NOVEL ACTIVITIES OF KINASE-FOLD ENZYMES FROM *LEGIONELLA PNEUMOPHILA*

APPROVED BY SUPERVISORY COMMITTEE

Vincent Tagliabracci, Ph.D.

Melanie Cobb, Ph.D.

Eric Olson, Ph.D.

Joshua Mendell, M.D., Ph.D.

DEDICATION

I dedicate this work to the outstanding mentors who have been sources of inspiration and tireless advocates throughout my training: Vinnie, for instilling an insatiable drive for discovery and nurturing an open environment of curiosity and collaboration unlike any other, Dr. Audrey Odom John and Dr. Douglas Chalker, my undergraduate mentors, for teaching me to strive for ambitious goals, and my parents, for always feeding my curiosity and even patronizing my very first scientific endeavor – an extensive documentation of snake specimens collected from New Jersey woodlands – before I was ten years old. Thank you all for being my proponents and exemplary role models.

NOVEL ACTIVITIES OF KINASE-FOLD ENZYMES FROM *LEGIONELLA PNEUMOPHILA*

by

MILES BLACK

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2020

NOVEL ACTIVITIES OF KINASE-FOLD ENZYMES FROM *LEGIONELLA PNEUMOPHILA*

Publication No._____

Miles Black

The University of Texas Southwestern Medical Center at Dallas, 2020

Supervising Professor: Vincent S. Tagliabracci, Ph.D.

Abstract

Protein kinases are fundamental mediators of cell signaling that transfer phosphate from ATP to their substrates. The protein kinase superfamily encompasses a vast and diverse trove of enzymes from all domains of life, including remote members that are barely recognizable by their primary amino acid sequence. SidJ (Substrate of *Icm/Dot* J) is a distant protein kinase homolog from the human

pathogen *Legionella pneumophila*. Contamination of water supplies with *Legionella* bacteria is a frequent cause of deadly pneumonia outbreaks (Legionnaire's disease). SidJ is a secreted *Legionella* virulence factor required for bacterial intracellular replication, but it is unknown how SidJ contributes to pathogenesis of Legionnaire's disease, or if SidJ has maintained the kinase fold or catalytic activity. In this work, I determine that SidJ is a calmodulin-binding protein which adopts a protein kinase fold. However, instead of phosphorylation, it catalyzes protein polyglutamylation. SidJ utilizes ATP to form an isopeptide bond between the amino group of free glutamate and the γ-carboxyl group of a glutamate of its substrate. During infection, SidJ polyglutamylates and inactivates a family of *Legionella* "all-in-one" ubiquitin ligases. Polyglutamylation is crucial step in the intracellular lifecycle of the bacterium and is required for full *Legionella* virulence in a eukaryotic host. SidJ reveals the unexpected catalytic versatility of the protein kinase fold, and highlights a unique strategy that pathogenic bacteria use to thrive within host cells. Interestingly, SidJ lacks key catalytic residues believed to be required for kinase activity. The discovery that SidJ is a polyglutamylating enzyme suggests that catalytically incompetent or 'pseudo' enzymes may lack activity only when assayed for the wrong reaction.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

1. **Black, M. H.,** Osinski, A., Gradowski, M., Servage, K. A., Pawlowski, K., Tomchick, D. R., & Tagliabracci, V. S. (2019). Bacterial pseudokinase catalyzes protein polyglutamylation to inhibit the SidE-family ubiquitin ligases. *Science*, 364(6442), 787-792. doi:10.1126/science.aaw7446

2. Ben Djoudi Ouadda, A., Gauthier, M. S., Susan-Resiga, D., Girard, E., Essalmani, R., **Black, M.,** . . . Seidah, N. G. (2019). Ser-Phosphorylation of PCSK9 (Proprotein Convertase Subtilisin-Kexin 9) by Fam20C (Family With Sequence Similarity 20, Member C) Kinase Enhances Its Ability to Degrade the LDLR (Low-Density Lipoprotein Receptor). *Arterioscler Thromb Vasc Biol*, 39(10), 1996-2013. doi:10.1161/ATVBAHA.119.313247

3. Cozza, G., Moro, E., **Black, M.**, Marin, O., Salvi, M., Venerando, A., . . . Pinna, L. A. (2018). The Golgi 'casein kinase' Fam20C is a genuine 'phosvitin kinase' and phosphorylates polyserine stretches devoid of the canonical consensus. *FEBS J*, 285(24), 4674-4683. doi:10.1111/febs.14689

LIST OF FIGURES

LIST OF TABLES

LIST OF DEFINITIONS

ACB – *Acanthamoea castellani* minimal salt buffer
ADP – Adenosine 5’ diphosphate
ADPR – ADP-Ribose
Amp – Ampicillin
AMP – Adenosine 5’ monophosphate
AMP-CPP – α,β-Methyleneadenosine 5′-triphosphate
AMP-PNP – Adenosine 5′-(β,γ-imido)triphosphate
Arg – Arginine
ART – ADP-Ribosyltransferase
Asp – Aspartate
ATP – Adenosine 5’ triphosphate
AYE ACES-buffered yeast extract (*Legionella* liquid growth media)
BLAST – Basic Local Alignment Search Tool
Bp – base pairs
BSA – Bovine Serum Albumin
CaM – Calmodulin
CCP – cytosolic carboxypeptidase
cDNA – coding DNA
CFU – Colony Forming Units
CLANS – CLuster ANalysis of Sequences
CTD – C-terminal Domain
CYE charcoal-buffered AYE (Legionella solid growth media)
DMEM – Dulbecco's Modified Eagle Medium
DMSO – dimethyl sulfoxide
DNA – Deoxyribonucleic acid
DTT – Dithiothreitol
DUB – De-Ubiquitinase
DupA/B – Deubiquitinase for pR-ubiquitination A/B
E1/E2/E3 – eukaryotic enzymes that catalyze ubiquitin activation and thioester cascade
EDTA – Ethylenediaminetetraacetic acid, metal chelator
EGTA – ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid, chelating agent more specific for calcium
ESI-MS – Electrospray Ionization-Mass Spectrometry
FBS – Fetal Bovine Serum
FcγRIIa – Fc receptor, binds immune complexes
FFAS – Fold and Function Assignment System
FIC – Filamentation induced by c-AMP
FLAG – epitope tag with the sequence DYKDDDDK
FPLC – Fast Protein Liquid Chromatography
GAPDH – Glyceraldehyde 3-phosphate dehydrogenase, loading control
GFP – Green Fluorescent Protein
Gln – Glutamine
Gly – Glycine
GroEL – bacterial chaperone, loading control
HA – hemagglutinin epitope tag with the sequence YPYDVPDYA
HEK293A – Human embryonic kidney cell line
His – Histidine
HPLC – High-performance liquid chromatography
Icm/Dot – Defective in Organelle Trafficking/Intracellular multiplication
IgG – Immunoglobulin G
IP – immunoprecipitation
IQ – Motif implicated in calmodulin binding
ITC – Isothermal Titration Calorimetry
Kan – Kanamycin
KD – Kinase Domain

