AMPK signalling in mitochondrially diseased Dictyostelium cells.* Dis Model Mech, 2009. **2**(9-10): p. 479-89.

31. Escoll, P., et al., *From amoeba to macrophages: exploring the molecular mechanisms of Legionella pneumophila infection in both hosts.* Curr Top Microbiol Immunol, 2013. **376**: p. 1-34.

32. Alli, O.A., et al., *Temporal pore formation-mediated egress from macrophages and alveolar epithelial cells by Legionella pneumophila.* Infect Immun, 2000. **68**(11): p. 6431-40.

33. Asrat, S., et al., *Bacterial pathogen manipulation of host membrane trafficking.* Annu Rev Cell Dev Biol, 2014. **30**: p. 79-109.

34. Fields, B.S., *The molecular ecology of legionellae.* Trends Microbiol, 1996. **4**(7): p. 286-90.

35. Qiu, J. and Z.Q. Luo, *Effector translocation by the Legionella Dot/Icm type IV secretion system.* Curr Top Microbiol Immunol, 2013. **376**: p. 103-15.

36. Lu, H. and M. Clarke, *Dynamic properties of Legionella-containing phagosomes in Dictyostelium amoebae.* Cell Microbiol, 2005. **7**(7): p. 995-1007.

37. Bhuin, T. and J.K. Roy, *Rab proteins: the key regulators of intracellular vesicle transport.* Exp Cell Res, 2014. **328**(1): p. 1-19.

38. Hardiman, C.A., et al., *The role of Rab GTPases in the transport of vacuoles containing Legionella pneumophila and Coxiella burnetii.* Biochem Soc Trans, 2012. **40**(6): p. 1353-9.

39. Schoebel, S., et al., *RabGDI displacement by DrrA from Legionella is a consequence of its guanine nucleotide exchange activity.* Mol Cell, 2009. **36**(6): p. 1060-72.

40. Neunuebel, M.R., et al., *De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila.* Science, 2011. **333**(6041): p. 453-6.

41. Tan, Y. and Z.Q. Luo, *Legionella pneumophila SidD is a deAMPylase that modifies Rab1.* Nature, 2011. **475**(7357): p. 506-9.

42. Ingmundson, A., et al., *Legionella pneumophila proteins that regulate Rab1 membrane cycling.* Nature, 2007. **450**(7168): p. 365-9.

43. Mukherjee, S., et al., *Modulation of Rab GTPase function by a protein phosphocholine transferase.* Nature, 2011. **477**(7362): p. 103-6.

44. Arasaki, K., D.K. Toomre, and C.R. Roy, *The Legionella pneumophila effector DrrA is sufficient to stimulate SNARE-dependent membrane fusion.* Cell Host Microbe, 2012. **11**(1): p. 46-57.

45. Kouranti, I., et al., *Rab35 regulates an endocytic recycling pathway essential for the terminal steps of cytokinesis.* Curr Biol, 2006. **16**(17): p. 1719-25.

46. Sohn, Y.S., et al., *Lpg0393 of Legionella pneumophila is a guanine-nucleotide exchange factor for Rab5, Rab21 and Rab22.* PLoS One, 2015. **10**(3): p. e0118683.

47. Rothmeier, E., et al., *Activation of Ran GTPase by a Legionella effector promotes microtubule polymerization, pathogen vacuole motility and infection.* PLoS Pathog, 2013. **9**(9): p. e1003598.

48. Simon, S., et al., *Icm/Dot-dependent inhibition of phagocyte migration by Legionella is antagonized by a translocated Ran GTPase activator.* Cell Microbiol, 2014. **16**(7): p. 977-92.

49. Franco, I.S., N. Shohdy, and H.A. Shuman, *The Legionella pneumophila effector VipA is an actin nucleator that alters host cell organelle trafficking.* PLoS Pathog, 2012. **8**(2): p. e1002546.

50. Michard, C., et al., *The Legionella Kinase LegK2 Targets the ARP2/3 Complex To Inhibit Actin Nucleation on Phagosomes and Allow Bacterial Evasion of the Late Endocytic Pathway.* MBio, 2015. **6**(3): p. e00354-15.

51. Di Paolo, G. and P. De Camilli, *Phosphoinositides in cell regulation and membrane dynamics.* Nature, 2006. **443**(7112): p. 651-7.

52. Saarikangas, J., H. Zhao, and P. Lappalainen, *Regulation of the actin cytoskeleton-plasma membrane interplay by phosphoinositides.* Physiol Rev, 2010. **90**(1): p. 259-89.

53. 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.

54. Toulabi, L., et al., *Identification and structural characterization of a Legionella phosphoinositide phosphatase.* J Biol Chem, 2013. **288**(34): p. 24518-27.

55. Hsu, F., et al., *Structural basis for substrate recognition by a unique Legionella phosphoinositide phosphatase.* Proc Natl Acad Sci U S A, 2012. **109**(34): p. 13567-72.

56. Weber, S.S., C. Ragaz, and H. Hilbi, *The inositol polyphosphate 5-phosphatase OCRL1 restricts intracellular growth of Legionella, localizes to the replicative vacuole and binds to the bacterial effector LpnE.* Cell Microbiol, 2009. **11**(3): p. 442-60.

57. Gaspar, A.H. and M.P. Machner, *VipD is a Rab5-activated phospholipase A1 that protects Legionella pneumophila from endosomal fusion.* Proc Natl Acad Sci U S A, 2014. **111**(12): p. 4560-5.

58. Abu Khweek, A., et al., *The Sphingosine-1-Phosphate Lyase (LegS2) Contributes to the Restriction of Legionella pneumophila in Murine Macrophages.* PLoS One, 2016. **11**(1): p. e0146410.

59. 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.

60. Neunuebel, M.R., et al., *Legionella pneumophila LidA affects nucleotide binding and activity of the host GTPase Rab1*. J Bacteriol, 2012. **194**(6): p. 1389-400.

61. Horenkamp, F.A., et al., *Legionella pneumophila subversion of host vesicular transport by SidC effector proteins.* Traffic, 2014. **15**(5): p. 488-99.

62. Finsel, I., et al., *The Legionella effector RidL inhibits retrograde trafficking to promote intracellular replication.* Cell Host Microbe, 2013. **14**(1): p. 38-50.

63. Jank, T., et al., *Domain organization of Legionella effector SetA.* Cell Microbiol, 2012. **14**(6): p. 852-68.

64. Harding, C.R., et al., *LtpD is a novel Legionella pneumophila effector that binds phosphatidylinositol 3-phosphate and inositol monophosphatase IMPA1*. Infect Immun, 2013. **81**(11): p. 4261-70.

65. Hershko, A. and A. Ciechanover, *The ubiquitin system.* Annu Rev Biochem, 1998. **67**: p. 425-79.

66. Kattah, M.G., B.A. Malynn, and A. Ma, *Ubiquitin-Modifying Enzymes and Regulation of the Inflammasome.* J Mol Biol, 2017.

67. Dubrez, L., *Regulation of E2F1 Transcription Factor by Ubiquitin Conjugation.* Int J Mol Sci, 2017. **18**(10).

68. de Poot, S.A.H., G. Tian, and D. Finley, *Meddling with Fate: The Proteasomal Deubiquitinating Enzymes.* J Mol Biol, 2017.

69. Dye, B.T. and B.A. Schulman, *Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins.* Annu Rev Biophys Biomol Struct, 2007. **36**: p. 131-50.

70. Scheffner, M., U. Nuber, and J.M. Huibregtse, *Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade.* Nature, 1995. **373**(6509): p. 81-3.

71. Hicke, L., *Protein regulation by monoubiquitin.* Nat Rev Mol Cell Biol, 2001. **2**(3): p. 195-201.

72. Komander, D. and M. Rape, *The ubiquitin code.* Annu Rev Biochem, 2012. **81**: p. 203-29.

73. D'Andrea, A. and D. Pellman, *Deubiquitinating enzymes: a new class of biological regulators.* Crit Rev Biochem Mol Biol, 1998. **33**(5): p. 337-52.

74. Bernassola, F., et al., *The HECT family of E3 ubiquitin ligases: multiple players in cancer development.* Cancer Cell, 2008. **14**(1): p. 10-21.

75. Aravind, L. and E.V. Koonin, *The U box is a modified RING finger - a common domain in ubiquitination.* Curr Biol, 2000. **10**(4): p. R132-4.

76. Maupin-Furlow, J.A., *Prokaryotic ubiquitin-like protein modification.* Annu Rev Microbiol, 2014. **68**: p. 155-75.

77. Liu, X., et al., *Dynamic regulation of innate immunity by ubiquitin and ubiquitin-like proteins.* Cytokine Growth Factor Rev, 2013. **24**(6): p. 559-70.

78. Wimmer, P. and S. Schreiner, *Viral Mimicry to Usurp Ubiquitin and SUMO Host Pathways.* Viruses, 2015. **7**(9): p. 4854-72.

79. Maculins, T., et al., *Bacteria-host relationship: ubiquitin ligases as weapons of invasion.* Cell Res, 2016. **26**(4): p. 499-510.

80. Lin, Y.H., et al., *Host Cell-catalyzed S-Palmitoylation Mediates Golgi Targeting of the Legionella Ubiquitin Ligase GobX.* J Biol Chem, 2015. **290**(42): p. 25766-81.

81. Kubori, T., et al., *Legionella metaeffector exploits host proteasome to temporally regulate cognate effector.* PLoS Pathog, 2010. **6**(12): p. e1001216.

82. Pandit, S., D. Wang, and X.D. Fu, *Functional integration of transcriptional and RNA processing machineries.* Curr Opin Cell Biol, 2008. **20**(3): p. 260-5.

83. Kubori, T., A.M. Hubber, and H. Nagai, *Hijacking the host proteasome for the temporal degradation of bacterial effectors.* Methods Mol Biol, 2014. **1197**: p. 141-52.

84. Zheng, N., et al., *Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex.* Nature, 2002. **416**(6882): p. 703-9.

85. 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.

86. Binici, J. and J. Koch, *BAG-6, a jack of all trades in health and disease.* Cell Mol Life Sci, 2014. **71**(10): p. 1829-37.

87. Perpich, J.D., et al., *Divergent evolution of Di-lysine ER retention vs. farnesylation motif-mediated anchoring of the AnkB virulence effector to the Legionella-containing vacuolar membrane.* Sci Rep, 2017. **7**(1): p. 5123.

88. Ivanov, S.S., et al., *Lipidation by the host prenyltransferase machinery facilitates membrane localization of Legionella pneumophila effector proteins.* J Biol Chem, 2010. **285**(45): p. 34686-98.

89. 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.

90. Bardill, J.P., J.L. Miller, and J.P. Vogel, *IcmS-dependent translocation of SdeA into macrophages by the Legionella pneumophila type IV secretion system.* Mol Microbiol, 2005. **56**(1): p. 90-103.

91. Vincent, C.D. and J.P. Vogel, *The Legionella pneumophila IcmS-LvgA protein complex is important for Dot/Icm-dependent intracellular growth.* Mol Microbiol, 2006. **61**(3): p. 596-613.

92. 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.

93. Qiu, J., et al., *Ubiquitination independent of E1 and E2 enzymes by bacterial effectors.* Nature, 2016. **533**(7601): p. 120-4.

94. Bhogaraju, S., et al., *Phosphoribosylation of Ubiquitin Promotes Serine Ubiquitination and Impairs Conventional Ubiquitination.* Cell, 2016. **167**(6): p. 1636-1649 e13.

95. Ao, X., L. Zou, and Y. Wu, *Regulation of autophagy by the Rab GTPase network.* Cell Death Differ, 2014. **21**(3): p. 348-58.

96. English, A.R., N. Zurek, and G.K. Voeltz, *Peripheral ER structure and function.* Curr Opin Cell Biol, 2009. **21**(4): p. 596-602.

97. Yang, Y.S. and S.M. Strittmatter, *The reticulons: a family of proteins with diverse functions.* Genome Biol, 2007. **8**(12): p. 234.

98. 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.

99. Ame, J.C., C. Spenlehauer, and G. de Murcia, *The PARP superfamily.* Bioessays, 2004. **26**(8): p. 882-93.

100. Morales, J., et al., *Review of poly (ADP-ribose) polymerase (PARP) mechanisms of action and rationale for targeting in cancer and other diseases.* Crit Rev Eukaryot Gene Expr, 2014. **24**(1): p. 15-28.

101. Nakasone, M.A. and D.T. Huang, *Ubiquitination Accomplished: E1 and E2 Enzymes Were Not Necessary.* Mol Cell, 2016. **62**(6): p. 807-809.

CHAPTER 2

THE MOLECULAR MECHANISM OF *LEGIONELLA* EFFECTOR-MEDIATED PHOSPHORIBOSYl-UBIQUITINATION

2.1 Abstract

Ubiquitination is a post-translational modification that regulates a myriad of cellular processes in eukaryotes[1-4]. 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[1, 5]. Recent studies of the *Legionella pneumophila* SidE family of effector proteins have revealed a novel mode of ubiquitination in which a phosphoribosylated ubiquitin (PR-Ub) is conjugated to a serine residue of substrates via a phosphodiester bond[6-8]. To uncover the molecular mechanism of this unique post-translational modification, we determined the crystal structure of a fragment of a SidE family member SdeA that retains its full ubiquitination activity. The structure reveals that the catalytic module contains three distinct domains: a phosphodiesterase domain (PDE), a middle α-helical domain (Mid) of unknown function, and a mono-ADP-ribosyltransferase (mART) domain. Biochemical analysis shows that the Mid domain is indispensable for mART domain-mediated ADP-ribosylation of Ub and presumably plays a role in Ub recognition. Our data also demonstrate that the conversion of Ub to ADP-ribosylated Ub (ADPR-Ub) and the ligation of PR-Ub to substrates are two independent activities of SdeA. Finally, we present two crystal structures of a homologous PDE domain from the SidE family member SdeD[9] in complex with Ub or ADPR-Ub, which reveal an intriguing mechanism for

how SdeA processes ADPR-Ub to PR-Ub and AMP and conjugates PR-Ub to a serine residue of substrates. Therefore, our results establish the molecular mechanism of phosphoribosyl-ubiquitination (PR-ubiquitination) and pave the way for future studies of this unusual type of ubiquitination in eukaryotes.

2.2 Introduction

Posttranslational protein modification by ubiquitin (Ub) is a central eukaryotic mechanism that regulates a plethora of physiological processes, including protein homeostasis [1], cell signaling [2], and membrane trafficking [3, 4]. Following the conventional scheme of ubiquitination, Ub is covalently coupled to lysine residues on target proteins via the sequential activities of a collection of enzymes known as E1, E2, and E3 [10]. The C-terminal glycine residue of Ub is first activated and covalently linked to the catalytic cysteine residue of the Ub activating enzyme E1 through a thioester bond with the consumption of an ATP. The activated Ub moiety is then transferred to the active site cysteine of an E2 Ub-conjugation enzyme. The resulting thioester-linked E2~Ub complex interacts with specific E3 Ub ligases, which promote the direct or indirect transfer of Ub to either the ε-amine of a lysine residue targeted proteins or the N-terminal amine of another Ub [11-13].

Given the vital role of ubiquitination in cell physiology, it is not surprising that a variety of microbial pathogens exploit this essential posttranslational modification pathway during the infection of their corresponding hosts [14]. For example, among the hundreds of effectors injected into host cells by the Dot/Icm transporter of the intracellular pathogen *L. pneumophila*, more than 10 proteins are involved in ubiquitin manipulation [15] including those that contain conserved eukaryotic F- or U-box domains found in E3 ubiquitin ligases [16-19]. Other E3 Ub ligases that

have a unique structural fold but similar catalytic chemistry to the HECT-type ligases have also been characterized [20, 21]. In addition to these Ub ligases which utilize the canonical

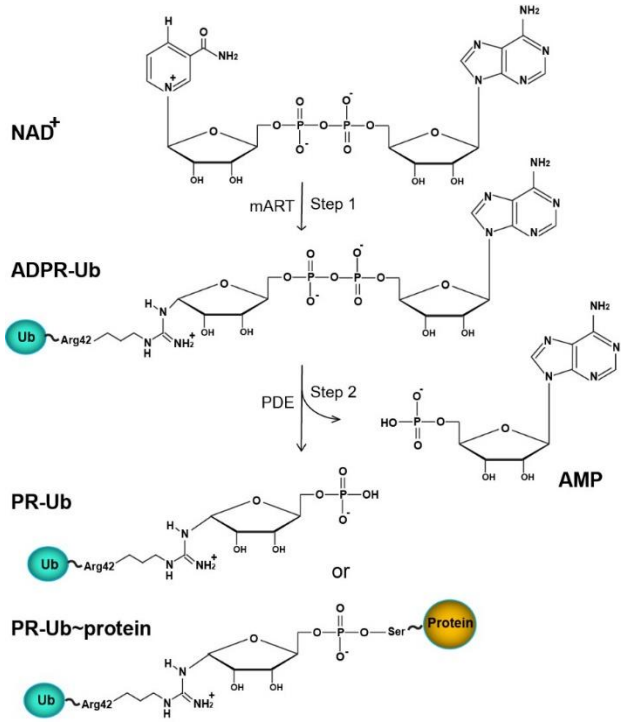


Figure 2-1: Schematic Diagram of the Enzymatic Pathway Catalyzed by SdeA. Serine ubiquitination catalyzed by SdeA involves two activities. First, the mART activity facilitates the ADP-ribosylation of Ub to generate ADPR-Ub by consuming an NAD^+. Second, the PDE activity catalyzes the cleavage of ADPR-Ub to PR-Ub if a water molecule is the Ub acceptor or the conjugation of PR-Ub to serine residues of substrate proteins if a substrate is present

host Ub machinery for ubiquitination, recent studies of the *L. pneumophila* Sde family effectors, such as SdeA uncovered a novel ubiquitination pathway that acts independently of E1 and E2 enzymes [6-8]. Instead, this unusual SdeA-catalyzed ubiquitination involves both mono-ADP-ribosyl transferase (mART) and phosphodiesterase (PDE) activities. SdeA first catalyzes the transfer of ADP-ribose from NAD^+ to the arginine 42 of ubiquitin to generate mono-ADP-ribosylated Ub (ADPR-Ub), using its mART activity. Subsequently, via the activity of the PDE domain, ADPR-Ub is transferred to a water molecule to generate phosphoribosylated ubiquitin (PR-Ub) and releases AMP in the absence of substrate proteins. Alternatively, ADPR-Ub can be conjugated to serine residues of substrate proteins to generate serine ubiquitinated products (Figure 2-1). Unlike conventional ubiquitination, this process is ATP independent. Instead, it consumes an NAD molecule in the activation of a Ub molecule. Moreover, there is no detectable difference in substrate ubiquitination when Ub mutants either carries substitutions on all surface lysine residues

or lacks the C-terminal di-glycine residues. Although the reaction products and the chemical linkage of SdeA-mediated ubiquitination have been established [6-8], the molecular mechanism of this unique ubiquitination pathway remains elusive, for example: how Ub is specifically recognized and modified by SdeA; how ADPR-Ub is processed and conjugated to substrates; and whether the intermediate ADPR-Ub generated by the mART is channeled to the PDE domain for the subsequent ligation to substrates. Detailed structural and biochemical studies are warranted to address these intriguing questions.