LC-MS/MS – liquid chromatography/mass spectrometry
LCV – Legionella-containing vacuole
Leu – Leucine
Lp02 – wild-type strain of *Legionella* pneumophila
Lp03 – strain of *Legionella* pneumophila lacking essential components of the T4SS
Lpg – *Legionella pneumophila* gene
Lys – Lysine
MAFFT – Multiple Alignment using Fast Fourier Transform
MCS – multicloning site
MOI – Multiplicity of Infection
Ms – mouse
NAD+ – Nicotinamide Adenine Dinucleotide
NAM – Nicotinamide
NEM – N-ethylmaleimide
NF-κB – Nuclear factor kappa-light-chain-enhancer of activated B cells
Ni-NTA – Nickel-Nitrilotriacetic acid
NRD – N-terminal regulatory domain
NTD – N-terminal Domain
ORF – Open reading frame
PAGE – Polyacrylamide Gel Electrophoresis
PBS – Phosphate-Buffered Saline
PCR – Polymerase Chain Reaction
PDB – Protein Data Bank
PDE – Phosphodiesterase
PEG – Polyethylene Glycol
P_i – inorganic phosphate (PO_4^{3-})
PKA – Protein kinase A
PP_i – pyrophosphate ($P_2O_7^{4-}$)
pR-Ub – Phosphoribosyl-ubiquitin
PUP – Prokaryotic ubiquitin-like protein
PYG – Peptone/Yeast Extract/Glucose media for amoeba
R6K – origin of replication which requires the pi protein (pir gene) for replication
Rb – rabbit
RSF1010 – Broad host range plasmid
RT – room temperature, 23 ºC
S17 – 1 λpir – *E. coli* strain harboring the pi protein, can replicate R6K plasmids
sacB – levansucrase gene which confers lethality in the presence of sucrose
SD – Synthetic dropout (media)
SdeA/B/C – Paralog A/B/C of SidE
SdjA – Paralog A of SidJ
SDS – Sodium Dodecyl Sulfate
SelO – Selenoprotein O
SidE - Substrate of Icm/Dot E
SidJ – Substrate of Icm/Dot J
SUMO – Small Ubiquitin-Like Modifier
T4SS – Type 4 Secretion System
TCA – Trichloroacetic Acid
TCEP – Tris(2-carboxyethyl)phosphine, reducing agent
TEV – Tobacco Etch Virus
Thy – Thymidine
ThyA – thymidine synthetase gene that allows bacteria to grow on media lacking thymidine
TNNC1 – Troponin C
Trp – Tryptophan
TTL – Tubulin Tyrosine Ligase

Ub - Ubiquitin
ULP – Ubiquitin-like protease
Ura – Uracil
V5 – epitope tag, sequence KPIPNPLLGLDST
WT – Wild-Type
yCaM – Yeast calmodulin
YdiU – *E. coli* homolog of SelO
β-ME - β-mercaptoethanol

Acknowledgements and Contributions

Many individuals contributed directly to this work, which would not have been possible without the collaborative environment in the Tagliabracci Lab, the Department of Molecular Biology, and UT Southwestern at large. First, I would like to thank my co-authors. The experiments in this study were designed by myself and my mentor, Dr. Tagliabracci. Experiments were performed by myself, Adam Osinski, and V.S.T. Specifically, A.O. performed ITC measurements of SidJ-CaM binding, glutamylation initial rate measurements for structure-guided SidJ mutants, acyl-adenylate formation, and some glutamylation assays, including the SdjA glutamylation assay. V.S.T. performed Ub laddering experiments in mammalian cells, co-expression of SidJ, SdeA and CaM in *E. coli*, and provided critical technical support in all aspects of the project. Dr. Krzysztof Pawłowski and his student Marcin Gradowski (Warsaw University of Life Sciences) identified the kinase fold in SidJ and performed bioinformatic analyses. LC-MS/MS for protein ID and PTM discovery was performed by Dr. Kelly Servage. The work of these individuals is included in these pages and I deeply thank them for these contributions.

Next, I thank the outstanding core facilities at UT Southwestern and the individuals who operate them for essential expertise and advice, assay development, and training. Protein intact mass analysis by ESI/MS was performed by Dr. Andrew Lemoff (UT Southwestern Proteomics Core Facility). Quantification of adenine nucleotide amounts by LCMS/MS was performed by Dr. Noelle Williams (UT Southwestern Preclinical Pharmacology Core Facility). X-Ray structure modeling and refinement was performed by Dr. Diana Tomchick (UT Southwestern Structural Biology Core Facility). ITC was performed with the assistance of Dr. Chad Brautigam (UT Southwestern Molecular Biophysics Core).

Dr. Ralph Isberg (Tufts) generously provided *Legionella* strains.

I thank members of the Tagliabracci lab for advice and outstanding discussions: Dr. Anju Sreelatha, Victor Lopez, Brenden Park, Dr. Ting-Sung Hsieh, Samantha Yee, and Christine Nolan. I thank my thesis advisory committee for their guidance, and other senior scientists who provided insights and direction at the most critical moments of the project: Drs. Kim Orth, Margaret Phillips, Joe Goldstein and Neal Alto.

I thank Dr. Andrew Zinn and the Medical Scientist Training Program for support and community throughout my training. I am forever grateful for the people who work tirelessly behind the scenes to keep labs running and funding flowing, including Robin Downing, Linda Gunn, Carla Childers, Greg Urquhart, Stephen Johnson and Barbara Waters. Finally, I thank the funding sources that have supported my training, including individual fellowship F30HL143859 and training grants T32 GM007062-44 (Pharmacological Sciences) and 2T32GM008014-34 (MSTP) from the NIH.