We have determined the crystal structure of a portion of SdeA (a.a. 211-910) that contains the predicted mART and PDE domains and is fully functional in catalyzing ubiquitination. Our structure reveals that the mART and PDE domains are packed tightly against each other, but the active sites of these two domains face opposite directions and are separated by a long distance. This structural organization corroborated with our biochemical results indicates that the ADP-ribosylation of Ub and the conjugation of ADPR-Ub to substrates are functionally independent. Our structure also unveils another conserved domain that is connected to the core of the enzyme via two long flexible loops. We show evidence that this domain may function to recruit and present Ub molecules to the mART domain for ADP-ribosylation. We have further determined the crystal structures of a homologous PDE domain from another *Legionella* effector SdeD [9] in its apo form and in complex with both Ub and ADPR-Ub. These structures show that Ub binds to the PDE domain and buries a large interface between the two molecules. More importantly, this interaction allows the ADPR moiety from ADPR-Ub to be positioned snugly in the deep catalytic groove of the PDE domain. Insights from our structural studies, together with biochemical and infection analysis, reveal an intriguing molecular mechanism underlying this unprecedented ubiquitination cascade.

2.3 RESULTS

Overall Structure of SdeA

It has been shown that substrate modification by PR-Ub is mediated by a portion of SdeA (a.a. 205-1000) that contains a predicted phosphodiesterase (PDE) and mono-ADP-ribosyltransferase (mART) domain [6, 7]. To probe the molecular mechanism of this unusual type of ubiquitination, we determined the crystal structure of SdeA (a.a. 221-910) by single-wavelength anomalous dispersion (SAD) phasing of crystals soaked with $HgCl_2$. The structure is comprised of three distinct domains, the PDE, mART, and an alpha-helical domain with unknown function (Figure 2-2).

The PDE domain mainly consists of alpha helices. The active site resides in a deep groove lined with mostly conserved residues, particularly the two invariant histidines. There is no significantly charged area on the surface of SdeA except for the PDE active site, which is highly positively charged (Figure 2-2 C and D). The concentrated positive charge at the PDE active site may facilitate the positioning of the phosphate groups of its substrate, ADPR-Ub. The overall structure of the SdeA PDE domain has a similar fold to that of another *Legionella* effector lpg1496 [22] with a root mean square deviation (rmsd) of 1.9 Å over 225 aligned Cα atoms. A prominent difference between the PDE domains is that the loops connecting the alpha helices vary both in primary sequence and in length (see Figure 2-8). Since these loops are protruding along the side of the active site groove, it is tempting to postulate that these loops may interact with targeted substrate proteins and play a role in substrate selection.

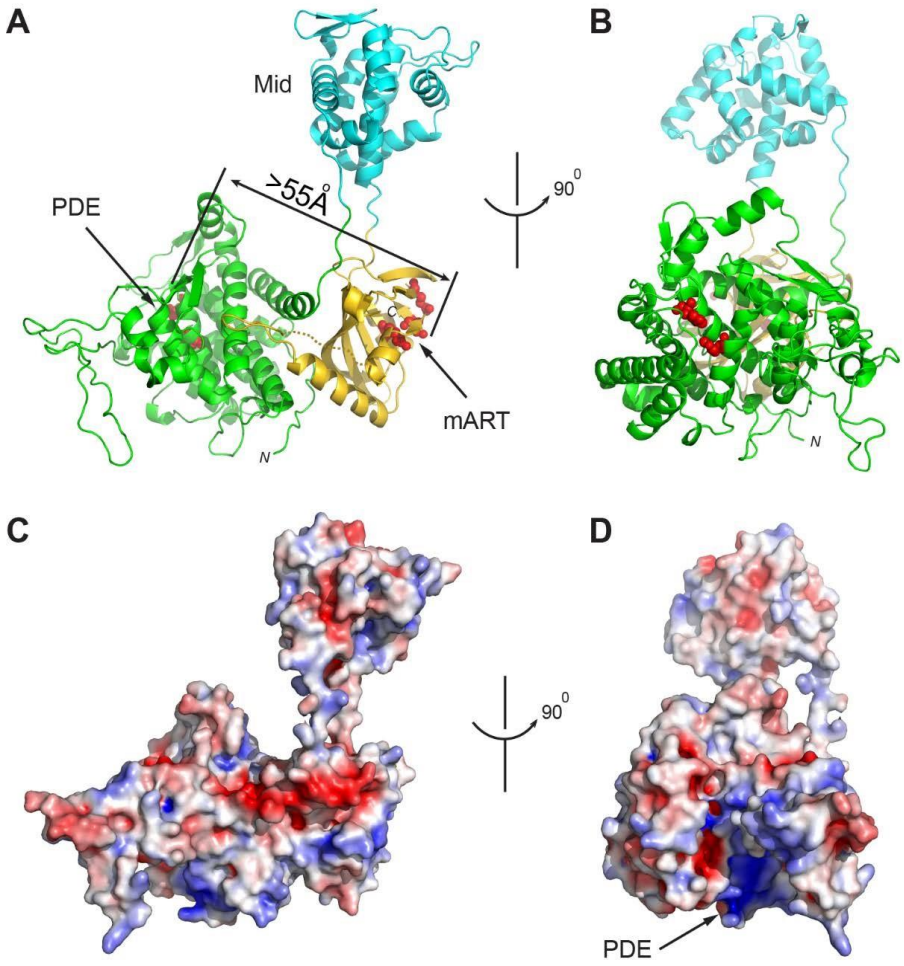


Figure 2-2: Overall Structure of SdeA. (A) Ribbon diagram of the overall structure of SdeA (211-901). This portion of SdeA has three distinct domains, including the PDE (shown in green), Mid (cyan), and mART (gold) domains. The active site residues of both the mART and PDE domains are shown in red spheres. The linear distance between these two active sites is about 55 Å. (B) An orthogonal view of (A). (C) Molecular surface of SdeA. 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). The orientation of the molecule is the same as shown in (A). (D) A 90^{o} rotated view of (C). The active site of the PDE domain is highly positively charged, which may facilitate the docking of the ADPR pyrophosphate group.

The mART domain of SdeA is composed of two nearly perpendicular β sheets. This two-layered β-sandwich core is flanked by three alpha helices bridging between the β strands. A

structural homology search with the Dali server [23] yielded a list of hits including the *Pseudomonas syringe* type III effector HopU1 (PDB ID: 3u0j) and the *Clostridium perfringens* toxin Iota-toxin (PDB ID 4H03). Both HopU1 and Iota-toxin are mono-ADP-ribosyltransferase that specifically modify the glycine-rich RNA-binding protein GRP7 [24] and actin [25] from the hosts, respectively. A structural comparison showed that the β-sandwich core is similar among all three of the mART proteins (Figure 2-3). All three mART proteins have three catalytically important conserved motifs: the (F/Y)-(R/H), STS, and EXE motif, suggesting a similar catalytic mechanism for ADP-ribosylation. However, the mART domain of SdeA displays two major differences compared to that of HopU1 and Iota-toxin. First, the ADP-ribosylating turn-turn loop (ARTT), which contains the EXE motif and is important for target protein recognition, is variable in sequence and highly flexible in conformation. Second, both HopU1 and Iota-toxin contain an additional alpha helical region at their N-terminus (colored in gray in Figure 2-3 b), which has been shown to be essential for substrate protein binding [24, 25]. However, the mART domain of SdeA lacks such an alpha helical portion. These two major structural differences suggest that the mART domain of SdeA may use a distinct mechanism to recognize its specific substrate, Ub.

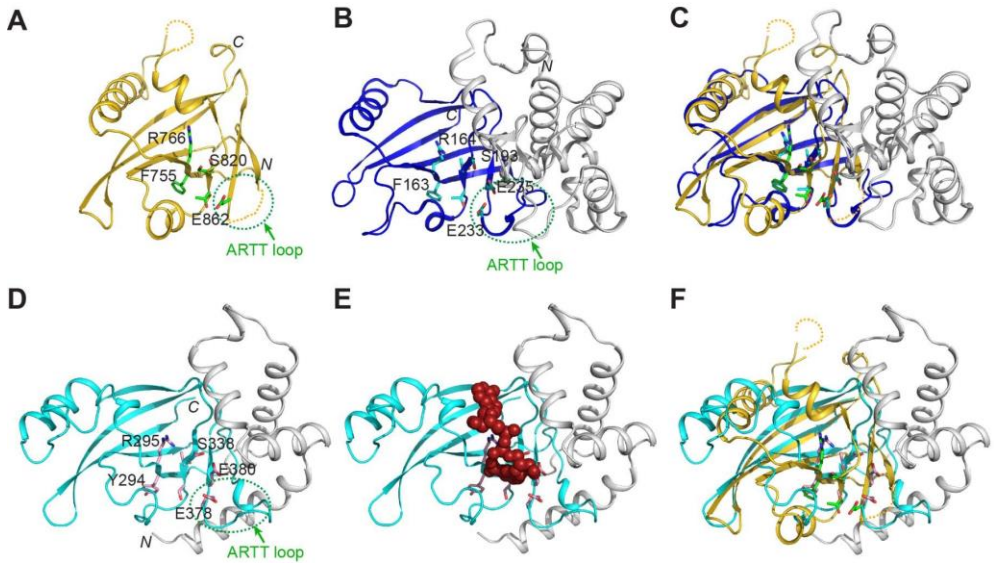


Figure 2-3: Structural Comparison of the mART Domain with Other mART Domains from Bacterial Toxins. (A) Ribbon diagram model of the mART domain of SdeA. (B) Ribbon diagram of HopU1 from *Pseudomonas syringe* (PDB ID: 3u0j). (C) Structural superimposition of the mART domains from SdeA (gold) and HopU1 (blue). (D) Ribbon diagram of Iota-toxin from *Clostridium perfringens* (PDB ID 4H03) and (E) Iota-toxin in complex with NAD (red spheres). (F) Structural overlay of the mART domains from SdeA (gold) and Iota-toxin (cyan). Residues comprising the catalytically important (F/Y)-(R/H), STS, and EXE motifs are shown in sticks and ARTT loop is indicated by a dash-lined circle. Note the structural core of the three mART domains have a similar fold, however, both HopU1 and Iota-toxin have additional alpha helical regions (colored in gray) while the mART domain of SdeA lacks such an equivalent structural component.

In addition to the PDE and mART domains, a third domain (a.a. 593-752), which we named the Mid domain, resides between the PDE and mART domains (Figure 2-2). The Mid domain contains 8 alpha helices of variable length. Interestingly, the longest alpha helix occupies the center of the domain forming a central axis around which the other 7 shorter alpha helices pack with different orientations (Figure 2-2 A and B). A structural homology search with the Dali server using the Mid domain yielded no hit above the threshold of meaningful similarity. A primary sequence analysis revealed that the Mid domain is only found in SdeA-related proteins upstream of the mART domain. These observations suggest that the Mid domain may have a unique function

relevant to the mART activity. Moreover, the Mid domain is connected to the core of the protein through two long flexible peptides (Figure 2-2). The observed conformation of the Mid domain relative to the rest of the protein is likely due to contacts in crystal packing. The flexibility of the Mid domain indicates that it can "communicate" with the rest of the protein by assuming different conformations. As discussed above, just as the mART domains of HopU1 and Iota-toxin have an N-terminal alpha-helical region that mediates substrate recognition, it is plausible that the Mid domain may function to recruit specific substrates for SdeA.

The Mid Domain of SdeA is Indispensable for ADP-ribosylation of Ub

We have shown that the Mid domain is required for the ADP-ribosylation of Ub. An intriguing question emerges as to how the Mid domain contributes to the modification of Ub. As proposed above, the Mid domain may function to recruit and present Ub to the mART domain, similar to the alpha helical regions found in other mART proteins. We first analyzed the sequence conservation of SdeA with the ConSurf server [26]. It is striking to note that the primary sequence of the Mid domain is highly conserved and that the most conserved residues form a prominent surface patch containing residues N723, Q727, and R729 (Figure 2-4 A). To further analyze the functional role of this conserved surface patch, we created single alanine mutations of these residues, as well as a nearby non-conserved F719 and a conserved residue D622 away from the patch. We also created an SdeA truncation with the entire Mid domain removed (ΔMid), and then assayed these proteins for enzymatic activity. As expected, ΔMid showed no activity in modifying Ub (Figure 2-4 B) while D622A displayed an activity level comparable to wild type SdeA. The F719A mutant, which is not conserved but is close to the conserved surface patch showed a significant impairment in producing ADPR-Ub. Strikingly, Ub ADP-ribosylation activity was completely abolished in SdeA carrying either of the N723A, Q727A, or R729A mutations. All the

aforementioned mutant proteins have similar expression levels and size-exclusion chromatographic behaviors comparable to wild type SdeA (Figure 2-4 C) and their PDE domains are fully functional in the conversion of ADPR-Ub to PR-Ub.

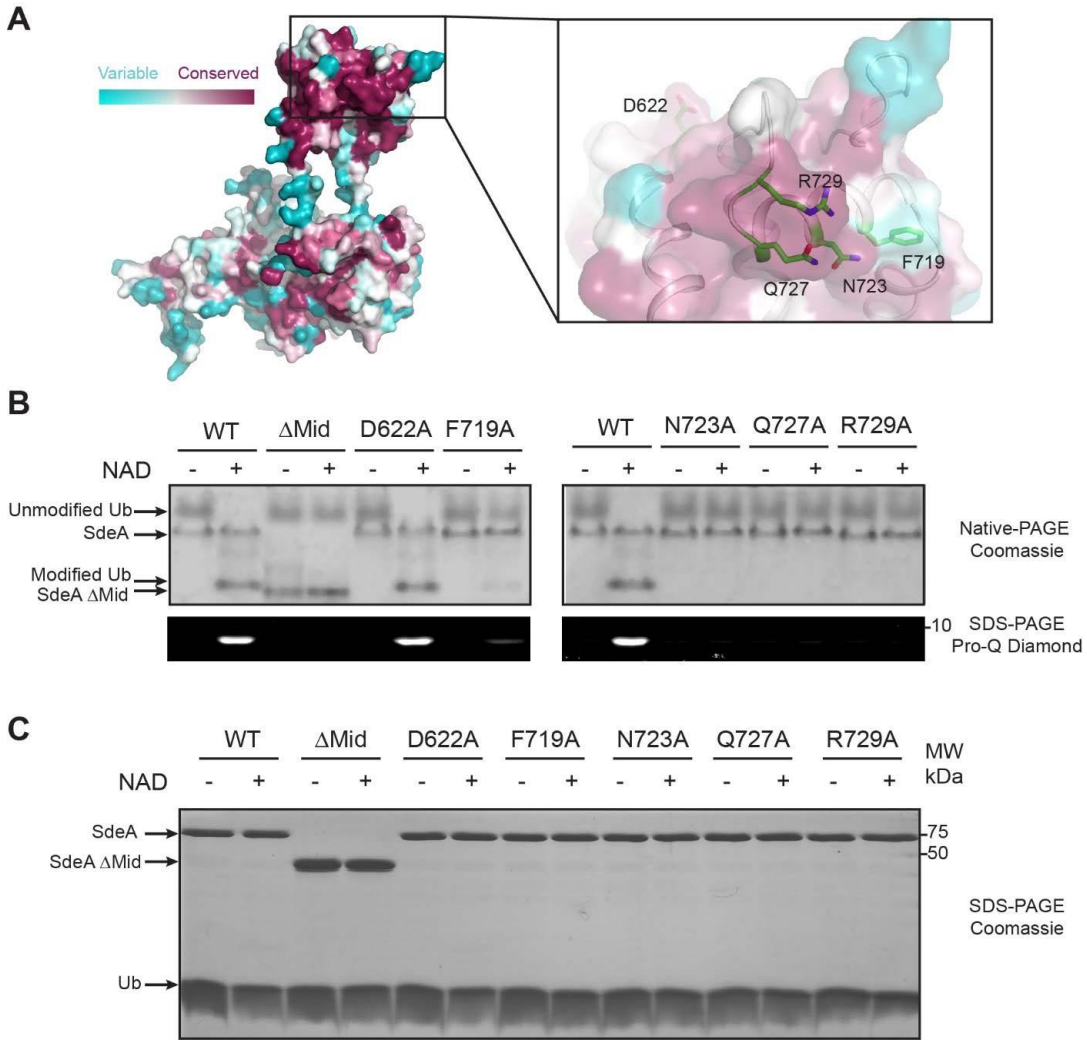


Figure 2-4: The Mid Domain of SdeA Is Indispensable for Ub ADP-ribosylation. (A) Sequence conservation of SdeA using a molecular surface representation (the most conserved residues are in purple and the least conserved residues in cyan). Sequence conservation is calculated by the ConSurf server (consurf.tau.ac.il). A zoomed-in view of a surface cluster comprised of the most conserved residues on the Mid domain. (B) Quantitative analysis of in vitro ubiquitin modification assays by SdeD mutants carrying mutations at its middle domain. The

reaction products were analyzed by native-PAGE and stained with Coomassie (top panel) and Pro-Q phosphoprotein stain (lower panel). (C) SDS-PAGE analysis of the proteins in the reaction mixture.

Taken together, our data show that the Mid domain is indispensable for the Ub ADP-ribosylation activity, and a surface patch which comprises of highly conserved residues on the Mid domain is crucial for this activity.

ADP-ribosylation of Ub and Serine Ubiquitination Are Two Independent Activities of SdeA

Our structure of SdeA revealed that the PDE and mART domains are packed against each other (Figure 2-2). It is worth noting that the catalytic sites of the PDE and mART domains are facing opposite directions with a straight linear distance of more than 55 Å between them. The geometric separation of these two catalytic sites raises an interesting question: How is the product from the mART, ADPR-Ub, loaded to the catalytic site of the PDE domain?

To address whether the catalytic activities of the mART and PDE domains are coordinated, we first set out to define the minimum region on SdeA that can ADP-ribosylate Ub. We generated a series of SdeA truncations and analyzed their ADP-ribosylation activity. Wild type SdeA spanning residues 211 to 910 efficiently modified Ub, as visualized on a native- PAGE gel (Figure 2-5). SdeA (563-910), which contains the last alpha helix (designated as αT) from the PDE domain, the Mid, and mART domains, could also modify Ub. However, SdeA (592-910), which spans the entire Mid and mART domains, was completely inactive, as was the mART domain alone (a.a. 752-910). These results demonstrate that both αT and the Mid domain are indispensable for ADP-ribosylation of Ub. The C-terminal portion of αT is enriched with hydrophobic residues and forms intensive hydrophobic and Van der Waals interactions with a hydrophobic patch on the outer surface of the mART β-sandwich core (Figure 2-5 A). It is possible that αT helps with structural

stability of the mART domain. However, it is also possible that αT provides a stable anchor point for the Mid domain, which can facilitate proper domain motions that may be required during the Ub ADP-ribosylation reaction (discussed below).

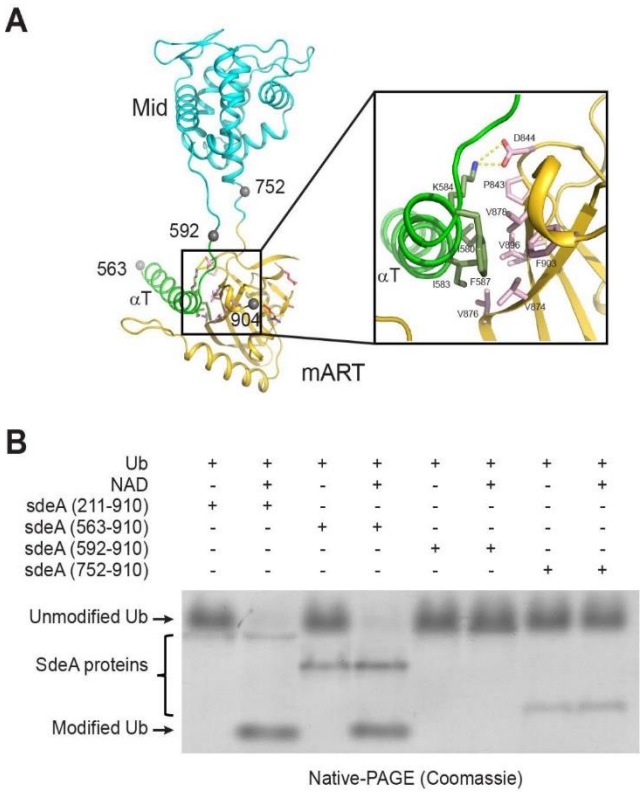


Figure 2-5: Mapping the Minimum Functional Region on SdeA Containing the mART Activity. (A) Ribbon diagram of minimum region that retains the ADP-ribosylation activity. This minimum region contains the last alpha helix (αT) from the PDE domain and the entire Mid and mART domain. αT packs against one side of the β sandwich of the mART domain though intensive hydrophobic interactions (zoomed-in view). Spheres labeled with residue numbers mark the boundary of constructs used for ADP-ribosylation activity assays. (B) ADP-ribosylation assays of Ub with indicated SdeA proteins. The reaction products were examined by native-PAGE.

We next set out to test whether the activities of the mART and PDE domains are interlinked. SdeA (211-910) was shown to convert Ub to ADPR-Ub and subsequently to PR-Ub,

as visualized both on the native-PAGE gel (the band labeled with modified Ub) and phosphoprotein specific Pro-Q Diamond stain due to the exposed phosphoryl group on PR-Ub [27]. In contrast, a SdeA (211-910) H227A mutant, which retains the ability to generate ADPR-Ub, is defective in processing ADPR-Ub to PR-Ub. Consequently, Ub band-shift caused by this mutant was evidently on the native-PAGE gel but no PR-Ub signal could be detected with the Pro-Q stain (Figure 2-6 B). Parallel reactions in which both SdeA (211-910) H227A and SdeA (211-592) (containing only the PDE domain) were added successfully converted Ub to PR-Ub. Similarly, reactions containing both SdeA (211-592) and SdeA (563-910) were also competent to generate PR-Ub. This data suggests that the conversion of Ub to PR-Ub can be achieved even when the mART and PDE activities are in two separated peptides. The independence of these two activities was further confirmed by SdeA-mediated ubiquitination of a substrate, Rab33b [6]. We observed that Rab33b was ubiquitinated to a similar extent compared to wild type SdeA when the PDE and mART activities were in two separate peptides (Figure 2-6 C). Interestingly, the SdeA PDE domain alone was able to ubiquitinate Rab33b comparable to wild type SdeA in the presence of purified ADPR-Ub (Figure 2-6 D). This data demonstrates that the PDE domain can use freely diffusing ADPR-Ub to ubiquitinate its substrates. To validate the independence of the two activities in vivo, we co-transfected Flag-tagged Rab33b and GFP-tagged SdeA constructs in HEK 293T cells. Ubiquitination of Rab33b was observed when plasmids expressing the PDE (211-592) and the portion containing the mART activity (563-910) were co-transfected (Figure 2-6 E). These results further support our conclusion that ADP-ribosylation of Ub and serine phosphoribosylated ubiquitination are two mechanistically independent activities.

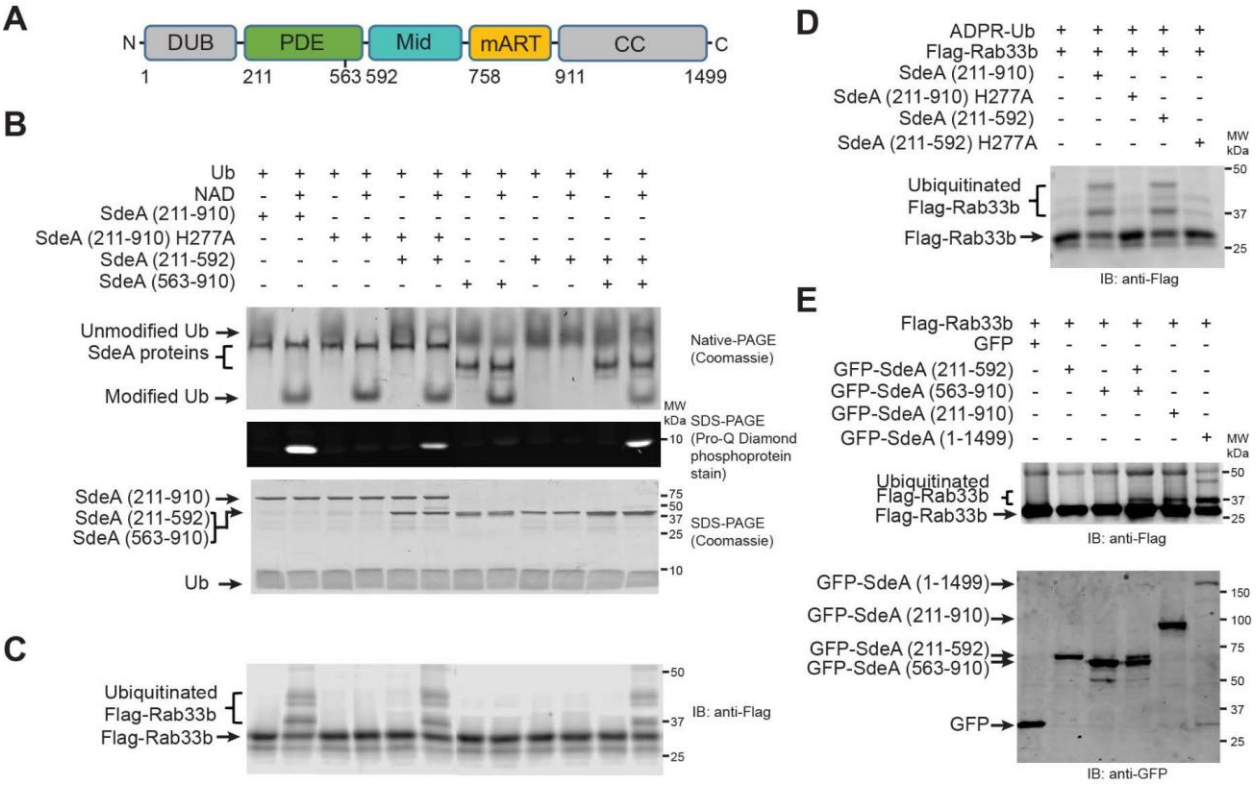


Figure 2-6: Phosphoribosylation of Ub and Serine Ubiquitination Are Two Independent Activities of SdeA. (A) Schematic diagram of the domain structure of SdeA. Some of the Sde family proteins are comprised of an N-terminal de-ubiquitinase (DUB) domain (gray), followed by a PDE (green), Mid (cyan), mART (gold), and a C-terminal coilded-coil (CC) domains (gray). (B) In vitro Ub-modification assays. Indicated SdeA proteins were mixed with Ub in the presence or absence of NAD. The modification of Ub to ADPR-Ub or PR-Ub can be monitored by the indicate band shift of Ub in native-PAGE (top panel). The production of PR-Ub can be visualized by phosphoprotein staining with Pro-Q Diamond stain (middle panel). All the proteins in the reaction mixtures were subjected to SDS-PAGE and visualized by Coomassie staining (lower panel). (C) In vitro ubiquitination assay of Rab33b by indicated SdeA proteins. Each reaction condition was the same as the corresponding lane in (B) except for the addition of recombinant Flag-tagged Rab33b. The reaction products were analyzed by anti-Flag Western-blot. (D) In vitro ubiquitination assay of Rab33b in the presence of purified ADPR-Ub. The reactions were performed with indicated SdeA proteins, Flag-Rab33b, and purified ADPR-Ub. The ubiquitination of Rab33b was analyzed by anti-Flag Western-blot. (E) In vivo ubiquitination assays of Rab33b by SdeA. Constructs expressing Flag-Rab33b, GFP or 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 probed with anti-Flag Western-blot (top panel). The expression of GFP-SdeA constructs were analyzed by anti-GFP Western-blot (Bottom panel).

Ub Recognition by the PDE Domain

We have shown that the isolated PDE domain of SdeA was able to process ADPR-Ub to PR-Ub and ubiquitinate its substrate proteins (Figure 2-6). This PDE domain has 23% sequence similarity to its counterpart in *Pseudomonas aeruginosa* PA4781, a well-characterized cyclic di-3',5'-GMP phosphodiesterase [28]. This lead us to ask how the PDE domain of SdeA can catalyze such a highly unusual reaction. To address this question, we first tested 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 2-7). The PDE domain of SdeA showed no detectable interaction with Ub, while the PDE domain of another SidE family member, SdeD displayed tighter interaction with Ub. SdeD was originally categorized as a member of the Sde family due to its high sequence homology in the PDE domain [9].

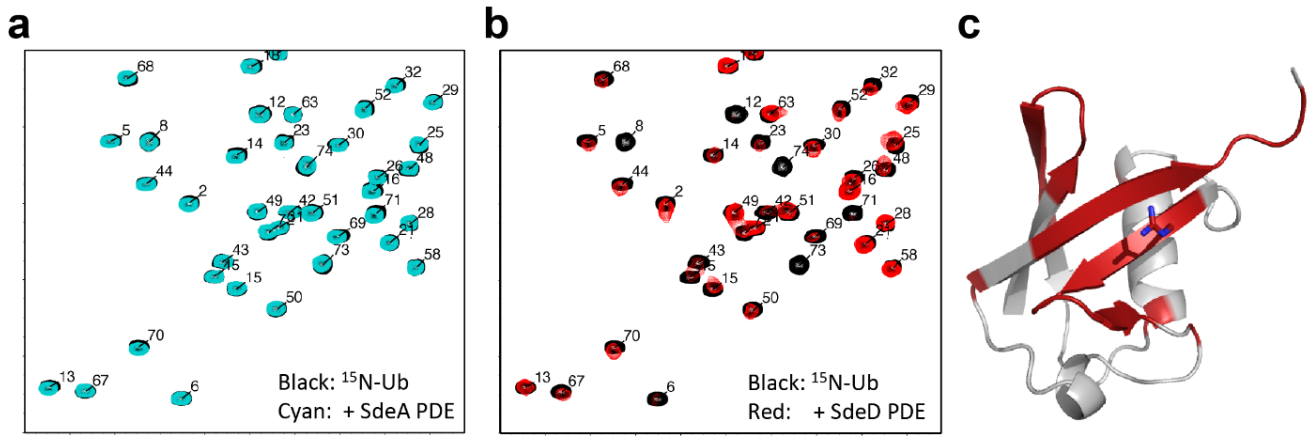


Figure 2-7: NMR titration analyses of the interaction between Ub and the PDE domains of SdeA and SdeD. (A) Spectral overlay of 150uM Ub (black) in the absence or presence of 300uM SdeA PDE (cyan) Ub binds weakly to SdeA as manifested by minimum changes of H-N peaks of

Ub. (B) Spectral overlay of 150uM Ub (black) with 75uM SdeD PDE. Ub binds much tighter to SdeD resulting in peak broadening and disappearance of resonances. (C) The most effected resonances in the presence of SdeD are mapped in red on a ribbon structure of Ub.

Through extensive crystallization trials, we succeeded in obtaining crystals of the PDE domain of SdeD both by itself and in complex with Ub.. The structure of the apo SdeD PDE domain was solved by selenomethionine SAD phasing. The overall structure of the SdeD PDE is very similar to that of SdeA with an rmsd of 1.73 Å over 251 overlaid Cα atoms. However, a major difference between SdeA and SdeD, similar to that between SdeA and lpg1496, resides on the size and position of the loops connecting the alpha helices (indicated with dashed-circles, Figure 2-8). Nevertheless, in all the PDE domains, the active site, containing the most conserved residues including two catalytic histidines, is located in a deep V-shaped groove. The groove is open at one end to a large and rather flat surface, which is nearly perpendicular to the groove (Figure 2-8).

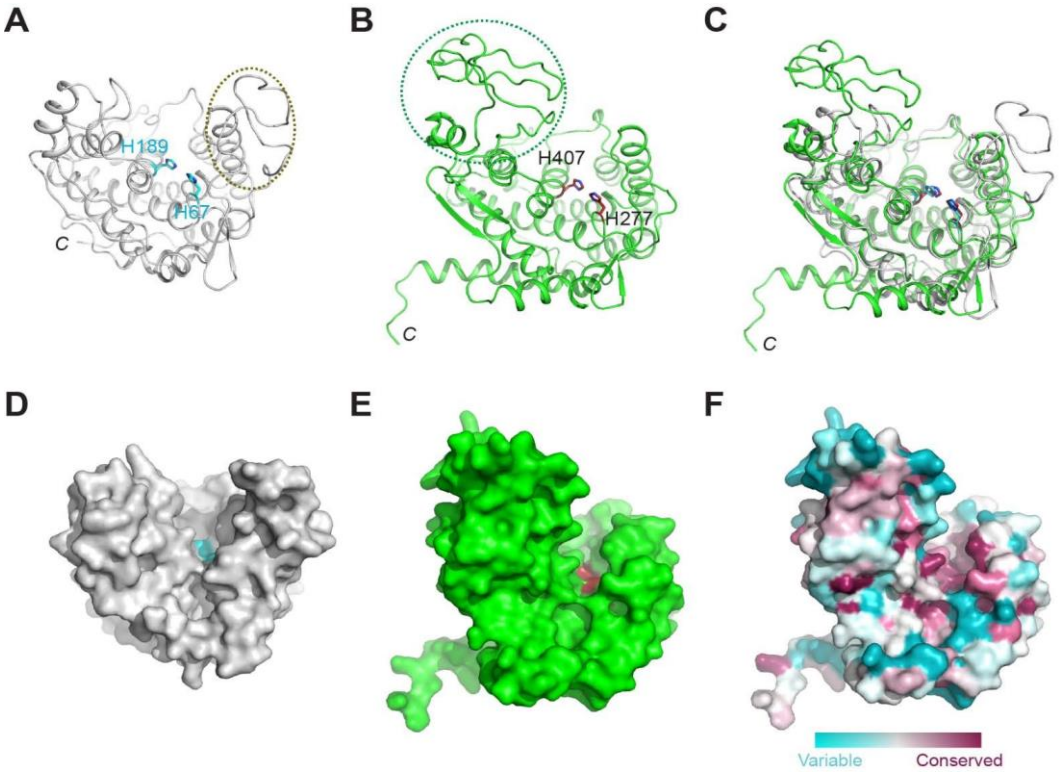


Figure 2-8: Overall Structure of the PDE Domain of SdeD and its Comparison with the PDE Domain of SdeA. (A) Ribbon model of the PDE domain of SdeD (gray). Two catalytic histidine residues shown in sticks (cyan) (B) The PDE Domain structure of SdeA. (C) Structural overlay of these two PDE domains. The variability unique to each PDE domain is indicated by a dash-lined circle. (D) Surface representative of the PDE domain of SdeD. The catalytic histidines are located at the bottom of a deep V-shaped groove. (E) Surface representative of the PDE domain of SdeA. (F) Surface residue conservation analysis of the PDE domain. The most conserved residues are colored in purple and the least conserved residues are in cyan.

We also solved the structure of the SdeD PDE domain in complex with Ub. To our surprise, we observed two Ub binding sites on the SdeD PDE domain. The first Ub (Ub1) binds on the aforementioned flat surface at the opening of the groove (Figure 2-9). A total of 1374 $Å^2$ surface area is buried at the Ub1 and SdeD interface. Three major regions of Ub1 are involved in the binding, including the loop region around T9 (T9 loop), the C-terminal end, and the vicinity of R42. At the T9 loop, L8 and the methyl group of T9 form hydrophobic interactions with SdeD while the T9 hydroxyl group and K6 form hydrogen bonds and electrostatic interactions with corresponding residues on SdeD (Figure 2-9 A). The C-terminus of Ub1 is also extensively engaged in the binding, which includes hydrophobic interactions mediated by L73, hydrogen bonding and electrostatic interactions contributed by R72 and the carboxyl terminus of G76 (Figure 2-9 A). The third interacting region on Ub1 involves the residues in the vicinity of R42. Most strikingly, the R42 sidechain of Ub1 sticks into the groove and forms hydrogen bonds and electrostatic interactions with the conserved residues Q52 and E126 at the catalytic site. It is very likely that the binding between Ub1 and SdeD represents the physiological interaction between Ub and the PDE domain of SdeA/SdeD. With this mode of binding, the ADPR moiety would be positioned into the groove ready for the reaction at the catalytic site. Ub1 buries a large surface

area on the PDE domain (Figure 2-9 A). This surface area is rather flat and nearly perpendicular to catalytic groove. Although the overall shape of this surface is also preserved in the PDE domain of SdeA (Figure 2-8 E), residues comprise this surface area are not well conserved (Figure 2.8 F). The large number of residues involved in close contact with Ub and the sequence variability of those contact sites hampered our attempts to design point mutations to dissect the roles of these residues on Ub binding. We thus focused on key Ub to investigate the recognition of Ub by the PDE domain.