CHAPTER ONE
Introduction

Enzymes with a protein kinase fold transfer phosphate from ATP to substrates in a reversible process known as phosphorylation. Krebs and Fischer first demonstrated in the 1960s that attaching a phosphate group to an enzyme can alter its activity[1], introducing a regulatory paradigm which is now understood to coordinate nearly all cellular responses to intracellular and external cues[2]. Approximately two-thirds of human proteins are modified by phosphorylation, establishing protein kinases as fundamental transducers of signaling. Kinase inhibitors have become the second-largest class of approved drugs, and kinase characterization remains a major focus of genomics-era research[3,4]. Kinases share the same core three-dimensional fold, and are identified by invariant catalytic residues that facilitate nucleotide binding and phosphotransfer[5]. Following the sequencing of the human genome, Manning and colleagues used these criteria to identify and classify nearly 500 kinases into the first draft of the human kinome.[6]

The proliferation of sequence data and modern bioinformatic strategies have enabled a search for novel protein kinases with remote sequence similarity to the 'canonical' human kinases. The group led by Krzysztof Pawłowski at the University of Warsaw has focused on manual bioinformatic identification of kinase domains so divergent that they escaped annotation by automated algorithms. The Tagliabracci lab then attempts to define catalytic activities and biological roles of these atypical

kinase families. This strategy has led to the discovery of ~30 new kinase families, including Fam20[7] CotH[8] HopBF1[9], Lpg2603[10], MavQ, and SelO[11] **(Figure 1)**. These discoveries highlight the structural diversity of the protein kinase superfamily and have uncovered new mechanisms of bacterial pathogenesis, kinase activation, and the cellular response to oxidative stress.

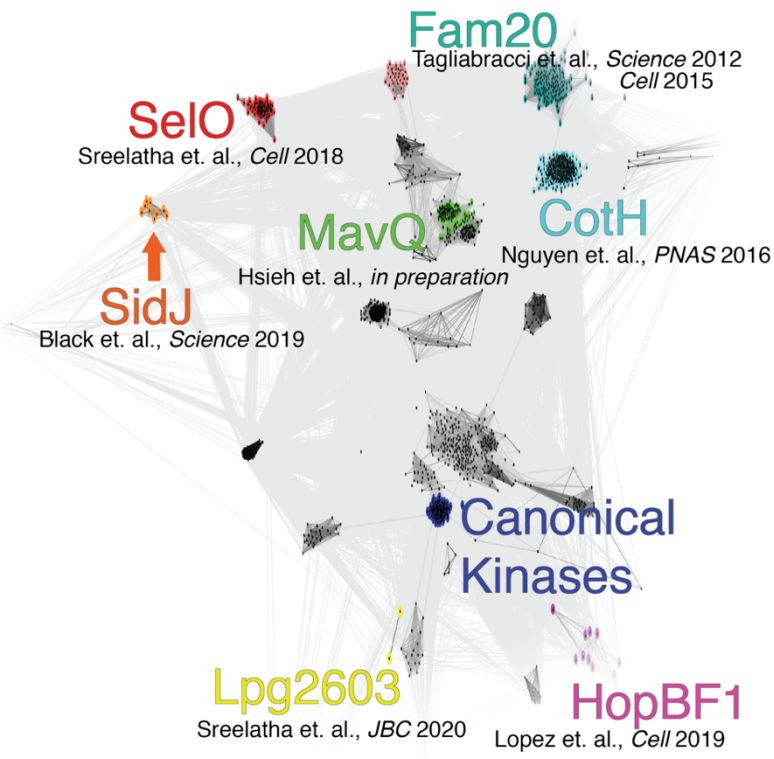


Figure 1. CLANS clustering analysis of selected atypical protein kinase families. Clustering is represented graphically as the network of BLAST-derived sequence similarities (edges) between representatives of known kinase families (dots). Select kinase families that the Tagliabracci lab has discovered and characterized are shown.

An unexpected outcome of this approach was the discovery that the protein kinase fold can catalyze reactions other than phosphorylation. The selenoprotein kinase SelO was predicted to be an inactive, or 'pseudo' kinase because it lacks one of the classic catalytic residues. Pseudokinases are widespread throughout nature, and account for 10% of the human kinome[12]. These 'dead' enzymes often play scaffolding or regulatory roles, but SelO revealed an unexpected catalytic activity: instead of phosphorylation, SelO uses ATP to catalyze the transfer of AMP onto protein substrates (AMPylation). This discovery was the first example of an enzyme with a protein kinase fold that catalyzes a reaction other than phosphorylation[13]. Furthermore, it revealed that AMPylation is part of the cellular response to oxidative stress[11].

The Tagliabracci lab, in collaboration with Pawłowski group, continues to search for new functions and catalytic activities for members of the protein kinase superfamily. The ultimate goal of this work is to uncover new paradigms of cell signaling and points for therapeutic intervention. This dissertation describes my work to characterize a novel kinase family, SidJ, from the pathogenic bacterium *Legionella pneumophila.*

CHAPTER TWO
Review of the Literature

Legionella pneumophila, an environmental pathogen

In July of 1976, an American Legion convention was held at the Bellevue-Straftford Hotel in Philadelphia, Pennsylvania. Days after its conclusion, convention attendees, mostly men in their 50s and 60s, began to rapidly succumb to nonspecific clinical symptoms such as fever, pain, and malaise. Within weeks, 182 people became sick and 29 died[14]. Using guinea pigs experimentally infected with human tissue samples, researchers identified the etiologic agent as a fastidious gram-negative bacillus[15]. *Legionella pneumophila* continues to cause outbreaks of serious disease, with health departments in the United States reporting nearly 10,000 cases in 2018[16].