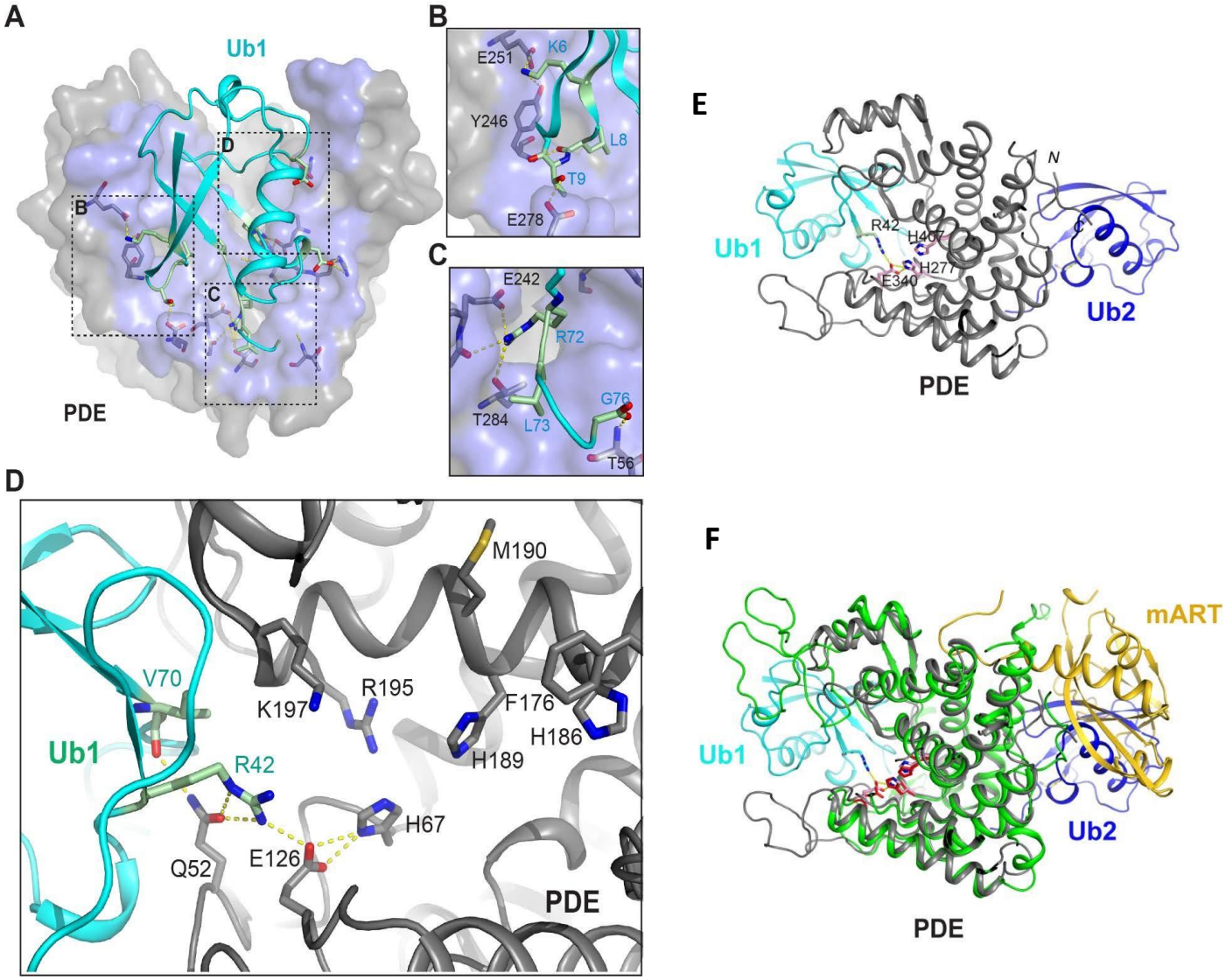


Figure 2-9: The Interaction of Ub with the PDE Domain of SdeD. (A) Overall complex structure of Ub (Ub1) with the PDE domain of SdeD. Ub1 is shown in cyan-colored ribbon. The PDE domain is shown in gray surface. The area that is within in the Van-der Waals distance range to Ub1 is colored in light blue. Ub1 binds at the opening of the PDE catalytic groove. Three major regions on Ub1 involved in the interaction with the PDE domain are marked by dash-lined squares. (B) A zoomed-in view of the T9-loop region of Ub1 interacting with the PDE domain. (C) A zoomed-in view of the C-terminus of Ub1 at the interface with the PDE domain. (D) A close view of the region near the Ub1 R42. The sidechain of the R42 of Ub1 sticks into the catalytic groove and forms hydrogen bonds and salt bridges with Q52 and E126. Conserved residues lining in the catalytic groove are shown in sticks. (E) The PDE domain of SdeD is bound with two Ub molecules Ub1 (cyan) and Ub2 (blue). Ub1 binds at the opening of the PDE catalytic groove with its R42 sidechain sticking into the groove. Ub2 binds a region on the opposite side of the PDE domain. (F) Structural superimposition of the complex structure with the structure of SdeA referenced on the PDE domain

The second Ub (Ub2) binds on the opposite side of SdeD (Figure 2-9 E) and occludes an area of 764 $Å^2$ at the interface. The physiological significance of this binding site is not clear. This binding site certainly does not exist in SdeA, as it would collide with the mART domain (Figure 2-9 F). It is interesting to note that the interaction between Ub2 and SdeD is also mediated mainly by the T9 loop and the carboxyl terminus of Ub2. A structural comparison of Ub1 and Ub2 revealed that the conformation of all the residues is nearly identical except for the T9 loop and the carboxyl terminus. These observations suggest that the conformational flexibility of the T9 loop and the carboxyl terminus bestows Ub with the versatility to interact with totally different protein surfaces using the same group of residues.

ADPR-Ub Binding to the Active Site of PDE

Our complex structure of SdeD with Ub revealed the mode of interaction between Ub and the PDE domain of SdeD, but how exactly the ADPR moiety fits in the groove remained elusive. To address this question, we sought to determine the complex structure of the SdeD (H67A) mutant with ADPR-Ub (generated by reacting Ub and NAD with mutant SdeA retaining the mART but not the PDE activity). Although we obtained crystals containing SdeD and ADPR-Ub, those crystals were of poor quality. We reasoned that the ADPR moiety in Ub2 may disrupt the crystal packing with adjacent molecules (Figures 2-9 E). We then tried to crystalize the SdeD PDE domain with 2:3 molar ratio of ADPR-Ub and Ub. We expected that ADPR-Ub would have a higher affinity for binding at the Ub1 site while wild type Ub would preferentially bind to the Ub2 site due to crystal packing contacts. Indeed, we obtained crystals with good diffraction quality using this strategy, and it was gratifying to observe that ADPR-Ub bound at the Ub1 site and wild type Ub bound at the Ub2 site.

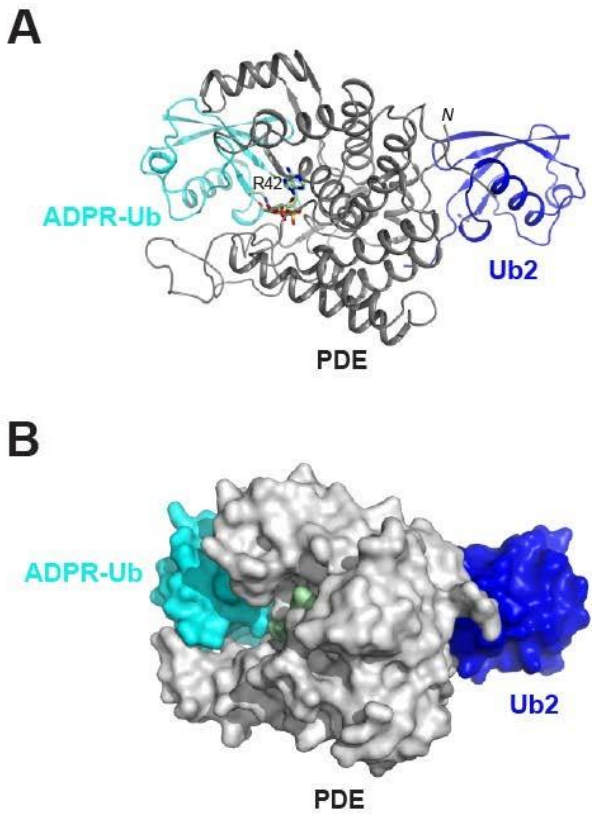


Figure 2-10: Overall Structure of the PDE Domain of SdeD in Complex with ADPR-Ub and Ub. (A) The PDE domain of SdeD is bound with both ADPR-Ub (cyan) and wild type Ub (Ub2 in blue). ADPR-Ub binds at the opening of the PDE catalytic groove with the ADPR moiety fitting in the catalytic groove. Ub2 binds a region identical to the Ub2 found in SdeD-Ub complex. (B) Surface representation of the complex structure shown in (A). Note that the ADPR-moiety shown in light green surface fits deeply in the catalytic groove.

ADPR-Ub engages a similar mode of binding as Ub1 with the SdeD PDE domain. The ADPR moiety extends into the large catalytic groove and resides on top of the two catalytic histidine residues, which are lined at the bottom of the groove (Figure 2-11 A and B). The adenine ring of the ADPR moiety is sandwiched between the hydrophobic sidechains of M190 and F176. The hydroxyl groups of the two pentose moieties form hydrogen bonds with K197, H186, and Y119. Two arginine residues, R123 and R195, accommodate the β phosphate group by forming salt bridges (Figure 2-11 C). Although the PDE domain in this complex carries the H67A mutation, based on the PDE-Ub complex structure, we can easily model back in the histidine in silico. Presumably, H67 would assume a similar conformation as seen in the PDE-Ub complex

and it, together with other key catalytic residues E126 and H189, is lined at the bottom of the groove within attacking distance to the β phosphate group of ADPR.

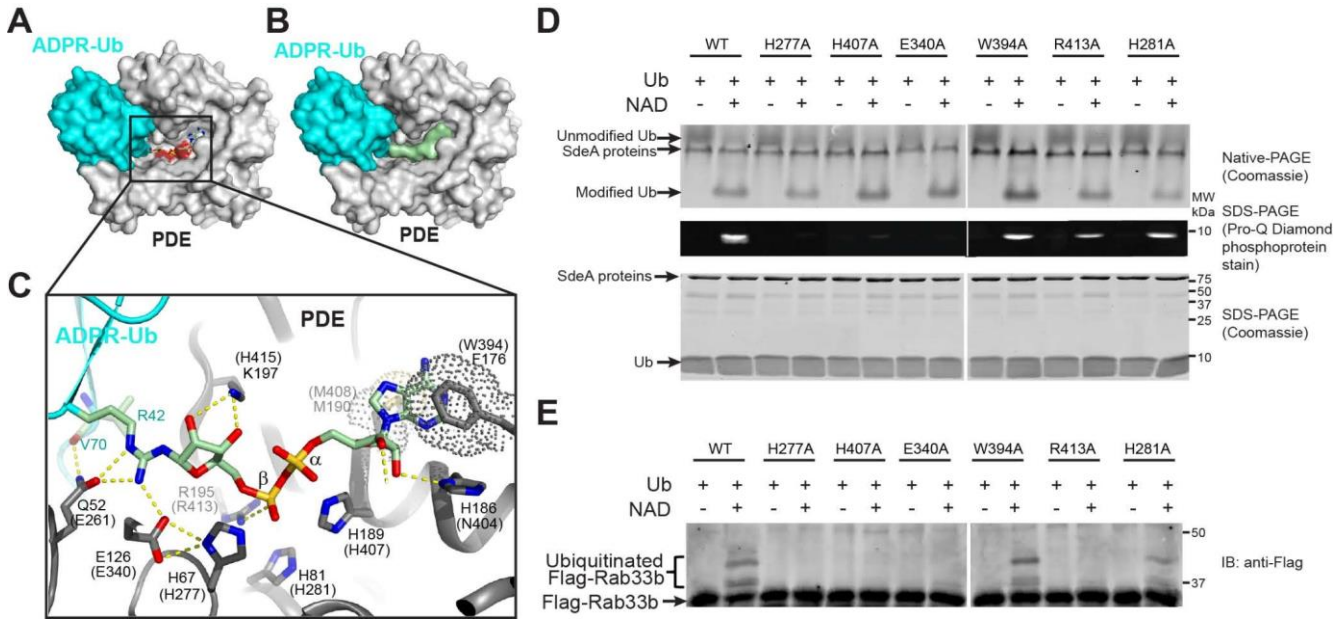


Figure 2-11: Complex Structure of ADPR-Ub with the PDE Domain of SdeD. (A) Surface representation of ADPR-Ub (cyan) in complex with the SdeD PDE domain (gray). The catalytic histidine residues (H67, H189) are colored in red. The ADPR moiety is shown in sticks and colored in light green. (B) Same view of the complex structure as in (A). The ADPR moiety is represented as a light green surface. Note that the catalytic site is completely sequestered from the bulk solvent when ADPR-Ub is docked. (C) A detailed interaction of the ADPR moiety within the catalytic groove of the PDE domain. Residues of the SdeD PDE domain that are involved in binding of the ADPR moiety are labeled and the corresponding residues in SdeA are labeled with parenthesis. Hydrogen bonds between ADPR-Ub and the PDE domain are specified by dashed yellow lines. The hydrophobic stacking of the adenine is depicted by dots. (D) Enzymatic activity analysis of conserved residues within the catalytic groove. All the procedures of the assay were like those described in Figure 2-6. (E) Rab33b ubiquitination assays by SdeA mutants carrying mutations on conserved residues within the catalytic groove.

Due to the structural similarity between the two PDE domains from SdeD and SdeA, we can easily model the binding of ADPR-Ub to the PDE domain of SdeA. Most of the residues in the catalytic groove that interact with ADPR are identical or functionally conserved between SdeD and SdeA. To test the role of the ADPR-interacting residues within the catalytic groove, we mutated several of those conserved residues in SdeA. Recombinant proteins (all containing residues 211-910) of the mutants were purified and their ability to modify Ub was analyzed (Figure 2-11 D and E). The PDE activity was completely abolished in the H227A, H407A, and E340A mutants as indicated by the lack of the Pro-Q phosphoprotein staining signal. However, the mART activity of these mutants was in intact as indicated by the Ub band-shift in native gels (Figure 2-11 D). In agreement with the defective PDE activity, these three mutants all failed to ubiquitinate Rab33b (Figure 2-11 E). The W394A mutant showed similar activity to wild type, suggesting the hydrophobic stacking of the adenine ring is not as critical. H281, which is completely invariable and is located near the β phosphate of ADPR, showed no obvious effect on activity when mutated to alanine. Conversely, the activity of R413A was significantly impaired, indicating the importance of proper positioning of the β phosphate for the reaction.

2.4 DISCUSSION

SdeD ADPR-Ub complex structure reveals that the large ADPR moiety occupies the entire catalytic groove and completely covers the three-essential catalytic resides E340, H277, and H281 and shields them from the bulk solvents (For simplicity, we only use the residue numbering in SdeA. Figure 2-11 a-c). This ADPR-Ub enzyme binding mode prohibits potential direct contacts of Ub acceptor substrates with the catalytic residues for activation. This observation suggests that the ubiquitination reaction must be initiated without the participation of Ub acceptor substrates. Based on our structural and biochemical analysis, we propose that the conjugation of PR-Ub to

serine residues of substrate proteins is a two-step reaction (Figure 2-12). In the first step, the ADPR moiety is nestled in the catalytic groove upon the binding of ADPR-Ub to the PDE domain of SdeA. The negatively charged E340 stabilizes R42 from ADPR-Ub and also forms hydrogen bonds with H277. Thus, E340 increases the pKa of H277 and renders H277 a stronger nucleophile. H227 attacks the β phosphate of ADPR and forms a transient covalent phosphor-histidine linkage. The nearby H407 functions as a general acid to donate a proton to the α phosphate of the releasing AMP molecule. As a consequence of this step, a transient PR-Ub-enzyme complex is formed via a mechanism similar to the that of histidine phosphorylation, which is well established as a critical component of cell signaling and metabolism in prokaryotes [29]. Enzymes, such as the two-component system histidine protein kinase, form a transient phosphor-histidine intermediate during the reaction [30]. In the second step, H407 deprotonates the hydroxyl of a serine residue of the approaching substrate. The activated hydroxyl then replaces H277 in the transient PR-Ub-enzyme intermediate by attacking the phosphoryl group to form a stable phosphor-serine linkage between the substrate protein and PR-UB. In the final stage of the reaction, E340 functions as a general acid and protonates the released H277 and thus recycles the enzyme to its initial state. According to this mechanism, if a water molecule is used as the Ub acceptor in the second step, the reaction simply results in the cleavage of ADPR-Ub to PR-Ub.

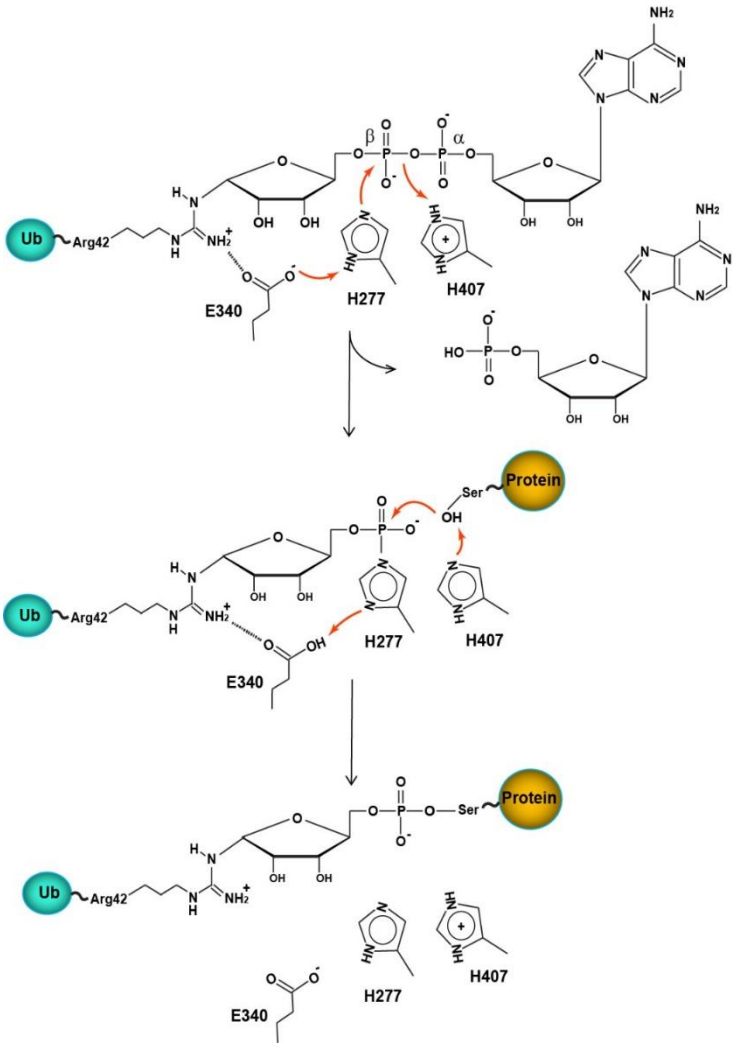


Figure 2-12. A Mechanistic Model of Serine Ubiquitination catalyzed by the PDE Domain of SdeA. A two-step reaction model of SdeA catalyzed serine ubiquitination is proposed here. The first step involves the formation of a transient PR-Ub~enzyme intermediate and the cleavage of the phosphoanhydride bond of ADPR-Ub. The second step involves the formation of a phosphoester bond between PR-Ub and a substrate serine residue.

Our data showed that the ubiquitination of Rab33b is overall fairly inefficient, which is in agreement with previously reported results that the cleavage of ADPR-Ub is apparently much more efficient than the conjugation of PR-Ub to substrate proteins [7]. This observation can be explained by the accessibility of the catalytic groove. A small water molecule is much easier to access the catalytic site located in a deep groove than a bulky protein substrate. Even for protein substrates, it is likely that only the serine residues located at protruding loops are able to access the catalytic site on the PDE domain. In fact, Residue S154 of Rab33b, which has been shown to be the ubiquitinated by SdeA [7], is located at the tip of a protruding loop on Rab33b [31].

SdeD and SdeA, in comparison, share a very similar fold. The major difference is the location of two different loops that cover the active cleft of these PDE domains: the loop between residues 26 and 48 of SdeD is missing within SdeA PDE domain and the loop between residues of 465 and 513 of SdeA is completely gone within the SdeD PDE domain. The difference of location of these loops in PDE domains might be the underlying reason why SdeD can bind to Ub much tighter than SdeA and how SdeA might recognize target proteins and SdeD can not. In addition to the loop difference, the ubiquitin recognition surface of SdeD is not conserved in SdeA and other large SidE family members. As discussed in more detail in Chapter 3, all these PDE domain differences might be important for the consecutive activity of these proteins during infection.

The presence of a Mid domain was not expected from previous bioinformatical predictions. Regarding the placement of the domain between two other functioning domains, we first predicted it to have a carrier function. The function independence experiment and mutational analysis on the Mid domain itself showed the importance of the domain only for the first step of ubiquitin modification. The evolutionary reason behind the presence of this Mid domain is not clear, because the domain's fold is not conserved among other protein structures. When we searched for a similar fold with Dali Server, we could not find a match with a score above the threshold. Also, when we compare the mART domain of SdeA to other mART domain containing bacterial effector proteins, we observe the target recognition motif of other effectors is missing in SdeA. Evolutionarily, *Legionella* might have obtained a fold that recognizes ubiquitin and over time this domain might have evolved to this unique fold. The essentiality of this domain can be tested by using the Mid domain deleted full-length construct for *Legionella* infection reactions.

SdeA protein has been shown to be highly toxic with previous yeast toxicity assays [32]. The reason behind the toxicity of the protein is prone to discussion: the toxicity might be either

because of the depletion of available ubiquitin to the conventional ubiquitination pathway or the PR-ubiquitination of the target proteins and their inactivation. The target inactivation by novel ubiquitination seems not to be real for Rab33b molecule since the protein's GTPase activity is not lost upon PR-ubiquitination *in vitro* [33]. The exact effect of the PR-ubiquitin modification needs to be elucidated thoroughly. The targets, upon modification, might change localization or their interaction with other proteins. These possibilities can be observed both *in vivo* with mammalian cell overexpression of SdeA and monitoring target proteins or even observing the target proteins during infection conditions. The possibility of ubiquitin depletion can be tested by looking at the modified levels of ubiquitin during infection. Ubiquitin levels don't fluctuate to a big degree within cell homeostasis [34], so the modified ubiquitin might be present only for a short period of time during infection. A pulse-chase experiment of ubiquitin modification might shed more light on this depletion possibility.

Given the fact that many *Legionella* effector proteins have eukaryotic origins evolutionarily[35], it is likely that eukaryotes also harbor an equivalent machinery to perform PR-ubiquitination. Future elucidation of such a eukaryotic enzyme system is of fundamental significance to our understanding of the versatile Ub code.

2.4 MATERIALS AND METHODS

Cloning and Mutagenesis.

DNA fragments encoding SdeA (aa. 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. DNA fragments encoding wild type or mutant Ub were PCR amplified and subcloned into a pET21a vector. All constructs were confirmed by DNA sequencing.

Protein Expression and Purification.

For expression of SdeA and SdeD proteins, E. coli BL21-DE3 strains harboring the expression plasmids were grown in Luria-Bertani medium supplemented with 50 μg/ml kanamycin to mid-log phase. Protein expression was induced for overnight at 18°C with 0.1 mM isopropyl-B-D-thiogalactopyranoside (IPTG). Harvested cells were resuspended in a buffer containing 20 mM Tris-HCl (pH 8.0), 150 mM NaCl) and were lysed by sonication. Soluble fractions were collected after centrifugation at 16,000 rpm for 30 min at 4°C and incubated with cobalt resins (Clonetech) for 1.5 h at 4°C. Protein bound resins were extensively washed with lysis buffer. The SUMO-specific protease Ulp1 was then added to the resin slurry to release SdeA proteins from the His-SUMO tag. Eluted protein samples were further purified by FPLC size exclusion chromatography. FPLC fractions corresponding to the protein peak were pooled and the purified proteins were concentrated to 10 mg/ml. Protocols for Ub expression and purification were adapted from publish literatures [36]. Briefly, harvested cells were resuspended in a buffer containing 20

mM ammonium acetate (pH 5.1). Cells were lysed by sonication and cell lysate was clarified by 18.000 rpm centrifugation for 30 min. Clarified supernatant was pooled and was titrated gradually with acetic acid to lower the pH to 4.8. Precipitated proteins after the titration were removed by centrifuged at 18.000 rpm for 30 min. The pH of the clarified supernatant was adjusted to 5.1 by NaOH addition. HiTrap SP column (GE Healthcare) was used for cation exchange with a buffer gradient from 20 mM ammonium acetate to 0.5 M ammonium acetate (both pH: 5.1). Fractions containing ubiquitin peak were pooled and was 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 prepare ADPR-Ub for both biochemical assays and crystallographic trials, 1 μM of SdeA (211-910) H277A (which lacks PDE activity, see Figure 5D) 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.

Protein Crystallization.

Generally, all protein crystallization screens were performed with a Crystal Phoenix liquid handling robot (Art Robbins Instruments) at room temperature. The crystallization, which yields 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 crystallization, SdeA was concentrated to 12 mg/ml and crystallized in 100 mM HEPES pH 7.9, 12% PEG 8000. Thin-plate shaped SdeA 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. The SdeD/Ub/ADPR-Ub co-crystal was formed by mixing a catalytically inactive mutant SdeD (a.a. 1-341, H67A) with ADPR-Ub and WT-Ub in a 1:2:3 molar ratio, with a final SdeD concentration of 12 mg/ml. Rod-shaped crystals appeared in 100 mM sodium cacodylate pH 6.7 and 21% PEG 8000.

X-ray Diffraction Data Collection, and Processing.

Diffraction data sets for SdeA, the SdeD-Ub complex, and the SdeD/Ub/ADPR-Ub complex data sets were collected at Cornell synchrotron light source MacChess beam line F1, while data sets for SdeD crystals were collected at the A1 beam line (Table 2-1). Before data collection, all crystals were soaked in cryo-protectant solutions containing their respective crystallization condition buffer supplemented with 20% glycerol and flash frozen in liquid nitrogen stream. All data sets were indexed, integrated and scaled with HKL-2000 [37].

A. Data collection statistics				
Protein name	SdeA	SdeD	SdeD-Ub	SdeD-ADPRUB-Ub
Space group	$P2_1$	R3	$P2_1$	$P2_1$
Cell dimensions	a = 69.8 Å, b = 80.6 Å, c = 85.6 Å, $\alpha = 90^0$, $\beta = 109.8^0$, $\gamma = 90^0$	a= 154.4 Å b= 154.4 Å c= 89.6 Å, $\alpha = 90^0$, $\beta = 90^0$, $\gamma = 120^0$	a= 64.8 Å b= 58.6 Å c= 74.1 Å, $\alpha = 90^0$, $\beta = 114.6^0$, $\gamma = 90^0$	a= 64.7 Å b= 58.8 Å c= 75.1 Å, $\alpha = 90^0$, $\beta = 114.2^0$, $\gamma = 90^0$
Synchrotron beam lines	MCCHESS F1	MCCHESS A1	MCCHESS F1	MCCHESS F1
Wavelength (Å)	0.9789	0.68	0.9789	0.9789
Maximum resolution	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 (%)[a]	99.3	99.5	99.4	100
$\langle I \rangle / \langle \sigma \rangle$[a]	7.98 (0.87)	29.2 (1.52)	25.4 (1.54)	19.28 (1.18)
R_{sym}[a,b] (%)	0.122(0.759)	0.07(0.622)	0.078(0.798)	0.093(1.105)
B. Phasing				
Phasing methods	SAD	SAD	MR	MR
Heavy atom type	Hg	Se	-	-
Number of proteins/ASU	1	2	1 SdeD, 2 Ub	1 SdeD, 1 Ub, 1 ADPR-Ub
Number of heavy atoms/ASU	9	8	-	-
C. Refinement statistics				
Resolution (Å)[a]	80.51(2.20)	77.174(1.51)	67.36(1.70)	68.93(1.85)
R_{crys} / R_{free} (%)[a,c]	0.192/0.241	0.167/0.195	0.172/0.28	0.021/0.249
Rms bond length (Å)	0.023	0.027	0.03	0.028
Rms bond angles (°)	2.29	2.47	2.47	2.43
Ramachandran plot				
Most favored/allowed (%)	96.01/3.99	100/0	97.38/2.62	96.23/3.77
Generous/Disallowed (%)	0	0	0	0

[a]Values in parenthesis are for the 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.

Table 2-1: X-Ray Crystallography Data Collection and Refinement Statistics

Structure Determination and Refinement.

The structure of SdeA was solved by single wavelength anomalous dispersion (SAD) method. Before data collection, SdeA crystals were soaked in cryoprotectant (0.1 M Hepes pH 7.9, 12% PEG8000 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 initial phase was calculated using the program HKL2MAP [38]. The structure of the PDE domain of SdeD was solved by SAD phasing with seleno-methionine incorporated SdeD crystals. The structures of the SdeD-Ub and SdeD/Ub/ADPR-Ub complexes were solved by molecular replacement with the AMoRe program

[39] of the CCP4 suite [40], 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 [41] and refmac5 [42] of the CCP4 suite.

Computational Analysis and Graphic Presentation of Protein Sequence and Structure.

SdeA homologous sequences were selected from results generated by BLAST server (NCBI). Edited sequences were aligned with Clustal Omega [43] and colored by 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) [26]. All structural figures were generated using PyMOL (The PyMOL Molecular Graphics System, Version 1.8.X, Schrödinger, LLC).

Ubiquitin Modification and Rab33b Ubiquitination Assays.

Ubiquitin modification reactions were carried out by mixing 1 μM of SdeA 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 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. Rab33b ubiquitination reactions were performed by the addition of 4 μM of recombinant Flag-Rab33b to the ubiquitin 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.

2.5 REFERENCES

1. Hershko, A., A. Ciechanover, and A. Varshavsky, *Basic Medical Research Award. The ubiquitin system.* Nat Med, 2000. **6**(10): p. 1073-81.

2. Chen, Z.J. and L.J. Sun, *Nonproteolytic functions of ubiquitin in cell signaling.* Mol Cell, 2009. **33**(3): p. 275-86.

3. Hurley, J.H. and H. Stenmark, *Molecular mechanisms of ubiquitin-dependent membrane traffic.* Annu Rev Biophys, 2011. **40**: p. 119-42.

4. 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.

5. Komander, D. and M. Rape, *The ubiquitin code.* Annu Rev Biochem, 2012. **81**: p. 203-29.

6. Qiu, J., et al., *Ubiquitination independent of E1 and E2 enzymes by bacterial effectors.* Nature, 2016. **533**(7601): p. 120-4.

7. Bhogaraju, S., et al., *Phosphoribosylation of Ubiquitin Promotes Serine Ubiquitination and Impairs Conventional Ubiquitination.* Cell, 2016. **167**(6): p. 1636-1649 e13.

8. 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.

9. 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.

10. Scheffner, M., U. Nuber, and J.M. Huibregtse, *Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade.* Nature, 1995. **373**(6509): p. 81-3.

11. 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.

12. Wenzel, D.M., et al., *UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids.* Nature, 2011. **474**(7349): p. 105-8.

13. Ciechanover, A. and R. Ben-Saadon, *N-terminal ubiquitination: more protein substrates join in.* Trends Cell Biol, 2004. **14**(3): p. 103-6.

14. 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.

15. Qiu, J. and Z.Q. Luo, *Legionella and Coxiella effectors: strength in diversity and activity.* Nat Rev Microbiol, 2017.

16. 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.

17. 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.

18. Kubori, T., et al., *Legionella metaeffector exploits host proteasome to temporally regulate cognate effector.* PLoS Pathog, 2010. **6**(12): p. e1001216.

19. 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.

20. 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.

21. 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.

22. 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.

23. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D.* Nucleic Acids Res, 2010. **38**(Web Server issue): p. W545-9.

24. 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.

25. 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.

26. 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.

27. 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.

28. 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.

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

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

31. Eathiraj, S., et al., *Structural basis of family-wide Rab GTPase recognition by rabenosyn-5.* Nature, 2005. **436**(7049): p. 415-9.

32. 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.

33. Nakasone, M.A. and D.T. Huang, *Ubiquitination Accomplished: E1 and E2 Enzymes Were Not Necessary.* Mol Cell, 2016. **62**(6): p. 807-809.

34. Hershko, A. and A. Ciechanover, *The ubiquitin system.* Annu Rev Biochem, 1998. **67**: p. 425-79.

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

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

37. Otwinowski, Z. and W. Minor, *Processing of X-ray diffraction data collected in oscillation mode.* Methods Enzymol., 1997. **276**: p. 307-326.

38. 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.

39. Trapani, S. and J. Navaza, *AMoRe: classical and modern.* Acta Crystallogr D Biol Crystallogr, 2008. **64**(Pt 1): p. 11-6.

40. Collaborative Computational Project, N., *The CCP4 suite: programs for protein crystallography.* Acta Cryst., 1994. **D**(50): p. 760-763.

41. 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.

42. 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.

43. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Mol Syst Biol, 2011. **7**: p. 539.

CHAPTER 3

SdeD-SdeF DEUBIQUITINATES SdeA TARGETED PROTEINS TO REGULATE SdeA ACTIVITY

3.1 ABSTRACT

During infection, *Legionella* manipulates host cellular processes in a highly regulated manner. Ubiquitination, a post-translational modification that is involved in the regulation of various cellular pathways, is modulated by *Legionella* effector proteins for the continuation of infection. SidE family effector proteins have recently been shown to possess a novel ubiquitination mechanism that includes addition of a phosphoribose moiety to a ubiquitin molecule with their mono ADP ribosyltransferase (mART) and phosphodiesterase (PDE) domains. The SidE protein family consists of four large members that have both mART and PDE domains, but there are five smaller proteins that share the PDE domain alone. The function of this stand-alone PDE domain is not known. In this study, we observed that the small SdeD and SdeF proteins from SidE family do not target the proteins that SdeA does, but instead act as deubiquitinase (DUB) enzymes that remove the phosphoribosyl-ubiquitin from target proteins of SdeA. The DUB activity is shown both by a single target, Rab33b, or targets from the whole cell lysate. We have also observed a Golgi fragmentation phenotype when SdeA is present in mammalian cells. We have shown the reversion of this phenotype by SdeD-SdeF proteins. We have found a candidate for ubiquitination and Golgi fragmentation link, Syntaxin 5. We have shown the modification of Syntaxin 5 by SdeA and the DUB activity of SdeD-SdeF on this protein, both *in vivo* and *in vitro*. Our results suggest

a DUB activity of SdeD-SdeF proteins against the target proteins of large SidE protein family members. This DUB activity may be used by *Legionella* to revert the effects of SidE proteins to create a distinctive regulation of activity before and after the formation of *Legionella*-containing vacuole (LCV).

3.2 INTRODUCTION

Ubiquitination is a posttranslational modification that is involved in a wide range of pathways including protein degradation, inflammation, endocytosis, histone regulation, DNA repair and cell cycle [1-4]. The conventional ubiquitination signal cascade follows a three-step reaction: a ubiquitin activation step by E1 enzyme, a ubiquitin conjugating step by E2 enzyme, and a ubiquitin ligation step by an E3 enzyme [5]. The C-terminal glycine residue of ubiquitin gets activated by linking to the active cysteine residue of E1 enzyme through a thioester bond. This reaction needs ATP and Mg as cofactors. E1 enzyme transfers Ub molecule to E2 again by thioester linkage. Finally, E3 enzyme directly or indirectly transfers the ubiquitin to the target protein from the activated E2 enzyme [6]. The ubiquitin pathway is found only in eukaryotic cells [7], but microbial pathogens have been observed to secrete proteins into host cells that interferes with this pathway [8]. Pathogenic bacteria *Salmonella, Shigella, Chlamydia, Pseudomonas* and *Legionella* secrete effector proteins that either act as a bacterial E3 ligase (BEL) or act on the ubiquitinated proteins and act as a deubiquitinase (DUB) [9]. SidE protein family secreted by *Legionella* have been reported to ubiquitinate target proteins via a novel mechanism without the need of conventional ubiquitination enzymes [10]. The monoADP-ribosylation (mART) domain of SidE proteins transfers an ADP-ribosyl moiety (ADPR) to the R42 residue of ubiquitin using the NAD molecule as a cofactor and releasing nicotinamide; and the phosphodiesterase (PDE) domain cleaves the AMP from the ADPR moiety and ligates the phosphoribosylated (PR) ubiquitin

to the serine residue of target proteins [11, 12]. SidE members also contain a conventional DUB domain that cleaves the K42 and/or K63 linked ubiquitin from a very broad range of proteins [13]. The biochemical mechanism of both mART and PDE domains remains elusive. In the previous study that was detailed in Chapter 2, our lab has successfully solved the mechanism of PDE and how the target is PR-ubiquitinated. We have identified a new domain between mART and PDE domains and named it as Mid domain that helps with the activity of mART domain.

The sole function of secreted effectors is not to act on host pathways. Metaeffectors are proteins that regulate the activity of other effector proteins during infection [14]. The metaeffector activity can be on the other effectors directly as in the case of LubX ubiquitinating other effector protein SidH [15], or the activity can be on the substrates or the products of the other effector proteins. A well-known example is the modulation of Rab1 during infection: AnkX effector protein adds a phosphocholine (PC) moiety to Rab1 [16] and Lem3, and at a later stage of the infection, removes that moiety [17]. SidM adds an AMP molecule to Rab1 [18] and SidD protein arrives and removes AMP again at a later stage [19]. Even earlier than the report on the activity of the SidE proteins, SidJ effector has shown to be a metaeffector for SidE family members in a yeast survival assay [20]. A later report on SidJ showed that the protein can cleave the PR-ubiquitin modification of SdeA from the target proteins and act as a deubiquitinase for SidE ubiquitination [21].

SidE protein family contains five other proteins (lpg1496, lpg2509, lpg2154, lpg2239, and lpg2523) that share the PDE domain with the other four members (SidE, sdeA-C) (see Figure 1-3). With the activity of SidJ cleaving the ubiquitin from the targets revealed, we wondered what the proteins that have the PDE domain alone can do during infection. In this study, we have shown that small SidE members, lpg2509 (sdeD) and lpg2154 (sdeF) can act on ADPR-ubiquitin to make

PR-Ub and these two proteins can cleave the PR-Ub molecule from SidE targeted proteins both *in vitro* and *in vivo*. We observed that overexpression of SdeA protein causes Golgi fragmentation and when SdeD or SdeF is overexpressed with SdeA together, the phenotype is gone. During mitosis, Golgi fragmentation is regulated by the ubiquitination state of SNARE proteins [22]. It is well established that the SNARE protein Syntaxin5 (Stx5) gets ubiquitinated by HACE1 during mitosis and deubiquitinated by the VCIP135 complex to form the fully functional Golgi [23]. We showed that SdeA PR-ubiquitinates Stx5 and SdeD, SdeF proteins deubiquitinate Stx5, both *in vivo* and *in vitro*. This "metaeffector" activity of smaller SidE members provides information about how the PR-ubiquitination of SidE members can be regulated. This might help to revert the initial infection steps during LCV formation and let the host cell recover until *Legionella* continues its development hiding inside the LCV.

3.3 RESULTS

SdeD and SdeF can act on ubiquitin, but not on target proteins

SidE protein family shares the PDE domain in all its members. When the sequences are aligned, the active site residues are observed to be conserved in all the proteins (Figure 3-1). With this conservation, we were curious if the small PDE proteins will also share the same activity. To test, we synthesized ADPR-Ub by incubating ubiquitin with PDE-dead mutant of SdeA core protein (210-911 aa.). When incubated with ADPR-Ub, PDE domains of SdeA, SdeD and SdeF can all remove the AMP moiety from the ADPR-Ub as shown with Pro-Q Diamond Phosphoprotein Stain (Figure 3-2 A).