Legionella is a remarkable pathogen which has evolved to shelter and proliferate within freshwater protozoans, organisms which normally graze on bacteria. After multiplying and consuming the nutrients of the protozoan, the progeny switch to a free-swimming, virulent phenotype and burst into the water supply to seek their next host[17,18]. When contaminated water is aerosolized and inhaled, *Legionella* bacteria can infect and replicate within human alveolar macrophages, causing disease[19] **(Figure 2).**

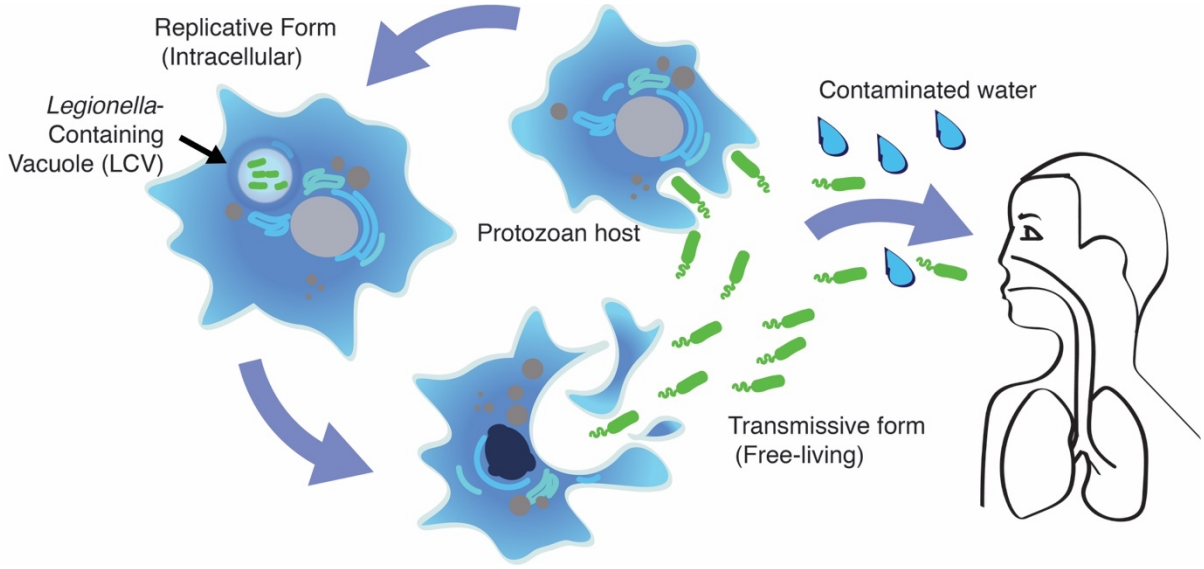


Figure 2. Life cycle of *Legionella*. *Legionella* bacteria (green) replicate in protozoan hosts (blue) within a membrane-bound compartment (LCV). Bacteria adopt a motile, infectious phenotype upon host cell egress. Aerosolized bacteria that are inhaled can cause Legionnaire's disease.

Legionella injects proteins into its host to promote infection

To survive and multiply within professional phagocytic cells, *Legionella* must suppress innate immune defenses, acquire nutrients, and prevent the apoptotic death of its host[20]. Ultimately, *Legionella* remodels the membranes of its host cell into a unique compartment within which it can replicate: the *Legionella*-containing vacuole (LCV)[21]. Like many other host-associated bacteria, *Legionella* transports macromolecules directly into the cytosol of the host to promote infection. A multiprotein machine called the Type 4 Secretion System (T4SS) threads unfolded protein cargo through a narrow pore, traversing the inner and outer bacterial membranes as well as the host plasma membrane[22]. The secretion system is a 27-component pore with 13-fold symmetry that has evolved from DNA conjugation

systems and uses at least three ATPases to power assembly and translocation[23]. Translocated proteins (effectors) harbor a C-terminal secretion signal which interacts with cytosolic coupling proteins, bringing the effectors into proximity of the core T4SS components. Once within the host cell, these effectors begin manipulating eukaryotic physiology to the benefit of the bacterium.

While the T4SS is well-conserved across different bacteria, the effector cargo varies considerably. Through horizontal gene transfer with its hosts and cohabiting bacteria, the *Legionella* genus has acquired over 18,000 translocated effectors representing 137 different eukaryotic-like domains[24]. *Legionella pneumophila* alone translocates more than 330 effectors, accounting for about 10% of its proteome[25]. This extensive effector repertoire is unmatched in the prokaryotic world. Because these effectors have evolved to target conserved processes, they represent an orthogonal approach to interrogate eukaryotic biology. Furthermore, effectors are a rich source of structural and biochemical diversity; of the 99 conserved protein domains identified in *Legionella pneumophila* effectors, 46 are entirely novel[25]. Even those effectors with recognizable protein folds sometimes catalyze unexpected reactions, such as a FIC domain that catalyzes phosphocholine transfer[26], a PP2C phosphatase that catalyzes de-AMPylation[27,28], and recently, an ADP-ribosyltransferase (ART) and phosphodiesterase (PDE) which together catalyze ubiquitination[29-31].

While the T4SS itself if required for virulence, most of the effectors are dispensable for intracellular replication[32]. This is thought to be due to redundancy between effectors[33]. A vast reserve of effectors likely helps the bacterium fine-tune the physiology of a breadth of protist hosts[34]. While several systems biology efforts have identified a 'core' set of effectors[25] and extensively minimized the Legionella genome[35], a true 'minimal set' of effectors required for virulence has yet to be identified.

Eukaryotic kinase domains in *Legionella* effectors

Eukaryotic kinase domains are among those appropriated by *Legionella* for its pathogenic intentions[36]. LegK1[37], LegK2[38], and LegK7[39] are well-known effector protein Ser/Thr kinases from *Legionella* that target the NF-κB pathway, the actin cytoskeleton, and Hippo signaling, respectively. A kinase domain was serendipitously discovered in the LepB effector which phosphorylates phosphatidylinositol lipids[40]. Two *Legionella* effectors appear to harbor tyrosine kinase domains[25] including Lpg2603, which is activated inositol hexakisphosphate (IP_6)[10]. A detailed search of the *Legionella* effector repertoire by the Pawłowski group uncovered several more with borderline similarity to protein kinases; these are active areas of investigation in the Tagliabracci lab. Among these candidates is an effector known as SidJ.

SidJ is a meta-effector of the SidE family

SidJ lies in a contiguous genomic locus with three effectors designated SdeA, SdeB, and SdeC. A fourth member, SidE, resides in a separate locus[32,41]. A SidJ homolog (SdjA) resides at yet another locus (**Figure 3**). Together, SdeA, SdeB, SdeC and SidE are referred to the SidE family, as they share the same domain architecture and up to 76% sequence identity. “Sid” (Substrate of Icm/Dot transporter) refers to their identification as some of the very first substrates interacting with components of the T4SS[31]. A growth defect caused by deletion of all four SidE members could be complemented by SdeA alone, revealing that the SidE effectors fulfil an important but likely redundant function during infection[42].