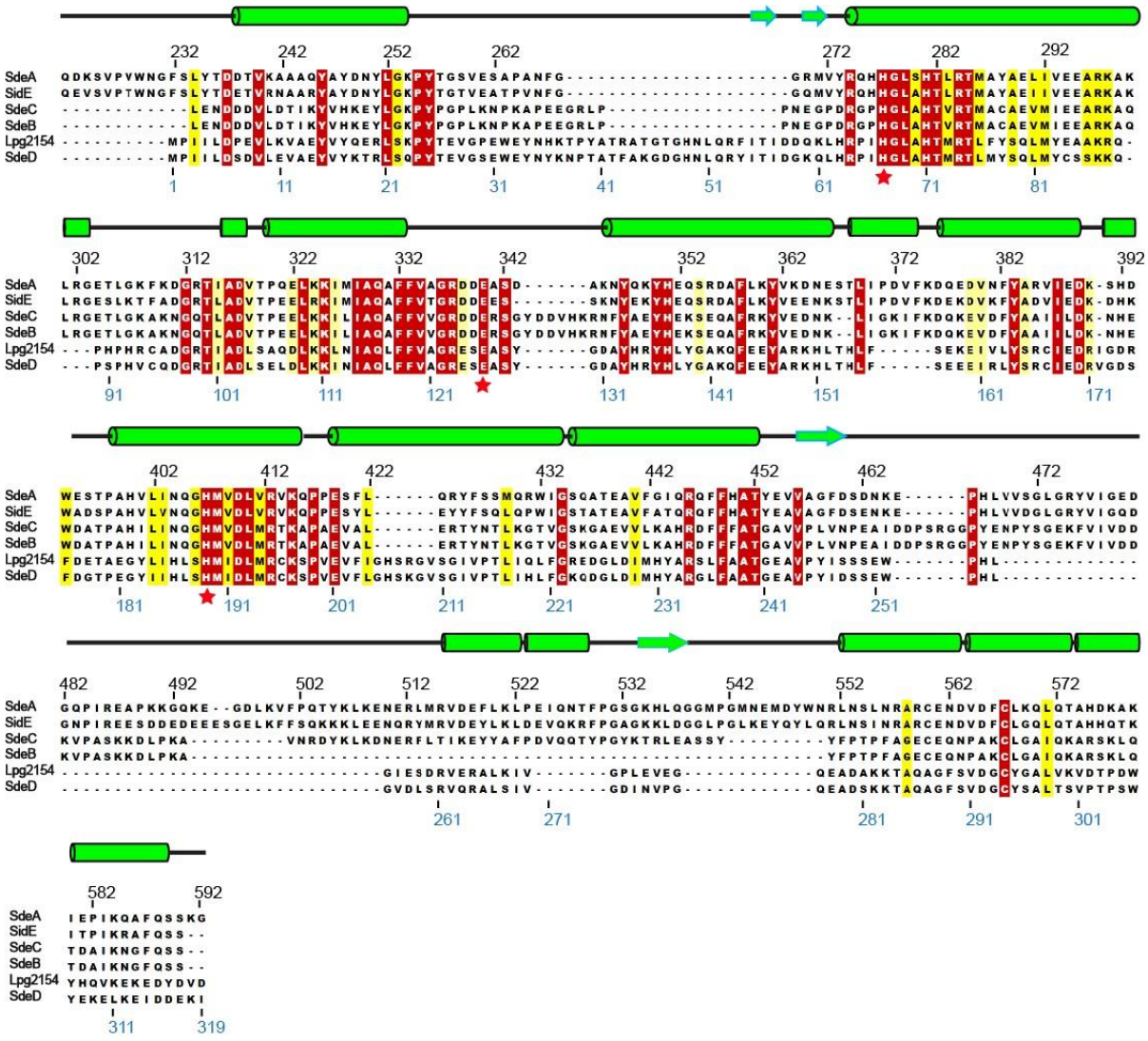


Figure 3-1: Multiple Sequence Alignment of the PDE Domain. Representative sequences corresponding to the PDE domain of SdeA (aa. 222-502) were aligned by MultiAlin [24]. Entrez database accession numbers are as follows: SdeA, GI: 1064303039; SidE, GI: 52840489; SdeB, GI: 52842367; SdeC, GI: 52842370; lpg2154, GI: 52842368; and SdeD, GI: 52842717. 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. The three essential catalytic residues are highlighted with stars.

Next, we wanted to see if SdeD and SdeF can ubiquitinate target proteins. Rab33b (shown to be target of SdeA [25]) and whole cell lysate (WCL) from HEK293T cells were incubated with the designated PDE domain proteins and ADPR-Ub proteins. To our surprise, we observed that only the PDE domain of SdeA can PR-ubiquitinate the target proteins in both Rab33b and WCL proteins (Figure 3-2 B and C).

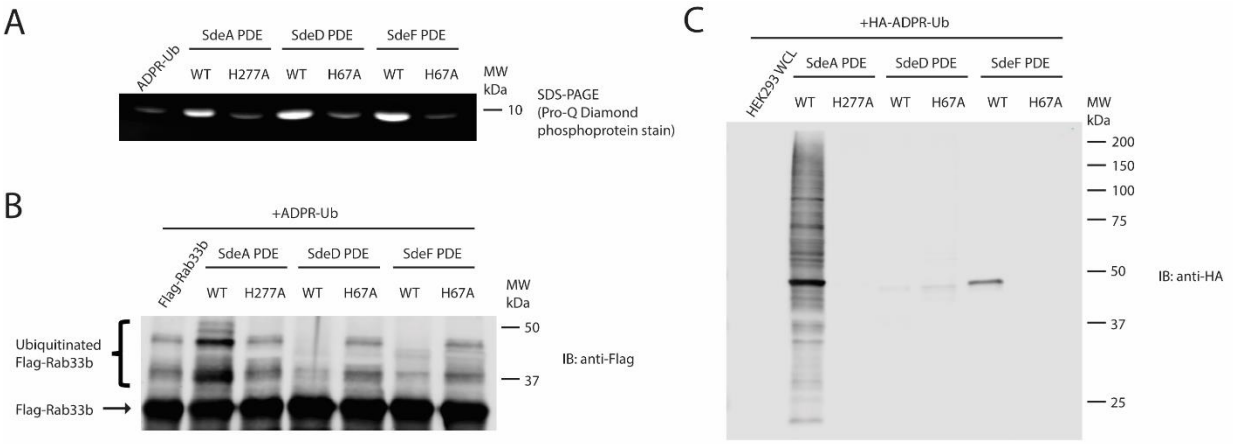
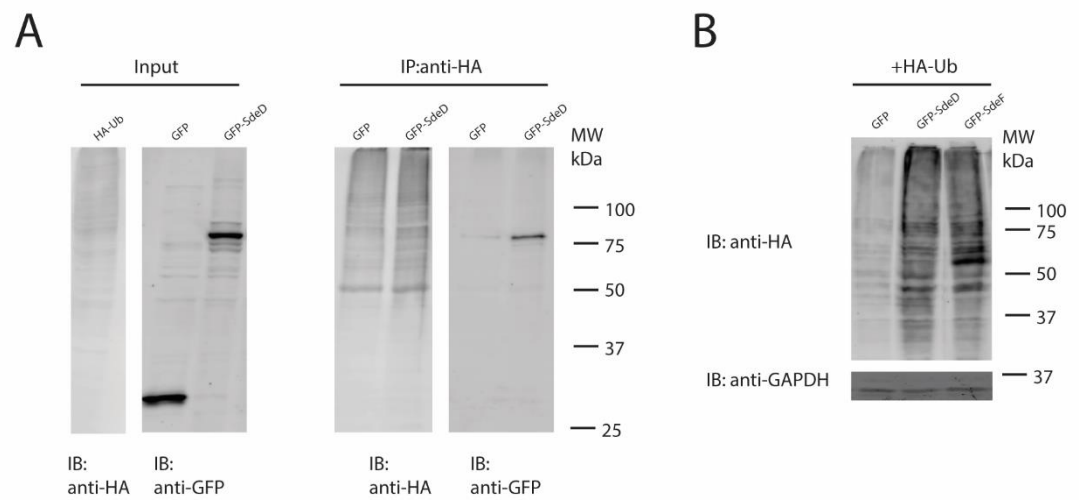


Figure 3-2: SdeD-SdeF Activity on Ubiquitin and SdeA Targets. (A) *In vitro* Phosphodiesterase activity assay. Cleavage of AMP by PDE domains can be monitored by Pro-Q Diamond Phosphoprotein staining of Ub on SDS gel. (B) Rab33b modification by PDE domains. Purified Flag-Rab33b was mixed with ADPR-Ub and indicated PDE domains of SidE family members. Reaction products were observed by immunoblotting against Flag antibody. (C) Modification of proteins in whole cell lysate. The same experimental set-up was used as in (B), but instead whole cell lysate from HEK293T cells were mixed with HA-ADPR-Ub and indicated PDE domains. The reaction products were run on SDS-PAGE and analyzed by immunoblotting against HA antibody (the PDE dead mutant proteins were used to compare with following results).

SdeD and SdeF interact with ubiquitin and ubiquitinated proteins

The activity difference of SdeA-SdeD PDE domains can be traced back to the structural differences. When we compare the two domains, we see two different "lids" above the catalytic grooves: the extensions from the catalytic grooves stay on different sides of the catalytic residues (see Figure 2-8 for details). This might be the structural constraint of SdeD to reach to the serine

residue of the target protein. The differing localization of the lid might also be the reason why SdeD can bind to very strongly to ubiquitin but SdeA cannot. Even though we could not crystallize the PDE domain of SdeA with ubiquitin, the complex could be obtained with SdeD PDE domain. To show Ub-sdeD physical interaction, we wanted to do co-immunoprecipitation. SdeD protein gets ubiquitinated when transfected in HEK293T cells (data not shown), so we did tandem immunoprecipitation. We incubated bead-bound HA-Ubiquitin transfected cell lysate with GFP-SdeD transfected cell lysate. Only SdeD could interact with ubiquitin and ubiquitinated proteins



(Figure 3-3 A).

Figure 3-3: SdeD-SdeF Can Interact with Ub and Ubiquitinated Proteins. (A) Interaction of SdeD with ubiquitinated proteins. Whole cell lysate from HEK293T cells with HA-Ub transfection were lysed and pulled down with HA-beads. The bound proteins were further mixed with either GFP or GFP-SdeD containing whole cell lysates. Beads were re-washed after second incubation and samples were run on SDS-PAGE. Analysis was done by immunoblotting against anti-HA and anti-GFP antibodies. (B) Ubiquitination pattern change by SdeD-SdeF o/e. HA-Ub and designated GFP tagged protein containing HEK293T cells were lysed and whole cell lysates were run on SDS-PAGE. The analysis was done by immunoblotting against anti-HA and anti-GFP.

SdeD and SdeF act as deubiquitinase on SdeA target proteins

When co-transfected with ubiquitin, SdeD and SdeF cause elevated levels of ubiquitinated proteins compared to PDE domain of SdeA alone (Figure 3-2 B). This was interesting since overexpression of SdeA PDE domain did not change the ubiquitination levels (data not shown). A recent finding about the metaeffector protein SidJ shows the deubiquitinase activity against PR-ubiquitinated proteins [21]. We wondered whether the PDE domains of SdeD and SdeF can perform deubiquitination activity. To test, we PR-ubiquitinated Rab33b protein with ADPR-Ub and SdeA PDE domain and later added SdeD or SdeF wild type or PDE-dead mutants. Both SdeD and SdeF PDE domains could cleave the PR-Ub from Rab33b (Figure 3-4 A). We repeated the same experiment with HEK293T cells, WCL was incubated with HA-ADPR-Ub and SdeA PDE domain and further incubated with SdeD and SdeF proteins. We showed that the PR-ubiquitinated proteins were deubiquitinated by both SdeD-SdeF proteins (Figure 3-4 B). In both experiments, the PDE-dead mutants of SdeD and SdeF could not remove the PR-Ub, meaning that the DUB activity is related to their PDE function (Figure 3-4 A and B).

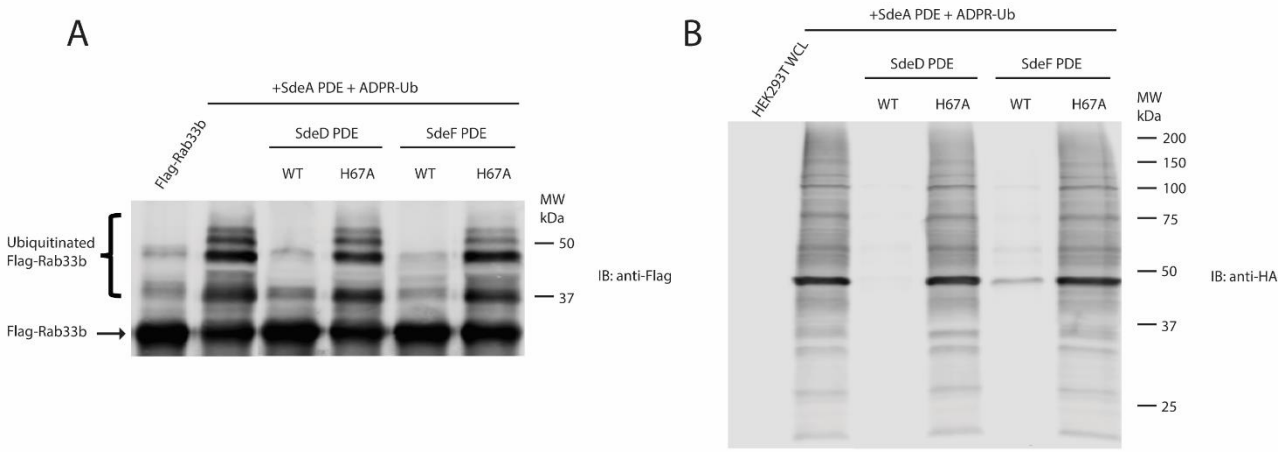


Figure 3-4: SdeD-SdeF act as Deubiquitinase for SdeA Target Proteins. (A) *In vitro* DUB activity of PDE domains. Pre-incubated Rab33b samples with ADPR-Ub and SdeA PDE domain were further incubated with indicated SdeD-SdeF proteins. Samples were run on SDS-PAGE and

results were analyzed by immunoblotting against anti-Flag antibody. (B) Deubiquitination of PR-ubiquitinated target proteins from whole cell lysate. Same experimental set up was used as in figure A. Whole cell lysate of HEK293T cells and HA-ADPR-Ub was used. Reaction products were run on SDS-PAGE and analysis was carried out by immunoblotting against anti-HA antibody.

SdeA causes Golgi fragmentation & SdeD and SdeF can revert the phenotype

The overexpressed SdeA protein is toxic to yeast and mammalian cells [20]. We wanted to see where SdeA is localized in HeLa cells to exert this toxic phenotype. To observe the localization, we co-transfected SdeA with different cellular markers and observed that the protein is localized mainly to the cytoplasm. Observing Golgi by immunostaining the Golgi marker GPP130 (Golgi Phosphoprotein 4 [26]), showed a more fragmented structure (Figure 3-5 A). This Golgi phenotype is dependent on the mART and PDE activities of SdeA: when we overexpressed the mART or PDE dead mutants of SdeA, there was no Golgi fragmentation observed (Figure 3-5 A). We wondered if SdeD and SdeF could revert the phenotype observed by SdeA overexpression. When we co-expressed SdeA with either SdeD or SdeF WT proteins, the fragmentation phenotype was gone (Figure 3-5 B). This reversion phenotype is PDE-dependent because when we co-expressed PDE dead mutants of SdeD and SdeF, the fragmentation phenotype by SdeA persisted (Figure 3-5 B).

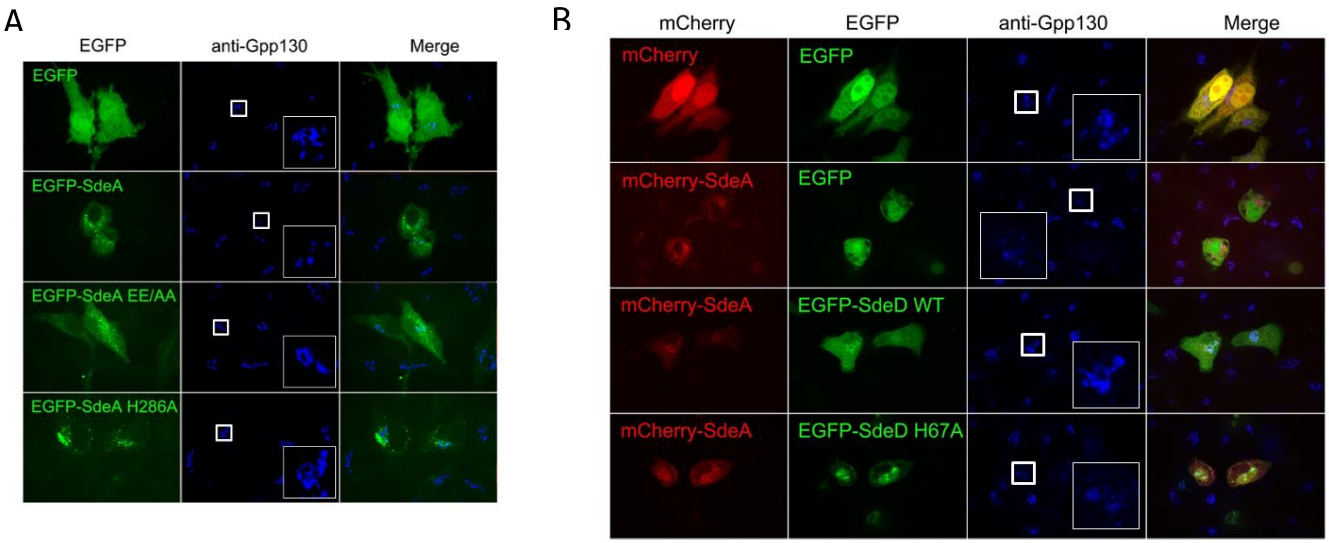


Figure 3-5: SdeA Causes Golgi Fragmentation & SdeD-SdeF can Revert the Phenotype. (A) Golgi fragmentation by SdeA. HeLa cells were transfected with the corresponding GFP-tagged proteins. Fixed cells were immunostained with anti-Gpp130 antibody. Images were obtained by confocal microscopy. (B) Phenotype reversion by SdeD-SdeF. HeLa cells were transfected with the corresponding GFP and mCherry tagged proteins. Fixed cells were immunostained with anti-Gpp130 antibody. Images were obtained by confocal microscopy. Zoomed-in boxes show the Golgi phenotype under designated transfection conditions.

Syntaxin 5 is a target of SdeA and SdeD-SdeF

We wanted to find a correlation between SdeA activity and Golgi fragmentation. A recent report suggested such a regulation of a SNARE protein, Syntaxin 5 (Stx5) during mitosis (Figure 3-6 a). This study showed that Stx5 gets monoubiquitinated during early mitosis stages by the HACE1 E3 ubiquitin ligase whose activity was previously shown to be important for Golgi disassembly [27]. This modification inhibits interaction of Stx5 with its complementing SNARE protein, Bet1. This interaction is also important for ER to Golgi transport [28]. At later stages of mitotic cycle, the VCIP135 deubiquitinase protein, in complex with p97 and p47, removes the monoubiquitin from Stx5. This allows Stx5 to interact with Bet1 and Golgi assembly can start in daughter cells [23]. With this information, we wanted to test if Stx5 can be modified by SdeA.

Flag-Stx5 was co-transfected with WT and mutant SdeA constructs. Stx5 was not in a high amount in whole cell lysates (data not shown), so we immunoprecipitated Stx5 from co-transfected cells and checked for modifications by SdeA. As seen from figure 3-6 b, Stx5 is ubiquitinated by SdeA WT, but not with mART or PDE domain mutants. For the *in vitro* modification, we purified MBP tagged Stx5 with its C-terminal transmembrane region truncated (1-279) and incubated it with SdeA with or without NAD. As can be seen from Figure 3-7 B (first two lanes), SdeA can modify Stx5 the same way as *in vivo* (Figure 3-7 B).

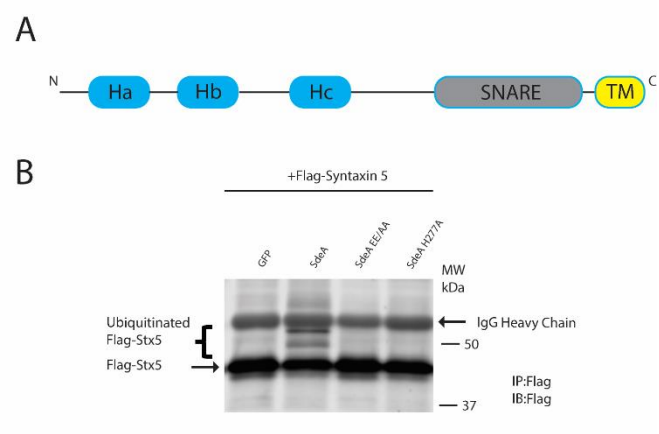


Figure 3-6: Syntaxin 5 is Modified by SdeA. (A) Schematic diagram of Syntaxin 5 domain structure. N' part of Stx5 contains three alpha helical bundles (Ha-c) and the C' part contains the SNARE domain and the Transmembrane (Tm) domain. (B) Modification of Stx5 by SdeA. Flag-Stx5 and designated GFP constructs co-transfected HEK293T cells were lysed and Flag-Stx5 was immunoprecipitated by anti-Flag beads. The eluates were run on SDS-PAGE and analyzed by immunoblotting against anti-Flag antibody.

SdeA modifies the serine residues of its targets [11]. In this context, we wanted to find the residue that SdeA can modify on Stx5 protein. We mapped the serine residues that are localized on the SNARE domain of Stx5 (Figure 3-6 A) [29]. Individual and collective mutation of residues S278, S285 and S320 did not alter the modification of SdeA, meaning that the modified serine lies on another part of Stx5 (data not shown).

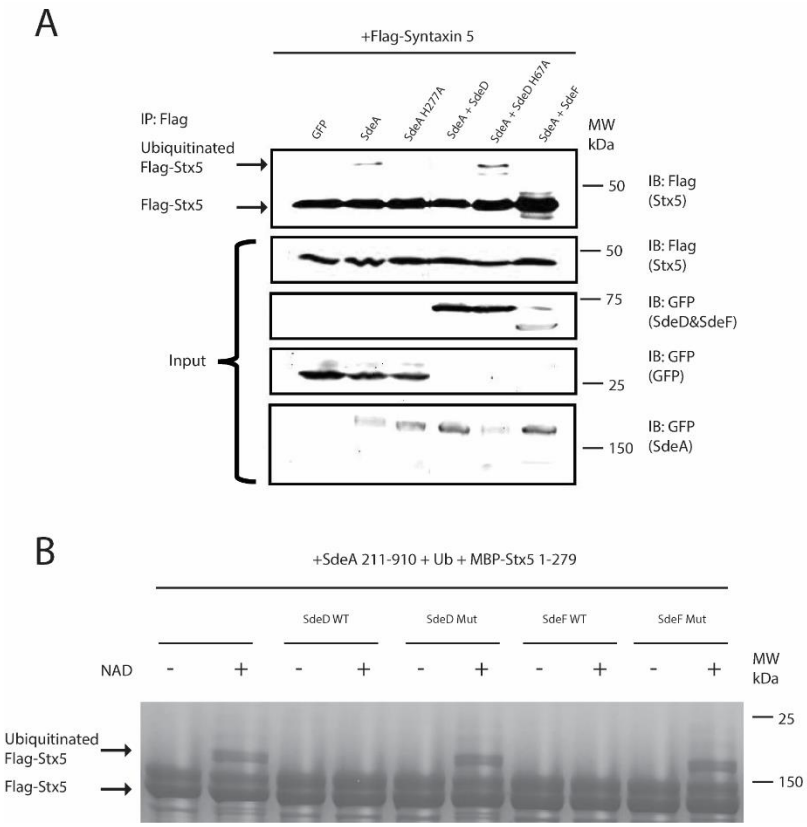


Figure 3-7: SdeD-SdeF can Remove PR-Ub from Stx5. (A) DUB activity of SdeD-SdeF *in vivo*. Flag-tagged Stx5 from cell lysates of indicated co-transfected HEK293T cells were immunoprecipitated. The eluates were run on SDS-PAGE and the analysis was carried out by immunoblotting against anti-GFP and anti-Flag antibodies. (B) *In vitro* PR-Ub removal by SdeD-SdeF. Maltose-binding-protein (MBP) tagged Stx5 was mixed with SdeA and Ub with or without NAD in all the samples. Additional SdeD-SdeF constructs were added as indicated within the lanes. The reaction mixtures were run on SDS-PAGE and analyzed by Coomassie staining. All proteins were induced in and purified from E. coli.

Next, we wanted to see if SdeD-SdeF can revert the SdeA modification on Stx5. We co-transfected HEK293T cells with SdeA, SdeD (or SdeF) and Stx5. After the IP of Stx5, we observed that PR-Ub modification of Stx5 is reversed by SdeD-SdeF in a function-dependent manner (Figure 3-7 A). We further wanted to see if this PR-Ub removal by SdeD-SdeF can be replicated *in vitro*. Again, MBP tagged Stx5 was incubated with SdeA with or without NAD for one hour, and further incubated with SdeD-SdeF for another hour. Similarly, we observed that PR-Ub modification by SdeA on Stx5 was removed by the addition of SdeD-SdeF depending on their function (Figure 3-7 B). It was previously shown that SidJ can remove the PR-Ub form SdeA targets [21], so we wanted to see if this reversion by SdeD-SdeF is observed by SidJ as well. To

our surprise, the modification of Stx5 increased compared to other co-transfection conditions, instead of diminishing. This might be related to the high specificity of SidJ targeting (data not shown) [21].

3.4 DISCUSSION

Previous studies on SdeA and other SidE family proteins show the biochemical activity of these proteins [10, 11]. There was no mention of the small SidE members, SdeD and SdeF, in these studies. The work described in this chapter argues that the proteins SdeD and SdeF acts as deubiquitinase enzymes against proteins that have a PR-Ub ligated by a SidE family member.

When exposed to ADPR-Ub, the PDE proteins SdeD-SdeF could easily remove the AMP moiety from ubiquitin, but when either whole cell lysate or previously shown target protein Rab33b was incubated with these two proteins, no modification occurs. This led us to think that there should be another function to these PDE domains. The comparison between the PDE domains of SdeA and SdeD also proves some differences between these proteins: the loop between residues 26 and 48 of SdeD is missing within SdeA PDE domain and the loop between residues of 465 and 513 of SdeA is completely gone within the SdeD PDE domain. Both loops that create a difference cover the space above the active cleft. This difference of loop location might be the reason behind why SdeA can ubiquitinate proteins but SdeD can not. Testing the activity of SdeD and SdeA after replacing these loops with the counterparts from the other PDE domain might give us more clues as to how the PDE domain of SdeA can interact with its target proteins and how SdeD can bind to ubiquitin much tighter than SdeA.

A recent paper published about effector protein SidJ showed that the protein acts as a deubiquitinase against the proteins that are modified by SidE family [21]. The cleavage is shown

to be on the β-phosphate of the ADPR moiety which is the same bond that PDE domain targets. This led us to think that the PDE domains that are conserved in the small members SdeD-SdeF might have a similar function. We showed that SdeD-SdeF can act as a deubiquitinase against PR-Ub conjugated proteins both *in vivo* and *in vitro*. This PR-Ub removal might be important during infection, since the first step of infection, LCV formation, requires different set of effectors compared to the second step of infection, LCV maturation and replication [30]. This antagonistic activity has already been reported for different *Legionella* effector proteins. The Rab1 host protein is first AMPylated by SidM molecule at early stages of LCV formation [19] and later gets deAMPylated by another effector protein SidD [17]. The activity of the targeted proteins by SdeA may be suspended by this modification during the formation of LCV, and at later stages of infection, SdeD and SdeF may reverse the suspension state of these proteins and let the host cell metabolism go back to normal. It is essential to pinpoint the exact secretion of these proteins during infection to shed more light on the activity regulation of these PDE domain-containing proteins.

ER-to-Golgi destined anterograde vesicles form the outer layer of the LCV [31]. The morphology of Golgi has been observed to be intact throughout the infection even though vesicles destined for the Golgi are recruited by *Legionella*. We have observed a Golgi fragmentation phenotype when SdeA protein is overexpressed. This introduction of a possible "artificial" phenotype with overexpression is observed by many other effector proteins. When the *Legionella* effector protein LpdA, a phospholipase D molecule, is overexpressed in mammalian cells, Golgi shows a fragmented phenotype [32]. When we overexpressed SdeD and SdeF proteins in HeLa cells, we observed ubiquitin puncta structures that were missing in control cells (data not shown). During infection, *Legionella* secretes a small amount of each effector to the host cell. A study on VipD phospholipase effector protein showed a total of 500 molecules during infection [33]. This

might mean that when some of the effectors are overexpressed in mammalian cells, effects of these molecules might get exaggerated and some artifact phenotypes might be observed. Although, these artifacts may hint on the actual roles of these proteins during infection conditions. Monitoring the Golgi phenotype in a time-dependent manner with SdeA overexpression may help to understand the exact effect of SdeA on Golgi.

Syntaxin 5 (Stx5) was predicted to be a candidate for SdeA Golgi phenotype regarding its ubiquitination state during mitosis. The ubiquitination of SNARE domain of Stx5 was shown to be crucial for interaction with its counterpart Bat1 SNARE protein [23]. For this reason, we mutated the serine residues that might be targeted by SdeA within the SNARE domain. The mono or triple mutation of these serine residues did not affect the ubiquitination by SdeA molecule. This means that another serine residue (or residues) gets modified by SdeA that might affect the activity of Stx5 in a different manner. The fragmentation of Golgi might not have anything with the modification of Stx5 protein. The next step should be trying to find more targets of SdeA by *in vitro* or *in vivo* methods that might relate the activity of SdeA to the Golgi phenotype.

3.5 MATERIALS and METHODS

Cloning and Mutagenesis

DNA fragments encoding SdeA (aa. 211-599), SdeD (a.a. 1-341) and SdeF (1-330) 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, SdeD and SdeF were introduced by site-directed mutagenesis using oligonucleotide primer pairs containing the appropriate base changes. DNA fragments encoding wild type or mutant Ub were PCR amplified and subcloned into a pET21a vector. All constructs were confirmed by DNA sequencing. The plasmids used for Rab33b and Syntaxin 5 experiments were provided by Zhao-Qing's lab in Purdue and Yanzhuang Wang's lab in University of Michigan.

Protein Expression and Purification

For expression of SidE family proteins, E. coli BL21-DE3 strains harboring the expression plasmids were grown in Luria-Bertani medium supplemented with 50 μg/ml kanamycin to mid-log phase. Protein expression was induced for overnight at 18°C with 0.1 mM isopropyl-B-D-thiogalactopyranoside (IPTG). Harvested cells were resuspended in a buffer containing 20 mM Tris-HCl (pH 8.0), 150 mM NaCl) and were lysed by sonication. Soluble fractions were collected after centrifugation at 16,000 rpm for 30 min at 4°C and incubated with cobalt resins (Clonetech) for 1.5 h at 4°C. Protein bound resins were extensively washed with lysis buffer. The SUMO-specific protease Ulp1 was then added to the resin slurry to release proteins from the His-SUMO tag. Eluted protein samples were checked by running SDS-PAGE. Protocols for Ub expression and purification were adapted from publish literatures [34]. Briefly, harvested cells were

resuspended in a buffer containing 20 mM ammonium acetate (pH 5.1). Cells were lysed by sonication and cell lysate was clarified by 18.000 rpm centrifugation for 30 min. Clarified supernatant was pooled and was titrated gradually with acetic acid to lower the pH to 4.8. Precipitated proteins after the titration were removed by centrifuged at 18.000 rpm for 30 min. The pH of the clarified supernatant was adjusted to 5.1 by NaOH addition. HiTrap SP column (GE Healthcare) was used for cation exchange with a buffer gradient from 20 mM ammonium acetate to 0.5 M ammonium acetate (both pH: 5.1). Fractions containing ubiquitin peak were pooled and was 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 prepare ADPR-Ub for biochemical assays, 1 μM of SdeA (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.

For purification of 6xHis-Flag-Rab33b, the same protocol as for SidE proteins was followed. Instead of cleaving the protein tag with a protease, the protein was eluted with Elution buffer (20mM Tris, pH:8.0, 150mM NaCl and 400mM Imidazole). The eluted protein was concentrated and further purified by FPLC size exclusion chromatography. The peaks corresponding to the protein was pooled down and concentrated before the reaction.

Purification of MBP-Syntaxin 5 followed the same protocol. Instead of using Cobalt beads, the protein was purified with Amylose Resin (New England Biolabs) and the protein was eluted with eluted with Elution buffer (20mM Tris, pH:8.0, 150mM NaCl and 100mM Maltose).

Computational Analysis and Graphic Presentation of Protein Sequence and Structure

SidE family protein sequences were selected from results generated by BLAST server (NCBI). Edited sequences were aligned with Clustal Omega [35] and colored by Multiple Align Show online server (http://www.bioinformatics.org/sms/index.html). All structural figures were generated using PyMOL (The PyMOL Molecular Graphics System, Version 1.8.X, Schrödinger, LLC).

Ubiquitin Modification and Rab33b & Whole Cell Lysate Ubiquitination Assays

ADPR-Ubiquitin modification reactions were carried out by mixing 1 μM of PDE protein with 25 μM ubiquitin in a reaction buffer containing 50 mM NaCl and 50 mM Tris pH 7.5. The reactions were incubated for 1h at 37°C and reaction products were assessed by SDS-PAGE. Gels were stained with Pro-Q Diamond phosphoprotein stain (Invitrogen) to assay for PDE activity. Rab33b ubiquitination reactions were performed by the addition of 4 μM of recombinant Flag-Rab33b to the ubiquitin 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. Whole cell lysate was collected from a 10cm 100% confluent HEK293T cells. The cells were lysed by Lysis buffer (50mM Tris pH:7.5, 150mM NaCl, 1% Triton-X100). The concentration of the lysate was measured by NanoDrop (Thermo Scientific). Whole cell lysate modification by PDE domains were carried out by adding 25 μM HA-ADPR-Ub to 100 μg of cell lysate with 1 μM PDE proteins. The reaction mixture was incubated at 37^{o}C for 1h and the samples were run on 12% SDS-PAGE. The analysis was done by Western blot and using anti-HA antibody (Sigma-Aldrich) at 1:10000 for immunoblotting.

SdeD-SdeF Deubiquitinase Assays

4 μM Rab33b molecule was mixed with 1 μM SdeA PDE domain with 25 μM ADPR-Ub and incubated at 37^0C for 1h. 1 μM SdeD-SdeF constructs were added to the reacted tubes and re-incubated at 37^0C for 1h. The reaction products were run on 12% SDS-PAGE and analysis was done by Western blotting by using anti-Flag antibody for immunoblotting. Whole cell lysate reactions were carried out by mixing 100 μg whole cell lysate with 1 μM SdeA PDE domain and 25 μM HA-tagged ADPR-Ub. The reactions were incubated at 37^0C for 1h. 1 μM SdeD-SdeF constructs were added to the reacted tubes and re-incubated at 37^0C for 1h. The reaction products were run on 12% SDS-PAGE and analysis was done by Western blotting by using anti-HA antibody (Sigma Aldrich) for immunoblotting.

Immunofluorescence of Golgi Fragmentation

Cells were seeded on glass coverslips and fixed in 4% PFA/PBS for 15 min. Cells were washed twice with PBS and permeabilized in PBS containing 0.05% saponin and 3% BSA for 30 min. Cells were immunostained with proper primary and secondary antibodies. Anti–rabbit GPP130 (Covance) antibodies were used at 1:500. Confocal images, acquired at room temperature, were taken using a CSU-X spinning disk microscope (Intelligent Imaging Innovations) with a spherical aberration correction device, 100×, 1.4 NA objective, on an inverted microscope (Leica), and acquired with a QuantEM EMCCD camera using Slidebook software (Intelligent Imaging Innovations).

In Vivo Syntaxin-5 Ubiquitination Assay

HEK293T cells were grown on 6-well plates to 80% confluency and the transfections were carried out with PEI (Thermo Fisher). Cells were lysed with Lysis buffer (50mM Tris pH: 8.0, 150 mM NaCl, 0.1% Deoxycholate, 1% Triton X-100, 1X PMSF, 1X EDTA-Free Protease Tablet from Roche) and cell lysate was collected after 20 min centrifugation. M2-Flag (Sigma Aldrich) beads were used to immunoprecipitate Syntaxin-5 protein. The samples were washed three times with Wash buffer (50mM Tris pH:8.0, 150mM NaCl, 1% Triton-X100) and the samples were eluted from the beads by boiling at 95^{o}C. The input and eluate were run on 8% SDS-PAGE and analysis was carried out by western blotting by using anti-Flag antibody (Sigma Aldrich) for immunoblotting.

3.6 REFERENCES

1. Hershko, A. and A. Ciechanover, *The ubiquitin system.* Annu Rev Biochem, 1998. **67**: p. 425-79.

2. Kattah, M.G., B.A. Malynn, and A. Ma, *Ubiquitin-Modifying Enzymes and Regulation of the Inflammasome.* J Mol Biol, 2017.

3. Dubrez, L., *Regulation of E2F1 Transcription Factor by Ubiquitin Conjugation.* Int J Mol Sci, 2017. **18**(10).

4. de Poot, S.A.H., G. Tian, and D. Finley, *Meddling with Fate: The Proteasomal Deubiquitinating Enzymes.* J Mol Biol, 2017.

5. Dye, B.T. and B.A. Schulman, *Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins.* Annu Rev Biophys Biomol Struct, 2007. **36**: p. 131-50.

6. Scheffner, M., U. Nuber, and J.M. Huibregtse, *Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade.* Nature, 1995. **373**(6509): p. 81-3.
7. Maupin-Furlow, J.A., *Prokaryotic ubiquitin-like protein modification.* Annu Rev Microbiol, 2014. **68**: p. 155-75.

8. Wimmer, P. and S. Schreiner, *Viral Mimicry to Usurp Ubiquitin and SUMO Host Pathways.* Viruses, 2015. **7**(9): p. 4854-72.

9. 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.

10. Nakasone, M.A. and D.T. Huang, *Ubiquitination Accomplished: E1 and E2 Enzymes Were Not Necessary.* Mol Cell, 2016. **62**(6): p. 807-809.