SidJ and SdjA are dissimilar to the SidE family, and it was apparent that they should perform a unique role during infection. A genetic interaction was soon identified that established SidJ as a negative regulator of the SidE family. In the absence of SidJ, excess SidE activity causes lethal intoxication of host cells and prevents bacterial replication[43,44]. This makes SidJ an example of a “meta-effector,” an effector which directly or indirectly modulates the activity of another. Three other well-characterized examples of *Legionella* meta-effectors have been documented[28,45,46], and a systems screen identified 14 additional genetic meta-effector pairs that await mechanistic decription[47]. At the time that the Pawłowski group identified a kinase fold in SidJ, the surprising activity of the SidE family itself had only recently been discovered.

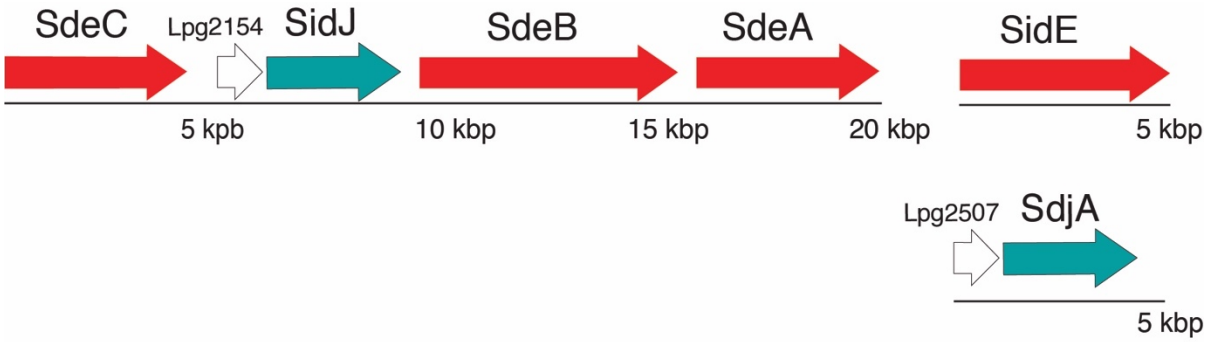


Figure 3. Genomic organization of the SidE, SidJ, SdjA, and neighboring effectors.

The SidE family catalyze non-canonical ubiquitination

The SidE family were thrust into the spotlight in 2016 with the discovery that SdeA can catalyze ubiquitination independent of E1, E2 or E3 ligases[29]. The mechanism of this noncanonical ubiquitination was subsequently revealed by a series of biochemical and structural studies[30,31,48-52]. The SidE effectors share the same domain organization. From the N-terminus, they encode a canonical deubiquitinase (DUB), a PDE, an ART, and a coiled-coil domain **(Figure 4, upper panel)**. The ART and PDE domains participate in non-canonical ubiquitination. While conventional ubiquitination is initiated by E1-catalyzed adenylation of the ubiquitin (Ub) carboxy terminus[53], the ART domain of SidE initiates ubiquitination by ADP-ribosylation of Ub on Arg42. Then, the PDE domain of SidE cleaves the phosphodiester bond to either release phosphoribosyl-Ub(pR-Ub) or link pR-Ub to serine residues of target proteins through a Ser-pR-Ub linkage, releasing AMP **(Figure 4, lower panel)**.

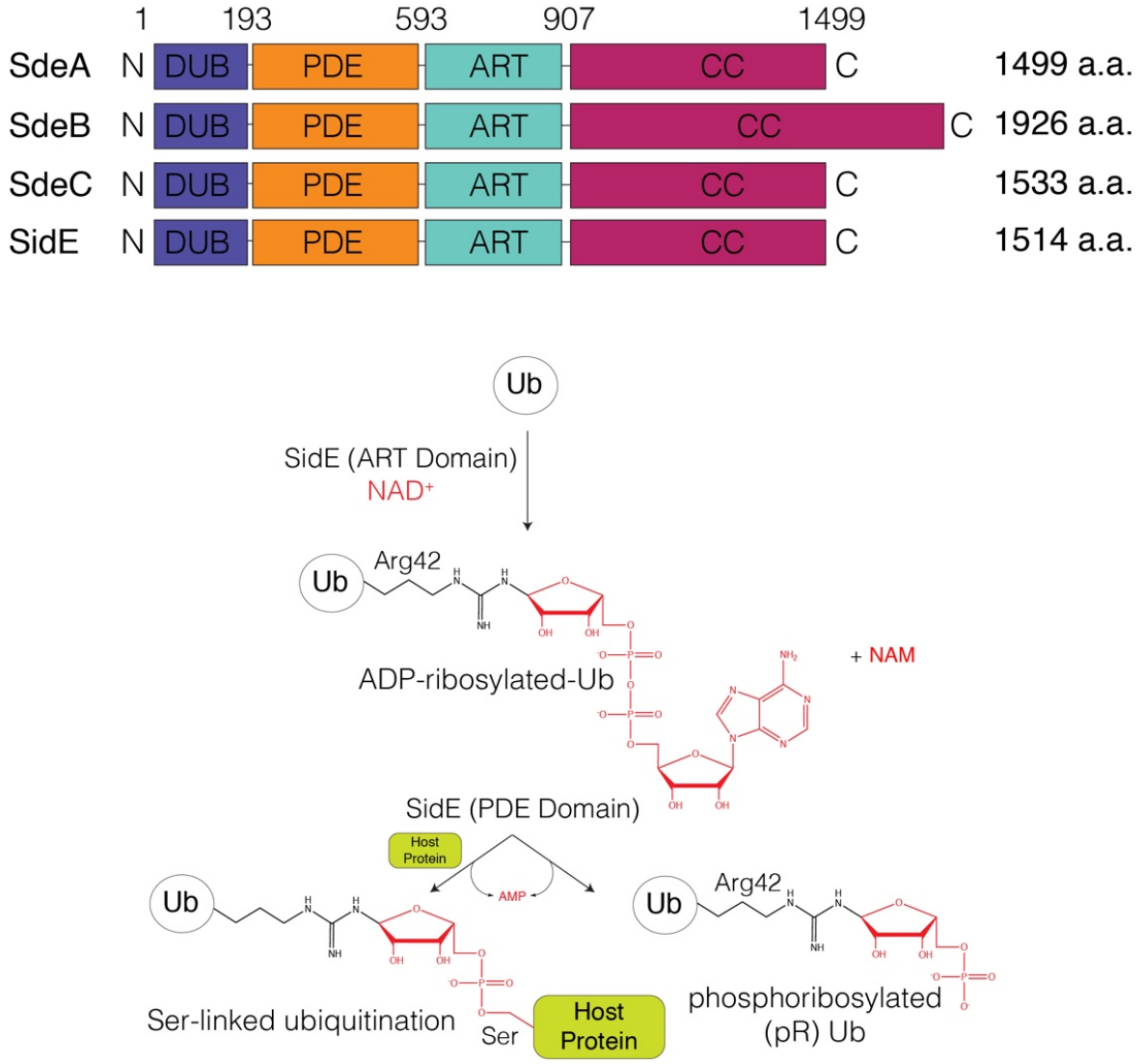


Figure 4. Domain architecture (upper panel) and reaction mechanism (lower panel) of the SidE effectors. DUB: deubiquitinase, PDE: phosphodiesterase, ART: ADP-ribosyltransferase, CC: coiled-coil. Amino acid (a.a.) boundaries for domains of SdeA are indicated, as reported in [50]. In the first step of the reaction, the ART domain in SidE uses NAD^+ as a cofactor to transfer ADP-ribose to Arg^{42} in Ub (white). The phosphodiesterase domain (PDE) in SidE then hydrolyzes the resulting

phosphodiester bond to generate phosphoribosylated (pR) Ub, or links the Ub to target proteins via a Ser-pR-Ub linkage. NAM; nicotinamide.

Almost 200 different host proteins have been identified as SidE substrates[29,30,48,49,54,55], most involved in vesicular transport and ubiquitination processes. SidE effectors are secreted early and localize to the *Legionella* vacuole minutes after the bacteria is engulfed, where they begin decorating the vacuole with pR-Ub-linked proteins[30,42]. This activity is required for proper formation of the LCV[49]. However, unrestrained SidE activity is harmful to the host and bacterium. pR-Ub, the side-product of the SidE reaction, fails to be activated by host E1 ligases. E1/E2-Ub conjugates that become phosphoribosylated stall the ubiquitination cascade. Cellular DUBs fail to cleave Ub conjugates modified by pR, and various types of polyUb chains accumulate[31]. Other *Legionella* effectors that manipulate the host Ub machinery may be blocked from functioning [31,43,44]. Together, these consequences intoxicate the host cell and prevent the bacterium from fully exploiting its resources.

The question of how SidJ antagonizes SidE remained open. SidJ may directly act on the SidE proteins, or indirectly suppress them by 'detoxifying' pR-Ub. Although one group claimed that SidJ possessed DUB activity against pR-Ub conjugates[56], this activity was only present in SidJ produced in *Legionella* and was likely due to co-purifying contaminants. The goal of this research was to determine if the kinase fold in SidJ possessed catalytic activity, and to establish the role of SidJ in *Legionella* virulence.

CHAPTER THREE
Methodology

Bioinformatic analysis of SidJ

The similarity of the SidJ protein to protein kinases was detected by analysing a collection of *Legionella* effector protein sequences using the FFAS system for remote sequence similarity recognition[57] and manually checking all non-significant similarities to kinases. The SidJ hit was verified using HHpred and Phyre2 servers for protein structure prediction[58,59]. According to standard significance thresholds specific for each method, similarity was not significant (or of borderline significance in the case of FFAS). The putative IQ CaM binding motif was also detected using the FFAS server. Location of the active site residues was predicted by analysing alternative FFAS alignments to different kinase-like proteins while considering sequence conservation among SidJ homologs and secondary structure predicted by the Jpred method[60].

To generate web logos of conserved kinase motifs in SidJ, homologues of the SidJ kinase-like domain were collected using Jackhmmer[61]. The sequences were aligned by MAFFT[62], the alignment was curated manually and sequence logos were created with Weblogo[63].

Cloning of *Legionella* effectors for expression in eukaryotic cells

The sequences of *Legionella* genes used in this study were obtained from the Philadelphia 1 strain genome[64] assembly GCA_000008485.1. Full-length open

reading frames (ORFs) of SidJ and SdeA were amplified by PCR using *Legionella pneumophila* Philadelphia-1 strain genomic DNA (gDNA) as a template. Primers were designed with adapters to ligate restriction enzyme sites and epitope tags 5' and 3' of amplified ORFs. SidJ (Lpg2155) was amplified with an N-terminal FLAG tag and cloned into a galactose-driven yeast expression vector, pESC-Leu (Agilent Technologies) to generate Flag-tagged SidJ. SdeA (Lpg2157) and SdeC (Lpg2153) were cloned into a modified galactose-driven yeast expression vector (pDGFP)[65] containing a C-terminal GFP-myc tag to generate SdeC-GFP. For expression in mammalian HEK293A cells, SidJ and SdeA ORFs were optimized for human codon bias and synthesized as gBlocks (SidJ; Integrative DNA Technologies, SdeA; Genscript). ORFs were amplified by PCR and cloned in frame with a C-terminal V5 tag (SidJ) or amplified with an N-terminal Myc tag (SdeA) into the CMV promoter-driven vector pcDNA (Invitrogen). pcDNA-HA-tagged UbB was a generous gift from Dr. Jenna Jewell and used as a template to generate the $Ub^{GG/AA}$ mutant.