11. Bhogaraju, S., et al., *Phosphoribosylation of Ubiquitin Promotes Serine Ubiquitination and Impairs Conventional Ubiquitination.* Cell, 2016. **167**(6): p. 1636-1649 e13.

12. 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.

13. 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.

14. Kubori, T., et al., *Legionella metaeffector exploits host proteasome to temporally regulate cognate effector.* PLoS Pathog, 2010. **6**(12): p. e1001216.

15. Kubori, T., A.M. Hubber, and H. Nagai, *Hijacking the host proteasome for the temporal degradation of bacterial effectors.* Methods Mol Biol, 2014. **1197**: p. 141-52.

16. Mukherjee, S., et al., *Modulation of Rab GTPase function by a protein phosphocholine transferase.* Nature, 2011. **477**(7362): p. 103-6.

17. Neunuebel, M.R., et al., *Legionella pneumophila LidA affects nucleotide binding and activity of the host GTPase Rab1*. J Bacteriol, 2012. **194**(6): p. 1389-400.

18. Brombacher, E., et al., *Rab1 guanine nucleotide exchange factor SidM is a major phosphatidylinositol 4-phosphate-binding effector protein of Legionella pneumophila.* J Biol Chem, 2009. **284**(8): p. 4846-56.

19. Neunuebel, M.R., et al., *De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila.* Science, 2011. **333**(6041): p. 453-6.

20. 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.

21. Qiu, J., et al., *A unique deubiquitinase that deconjugates phosphoribosyl-linked protein ubiquitination.* Cell Res, 2017. **27**(7): p. 865-881.

22. Wang, Y., et al., *VCIP135 acts as a deubiquitinating enzyme during p97-p47-mediated reassembly of mitotic Golgi fragments.* J Cell Biol, 2004. **164**(7): p. 973-8.

23. Huang, S., D. Tang, and Y. Wang, *Monoubiquitination of Syntaxin 5 Regulates Golgi Membrane Dynamics during the Cell Cycle.* Dev Cell, 2016. **38**(1): p. 73-85.

24. Corpet, F., *Multiple sequence alignment with hierarchical clustering.* Nucleic Acids Res, 1988. **16**(22): p. 10881-90.

25. Qiu, J., et al., *Ubiquitination independent of E1 and E2 enzymes by bacterial effectors.* Nature, 2016. **533**(7601): p. 120-4.

26. Masuda, M., et al., *Golgi phosphoprotein 4 (GPP130) is a sensitive and selective cellular target of manganese exposure.* Synapse, 2013. **67**(5): p. 205-15.

27. Tang, D., et al., *The ubiquitin ligase HACE1 regulates Golgi membrane dynamics during the cell cycle.* Nat Commun, 2011. **2**: p. 501.

28. Zhang, T. and W. Hong, *Ykt6 forms a SNARE complex with syntaxin 5, GS28, and Bet1 and participates in a late stage in endoplasmic reticulum-Golgi transport.* J Biol Chem, 2001. **276**(29): p. 27480-7.

29. Shorter, J., et al., *Sequential tethering of Golgins and catalysis of SNAREpin assembly by the vesicle-tethering protein p115.* J Cell Biol, 2002. **157**(1): p. 45-62.

30. Bruggemann, H., C. Cazalet, and C. Buchrieser, *Adaptation of Legionella pneumophila to the host environment: role of protein secretion, effectors and eukaryotic-like proteins.* Curr Opin Microbiol, 2006. **9**(1): p. 86-94.

31. Heidtman, M., et al., *Large-scale identification of Legionella pneumophila Dot/Icm substrates that modulate host cell vesicle trafficking pathways.* Cell Microbiol, 2009. **11**(2): p. 230-48.

32. Schroeder, G.N., et al., *Legionella pneumophila Effector LpdA Is a Palmitoylated Phospholipase D Virulence Factor.* Infect Immun, 2015. **83**(10): p. 3989-4002.

33. Gaspar, A.H. and M.P. Machner, *VipD is a Rab5-activated phospholipase A1 that protects Legionella pneumophila from endosomal fusion.* Proc Natl Acad Sci U S A, 2014. **111**(12): p. 4560-5.

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

35. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Mol Syst Biol, 2011. **7**: p. 539.

CHAPTER 4

CONCLUDING REMARKS AND FUTURE DIRECTIONS

Legionella pneumophila manipulates a wide array of host cellular processes during its lifecycle inside the host cell [1]. More than 300 effector proteins are secreted during this invasion phase and all these effector proteins are involved in forming the *Legionella*-containing vacuole (LCV) to let the bacteria replicate and develop within the host cytoplasm [2]. These effector proteins are tightly regulated and have a consecutive order of secretion/activity during infection [3-5]. They also possess a high redundancy as shown by various genomic deletion and infection studies [6-8]. There is a wide array of pathways modulated by these effector proteins: Small GTPases get post-translationally modified [9-12], autophagy is inhibited or enhanced according to the stage of infection [8, 13, 14], phospholipids, mostly phosphoinositides, get modified [15-17], and such. Ubiquitination machinery is one of the pathways that is "sabotaged" by *Legionella* [18]. Conventional ubiquitination needs a three enzyme cascade to modify target protein; E1, E2 and E3 [19], and bacterial effector proteins have been observed to mimic the activity of E3 ligase proteins [18, 20-23]. Recent studies proved evidence for a new mode of ubiquitination by SidE effector family: a mono-ADP-ribosyl transferase (mART) domain adds a ADP-ribose moiety to ubiquitin molecule by using an NAD molecule, and a phosphodiesterase (PDE) domain removes the AMP molecule from the ADPR moiety and ligates the remaining phosphoribosylated ubiquitin to target proteins; Rab33b and Reticulon 4 have been proven to be targeted by SidE family by this

modification [24-26]. Even though SdeA's function does is established, the question "how?" remains elusive.

This study aims to understand the molecular mechanism of SdeA and how the activity is regulated. In the second chapter, we (with help from a PhD student in Mao lab, Jon Wasilko) show how the SdeA protein is acting on the ubiquitin molecule. First, we have solved the crystal structure of SdeA protein with its mART, PDE and Mid domains that together form the functional core. The Mid domain was not expected to be there because previous bioinformatical studies did not predict a domain between mART and PDE domains [27, 28]. The Mid domain is localized between PDE and mART domains with two long loops connecting it to these domains. Initial thought was the possibility of Mid domain transporting ubiquitin between the two domains, but we have shown that PDE and mART domains can function separately. We also wanted to see the minimal active part of the mART domain and showed that mART can function only when Mid domain is localized to a certain position by a part of the PDE domain anchoring Mid domain. In short, our studies showed that this new Mid domain is indispensable for the mART activity of SdeA. How Mid domain participate in the first reaction remains elusive. Different constructs that extend more to the C' part of SdeA are being tested to see if it changes the activity. The essentiality of the Mid domain can also be tested by cloning the SdeA protein that lacks the Mid domain into *Legionella* and observing the survivability of the bacteria during infection. The mechanism of mART and how the protein recognizes Ub in the first place also remains elusive. More crystallization trials should be carried out to capture the mART domain with Ub and/or with NAD molecule. To help solve this, Small Angle X-Ray Scattering (SAXS) methods can also be utilized. With SAXS, the mART domain-ubiquitin interaction and the relative position of Ub to the active site can be observed.

To understand the molecular mechanism behind both mART and PDE reactions, we tried crystallization of mART and PDE domains extensively with their target molecules NAD, Ub and ADPR-Ub respectively. No crystals could be obtained. Previously, we had solved the crystal structure of another SidE protein family member, SdeD, that contains the PDE domain alone. When we compared the ubiquitin affinity of two PDE domains from SdeA and SdeD, we observed that PDE from SdeD can bind to ubiquitin whereas PDE domain from SdeA cannot. With the help of this binding information, we tried to and could obtain and solve the complex structure of SdeD with unmodified and ADPR-moiety ligated ubiquitin molecules. The ADPR moiety was captured within the active cleft of the protein and with further mutagenesis analysis, we had deciphered the reaction mechanism for the PDE activity. The mechanism of target protein recognition and transfer of PR-Ub to the target remain elusive. For this purpose, co-crystallization with the target proteins, i.e. Rab33b, can shed light on the mechanism. SAXS analysis can additionally give an idea to the binding site of target proteins to SdeA.

In the third chapter, we (with help from a Post-Doctoral researcher in Mao lab, Min Wan) tried to elucidate the activity of SdeD-SdeF. SdeD-SdeF are members of SidE family and contain the shared PDE domain alone [6]. We initially proposed that these proteins can have the same function as other SidE PDE domains. SdeD-SdeF were both found to be active towards ADPR-Ub to cleave the AMP molecule, but they were not active against target protein, Rab33b, or any other protein in the whole cell lysate. A recent study showed the deubiquitinase (DUB) activity of another effector protein, SidJ, towards the target proteins of SdeA [25]. With this information, we wanted to test the possibility of SdeD-SdeF acting as specific DUB for PR-ubiquitinated proteins. Indeed, when incubated with PR-ubiquitinated proteins, SdeD-SdeF can cleave the modification from the β-phosphodiester bond. It was verified both for a single target protein Rab33b and

modified proteins in the whole cell lysate experiment. When the two PDE domains of SdeD and SdeA are compared, each PDE domain has a loop over the active cleft that the other PDE domain does not. This structural difference can be the reason why one domain modifies the targets and one removes that modification. Replacing the loops on these PDE domains or crystallizing SdeA PDE domain with a target protein can help to find the structural reason of this difference. Two different activities of the same PDE domain within different members of the SidE family can be a regulatory mechanism as well. SdeA-C can PR-ubiquitinate different sets of proteins for the initial stages of infection and SdeD-SdeF can revert this modification for the progression of the infection after the LCV is established and *Legionella* starts to "hide" from the host cytoplasm. This phenomenon can be studied within infection conditions. If the PR modification of ubiquitin and PR-Ub modification of target proteins can be traced during infection, this should shed more light on the regulatory nature of this whole novel ubiquitination machinery.

During our studies for the localization of SdeA inside the mammalian cells, we observed that the Golgi is fragmented into smaller vesicles. Activity of SdeA was found to be crucial for this phenotype since the mutant SdeA proteins failed to create fragmentation. When SdeD-SdeF was present in the cell with SdeA, this fragmentation phenotype was reverted. We tried to find a link between ubiquitination and Golgi fragmentation and found out about a SNARE protein named Syntaxin 5 [29]. This SNARE protein was observed to be mono-ubiquitinated during early stages to inhibit its interaction with its corresponding SNARE molecule, Bet1 [30]. Within the later stages of mitosis, Syntaxin 5 was observed to be deubiquitinated and interact with Bet1 to start the defragmentation [31]. With this information, we wanted to see if SdeA can modify Syntaxin 5. We have shown the modification of Syntaxin 5 both *in vitro* and *in vivo* by SdeA. SdeD-SdeF can also revert the modification on Syntaxin 5 as well. This modification of Syntaxin 5 can be the link

between Golgi phenotype and SdeA activity. This actual link yet remains to be elucidated. Modification status of Syntaxin 5 can be monitored during infection. Golgi is reported to stay intact during infection [32], but SdeA may exert a short window of phenotype during infection that could have been missed by previous studies. A time-lapse monitoring of both Golgi phenotype and Syntaxin 5 modification may shed more light on the role of SdeA during infection.

In summary, in this thesis I have presented data that shows how SdeA modifies ubiquitin and target proteins. We have shown that PDE and mART domains can function independently and that a newly found Mid domain is essential for the mART activity. With the help of SdeD-Ub complex crystal structures, we have deciphered how the PDE domain functions mechanistically. In addition to the activity of SdeA, we have shown how SdeD-SdeF proteins can act on ADPR-Ub but not add PR-Ub moiety to target proteins. We provided evidence that SdeD-SdeF proteins can act as DUB molecules to remove the PR-Ub modification from SdeA targeted proteins. With a cellular biology approach, we have shown the Golgi fragmentation caused by SdeA and how this phenotype is reversed by SdeD-SdeF proteins. We have found a possible link, Syntaxin 5, between the Golgi phenotype and SdeA activity. Altogether, this protein suggests a roadmap to understand how the ubiquitin pathway is modulated by SidE family and gives an additional role for SidE family proteins during infection (Figure 4-1).

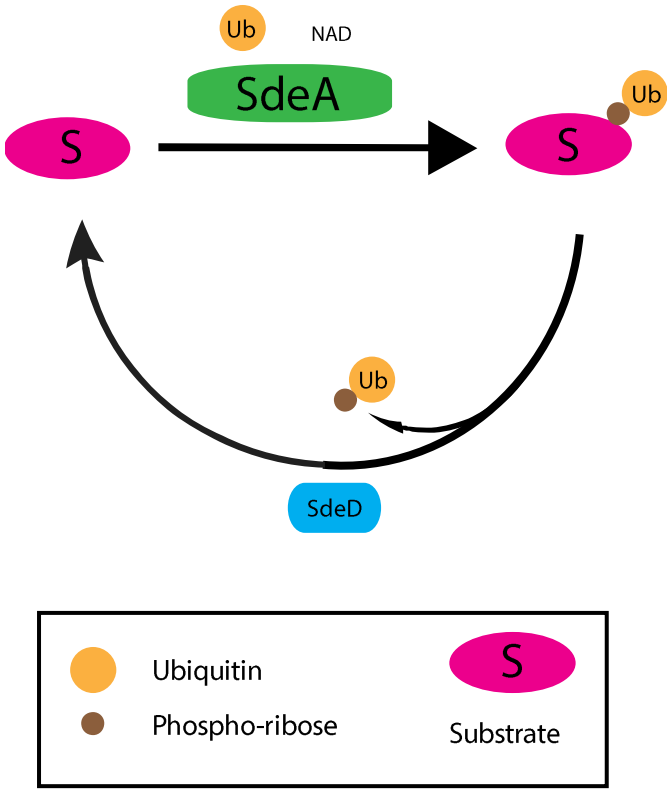


Figure 4-1: Substrate Modification by SidE Family. Diagram representation of how SidE family might modify target proteins to regulate their activity during infection.

REFERENCES

1. Fields, B.S., *The molecular ecology of legionellae.* Trends Microbiol, 1996. **4**(7): p. 286-90.

2. Escoll, P., et al., *From amoeba to macrophages: exploring the molecular mechanisms of Legionella pneumophila infection in both hosts.* Curr Top Microbiol Immunol, 2013. **376**: p. 1-34.

3. Kubori, T., et al., *Legionella metaeffector exploits host proteasome to temporally regulate cognate effector.* PLoS Pathog, 2010. **6**(12): p. e1001216.

4. 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.

5. Abu Khweek, A., et al., *The Sphingosine-1-Phosphate Lyase (LegS2) Contributes to the Restriction of Legionella pneumophila in Murine Macrophages.* PLoS One, 2016. **11**(1): p. e0146410.

6. 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.

7. Weber, S.S., C. Ragaz, and H. Hilbi, *The inositol polyphosphate 5-phosphatase OCRL1 restricts intracellular growth of Legionella, localizes to the replicative vacuole and binds to the bacterial effector LpnE.* Cell Microbiol, 2009. **11**(3): p. 442-60.

8. Choy, A., et al., *The Legionella effector RavZ inhibits host autophagy through irreversible Atg8 deconjugation.* Science, 2012. **338**(6110): p. 1072-6.

9. Ingmundson, A., et al., *Legionella pneumophila proteins that regulate Rab1 membrane cycling.* Nature, 2007. **450**(7168): p. 365-9.

10. Brombacher, E., et al., *Rab1 guanine nucleotide exchange factor SidM is a major phosphatidylinositol 4-phosphate-binding effector protein of Legionella pneumophila.* J Biol Chem, 2009. **284**(8): p. 4846-56.

11. Schoebel, S., et al., *RabGDI displacement by DrrA from Legionella is a consequence of its guanine nucleotide exchange activity.* Mol Cell, 2009. **36**(6): p. 1060-72.

12. Mukherjee, S., et al., *Modulation of Rab GTPase function by a protein phosphocholine transferase.* Nature, 2011. **477**(7362): p. 103-6.

13. Ivanov, S.S., *The tug-of-war over MTOR in Legionella infections.* Microb Cell, 2017. **4**(2): p. 67-68.

14. De Leon, J.A., et al., *Positive and Negative Regulation of the Master Metabolic Regulator mTORC1 by Two Families of Legionella pneumophila Effectors.* Cell Rep, 2017. **21**(8): p. 2031-2038.

15. Schroeder, G.N., et al., *Legionella pneumophila Effector LpdA Is a Palmitoylated Phospholipase D Virulence Factor.* Infect Immun, 2015. **83**(10): p. 3989-4002.

16. Toulabi, L., et al., *Identification and structural characterization of a Legionella phosphoinositide phosphatase.* J Biol Chem, 2013. **288**(34): p. 24518-27.

17. Hsu, F., et al., *Structural basis for substrate recognition by a unique Legionella phosphoinositide phosphatase.* Proc Natl Acad Sci U S A, 2012. **109**(34): p. 13567-72.

18. 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.

19. Hershko, A. and A. Ciechanover, *The ubiquitin system.* Annu Rev Biochem, 1998. **67**: p. 425-79.

20. Lin, Y.H., et al., *Host Cell-catalyzed S-Palmitoylation Mediates Golgi Targeting of the Legionella Ubiquitin Ligase GobX.* J Biol Chem, 2015. **290**(42): p. 25766-81.

21. Kubori, T., A.M. Hubber, and H. Nagai, *Hijacking the host proteasome for the temporal degradation of bacterial effectors.* Methods Mol Biol, 2014. **1197**: p. 141-52.

22. Perpich, J.D., et al., *Divergent evolution of Di-lysine ER retention vs. farnesylation motif-mediated anchoring of the AnkB virulence effector to the Legionella-containing vacuolar membrane.* Sci Rep, 2017. **7**(1): p. 5123.

23. 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.

24. Bhogaraju, S., et al., *Phosphoribosylation of Ubiquitin Promotes Serine Ubiquitination and Impairs Conventional Ubiquitination.* Cell, 2016. **167**(6): p. 1636-1649 e13.

25. Qiu, J., et al., *A unique deubiquitinase that deconjugates phosphoribosyl-linked protein ubiquitination.* Cell Res, 2017. **27**(7): p. 865-881.

26. 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.

27. 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.

28. Qiu, J. and Z.Q. Luo, *Legionella and Coxiella effectors: strength in diversity and activity.* Nat Rev Microbiol, 2017. **15**(10): p. 591-605.

29. Shorter, J., et al., *Sequential tethering of Golgins and catalysis of SNAREpin assembly by the vesicle-tethering protein p115.* J Cell Biol, 2002. **157**(1): p. 45-62.

30. Zhang, T. and W. Hong, *Ykt6 forms a SNARE complex with syntaxin 5, GS28, and Bet1 and participates in a late stage in endoplasmic reticulum-Golgi transport.* J Biol Chem, 2001. **276**(29): p. 27480-7.

31. Huang, S., D. Tang, and Y. Wang, *Monoubiquitination of Syntaxin 5 Regulates Golgi Membrane Dynamics during the Cell Cycle.* Dev Cell, 2016. **38**(1): p. 73-85.

32. Neunuebel, M.R., et al., *De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila.* Science, 2011. **333**(6041): p. 453-6.