Cloning

PCR products were amplified in 25 μL reactions containing 0.5 μM of each primer, template DNA (1ng of purified template or 1ng – 1μg genomic DNA) using Q5 DNA polymerase (New England Biolabs) with 200 µM of each deoxynucleotide, 2.0 mM Mg^{2+} and buffer components as specified by the manufacturer. An initial denaturation (98°C) of 30 s (for purified templates) or 3 min (genomic DNA) was followed by 32 – 36 cycles of denaturation (98°C for 10 s) annealing (temperature as

determined for each primer pair using the NEB Tm calculator v1.12.0) for 20 s, and extension at 72°C for 30 s/kbp. Final extension was performed for 3 min. PCR products were bound to silica matrix spin columns (Zymo research) and cleaned of contaminating reaction products using the supplied buffers. The cleaned PCR product was subjected to restriction enzyme digestion in 30 μL volumes according to the manufacturer's instructions for each enzyme combination (New England Biolabs). Destination plasmid DNA (2-5 μg) was cut in the same conditions. Digested DNA was resolved by electrophoresis in 1% TAE-agarose gels. Band were excised from the gel with a razor and DNA was recovered using the Zymoclean gel DNA recovery kit (Zymo Research) and eluted in 15 μL volume. Ligations (5.5 μL inserts mixed with 2.5 μL digested vector for a typical 0.5-4 kbp insert) were performed 10 min to overnight in 10 μL volume using T4 DNA ligase (New England Biolabs). 5 μL of the ligation mixture was transformed into chemically competent DH5α *E. coli* by 45 s heat shock at 42°C. Cells were briefly incubated on ice, then allowed to recover rotating at 37°C for 15 min (ampicillin selection) or 30 min (kanamycin selection) before plating on prewarmed LB plates with ampicillin (100 μg/mL final) or kanamycin (50μg/mL final). To recover plasmid DNA from *E coli*, single colonies were picked and used to inoculate 5 mL of LB with appropriate selection. Cultures were grown overnight, then collected by centrifugation. Plasmid DNA was purified using the GeneJet miniprep kit (Thermo). All constructs were verified by sanger sequencing.

Site-directed mutagenesis

Amino acid mutations were introduced via Quick Change site-directed mutagenesis. primers were designed using the Agilent Quick Change Primer design tool: https://www.genomics.agilent.com and used in PCR reactions with PfuTurbo DNA polymerase (Agilent) in 50 μL reactions with 50ng template DNA according to the manufacturer's protocol, but with DMSO added to 4%. Extension times were increased to 30 min and 12 to 15 cycles were performed in total. Reaction products were digested with DpnI endonuclease overnight before 5 μL was transformed into DH5α *E. coli* as described. Mutations were confirmed by Sanger sequencing.

Culture of *Saccharomyces cerevisiae*

All experiments in this study were performed with the parent strain BY4741[*Mata leu2Δ0 met15Δ0 ura3Δ0 his3Δ1*]. BY4741 was maintained on YPD (20 g/L peptone, 10 g/L yeast extract, 2% glucose) agarose plates at 30°C. Yeast expression plasmids were introduced into BY4741 via lithium acetate-mediated transformation[66]. Transformants were selected by plating on SD dropout media (6.7 g/L yeast nitrogen base without amino acids, 1.4 g/L yeast synthetic drop-out medium supplement) containing 2% [w/v] glucose. Amino acids His (0.02 mg/mL), Trp(0.02 mg/mL), Ura (0.02 mg/mL), and Leu (0.1 μg/mL), were added or omitted to maintain pESC Leu (prototrophy for leucine) or pDGFP (prototrophy for uracil). For solid media, 2% w/v agar and 0.05% v/v 5M NaOH was added to the media before autoclaving.

Culture of mammalian cells

HEK293A cells were grown in DMEM containing 10% (vol/vol) FBS with 100 µg/mL penicillin/streptomycin (GIBCO) at 37 °C with 5% CO2. Cells were routinely tested for Mycoplasma contamination by PCR based methods.

Expression of SidJ and SidE effectors in yeast

For growth inhibition assays by re-streaking agar plates, yeast were grown on the appropriate SD agar dropout plates supplemented with 2% glucose. A single colony collected with a sterile inoculating loop and struck onto plates supplemented with either 2% glucose or 2% galactose and 1% raffinose. Restruck plates were grown at 30°C for 2-5 days.

For spot dilution assays, a streak of yeast harboring SidJ, SdeC or SdeA from SD glucose plates was used to inoculate the appropriate SD dropout media supplemented with 2% glucose and grown at 30°C overnight in an orbital shaker set to 250 rpm. The following day, cultures were normalized to 1 OD_{600} in sterile water and diluted serially to 10^{-3} OD_{600}. 10 µL of each dilution were then spotted on the appropriate SD dropout agar plates using either 2% glucose or 2% galactose and 1% raffinose. Spotted plates were grown at 30°C for 2-5 days.

Expression of Ub, SidJ and SidE in mammalian cells

Near-confluent HEK293A cells were collected and plated into individual wells of a 6-well dish at ~35-50% confluency. The following day, individual wells were transfected with pcDNA-HA-Ub, pcDNA-HA-$Ub^{GG/AA}$, pcDNA-SidJ-V5 (or mutants),

pcDNA-Myc-SdeA (or mutants), or empty vector using PolyJet transfection reagent (SignaGen) with a total of 2.1μg plasmid DNA per well. Individual plasmids were used at the following concentrations: HA-Ub (1 μg): SidJ-V5 (1μg): myc-SdeA (0.1 μg). The premixed DNA was added to 100 μL of serum free, antibiotic free DMEM containing 8 μL of Polyjet transfection reagent and the mixture was incubated at RT for 20 min before it was added dropwise to cells. The culture medium was replaced 5 h after transfection with DMEM containing 10% FBS and PenStrep. ~16-20 h after transfection cells were analyzed for protein expression.

Preparation of cell lysates and detection of proteins by western blot

To verify protein expression in yeast, yeast cultures were grown overnight in the appropriate SD dropout media supplemented with 2% glucose. The overnight cultures were then diluted in the appropriate SD dropout media with 2% galactose and 1% raffinose. 1 mL of the induced culture (1 OD_{600}) was pelleted and resuspended in 100 μL yeast lysis buffer (4% 5N NaOH and 0.5% β-MErcaptoethanol (β-ME). The cells were incubated for 30 min on ice and then neutralized with 5N HCl (~5 μL). SDS-PAGE loading buffer was added to 1x (12.5 mM TrisPO4 pH 6.8, 10% (w/v) glycerol, 1.25% (w/v) SDS, 0.02% (w/v) bromophenol blue and 1% β-ME). To detect protein expression in mammalian cells, ~16-20 h after transfection, cells were washed twice with ice cold PBS and lysed directly on the plate with 2X SDS-PAGE loading buffer.

Proteins were resolved by SDS-PAGE. To resolve high-molecular weight proteins, including SidJ and SdeA, 6% acrylamide gels were prepared. To resolve Ub and CaM, 12% acrylamide was used. Proteins were electrophoresed at 185 volts in SDS running buffer (3.03 g/L Tris base, 14.4 g/L glycine, 1 g/L SDS) for 55 min. Proteins were then transferred to nitrocellulose membranes in transfer buffer (3.03 g/L Tris base, 14.4 g/L glycine, 20% v/v MeOH) at a constant current of 370 mA. Immunoblotting was performed as follows: the nitrocellulose membranes were blocked by rocking in 5% fat-free milk in TBS-T (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween-20) for 30 min. Then, primary antibody diluted in 2% milk/TBS-T was applied for 1 h at RT or overnight at 4C. Milk solutions were filtered through cheesecloth before use. Primary antibodies were typically used at 1:5000 dilution. After three ten-minute washes in TBS-T, HRP-conjugated secondary antibody diluted 1:5000 in 2% milk was applied for 1 h at RT. After three final ten-minute washes in TBS-T, membranes were incubated for five min in Enhanced chemiluminescence reagents before exposure to X-ray film.

Identification of the SidJ-CaM interaction by immunopurification and mass spectrometry

Five mL of yeast were induced to express Flag-SidJ by growing overnight in SD dropout media with 2% galactose and 1% raffinose. The whole culture was pelleted the following morning and resuspended in approximately an equal volume of ice-cold yeast immunoprecipitation (IP) buffer (50 mM Na-HEPES, 200 mM NaOAc, 1

mM EDTA, 1 mM EGTA, 5 mM MgOAc, 5% glycerol, 0.25% NP-40, 3 mM DTT, 1 mM PMSF, Roche protease inhibitor cocktail (PIC), pH 7.5). Cleared cell lysates were obtained by beating with acid-washed glass beads, followed by two subsequent spins at 3,000 xg (2 min. at 4°C) and 20,000 xg (10 min. at 4°C). Lysates were then incubated on an orbital shaker for 3 h at 4°C with Anti-Flag M2 agarose affinity resin that had been pre-blocked with 1% BSA in IP buffer. After 3 h, the agarose was pelleted and washed four times with ice-cold yeast IP buffer without protease inhibitors. Immunoprecipitated proteins were then eluted using 1XFlag peptide diluted to 100 μg/mL in yeast IP buffer and SDS loading buffer containing 1% β-ME was added to the samples and boiled for 10 min. Samples were run into an SDS-PAGE gel for concentration and the entire sample was excised with a razor for protein identification by mass spectrometry.

For IP of SidJ-V5 from HEK293A cells, 3 X 10^6 cells were seeded into 10 cm dishes and transfected ~18 h later with empty vector, codon optimized pcDNA-SidJ-V5, or pcDNA-SidJ-V5 with mutations in kinase or CaM-binding residues using PolyJet transfection reagent (4.5 μg DNA / 35 μL PolyJet, 200 μL of DMEM). Cells were harvested the following day and lysed in 1.5 mL ice-cold mammalian IP buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100) supplemented with 1 mM PMSF and 1 X protease inhibitor cocktail (PIC; Roche) by intermittent vortexing for 10 min on ice. Cell lysates were cleared by centrifugation at 20,000 xg for 10 min at 4°C, at which point aliquots of the extract were collected and boiled in 1X SDS loading

buffer + 1% B-ME. 1 µL mouse anti-V5 antibody was added to the cleared lysate and the samples were incubated overnight at 4°C with nutating. The following morning, 50 µL resin volume of Protein G agarose pre-blocked with 1% BSA in IP buffer was added to each sample and the antigen-antibody complex was bound for 2h at 23°C with nutating. The agarose was collected by centrifugation and washed 4 times with cold mammalian IP buffer. After the final wash, mammalian IP buffer was removed and the agarose was resuspended in 2X SDS loading buffer + 2% β-ME and boiled. Samples were resolved by SDS-PAGE, transferred to a nitrocellulose membrane, and immunoblotted with rabbit anti-V5, rabbit anti-CaM and mouse anti-GAPDH antibodies.

Anti-Flag immunoprecipitates from yeast extracts, prepared as described above, were reduced with DTT and alkylated with iodoacetamide prior to overnight enzymatic digestion with trypsin at 37°C. Tryptic peptides were de-salted via solid phase extraction (SPE) prior to LC-MS/MS analysis. Experiments were performed on a Thermo Scientific EASY-nLC liquid chromatography system coupled to a Thermo Scientific Orbitrap Fusion Lumos mass spectrometer. MS1 spectra were acquired in the Orbitrap mass analyzer (resolution 120,000) and precursor ions were subjected to high-energy collision-induced dissociation (HCD) for fragmentation. MS2 spectra of fragment ions were collected in the ion trap. MS/MS spectral data from anti-Flag immunoprecipitates from yeast expressing Flag-tagged SidJ, SidJ^{K367A}, and empty

vector control were searched using the Mascot search engine (Matrix Science) against entries included in the *S. cerevisiae* Uniprot protein database for peptide identification.

Culture of *Legionella* bacteria

L. pneumophila strains Lp02, Lp03 (Lp02 ΔdotA), and thymidine auxotrophic derivatives used in this study were derived from Legionella pneumophila Philadelphia-1 strain (27) and were generous gifts from Dr. Ralph Isberg. Legionella bacteria were maintained on ACES [N-(2-acetamido)-2- aminoethanesulfonic acid]-buffered charcoal yeast extract (CYE) agar plates (10g/L ACES, 10g/L yeast extract, 2g/L charcoal), or grown in ACES-buffered yeast extract (AYE) liquid cultures supplemented with ferric nitrate (0.135 g/L) and cysteine (0.4 g/L). Thymidine was added to a final concentration of 100 µg/mL for maintenance of the thymidine auxotrophic strains. Single colonies grew in approximately four days under these conditions. For experiments where colonies were counted, the plates were incubated in a humidified chamber at 37°C.

Generation of gene deletions in *Legionella*

SidJ knockout strains were generated using the R6K suicide vector pSR47s (Kan^R, sacB), a generous gift from Dr. Shaeri Mukherjee, UCSF. ~800bp regions flanking the SidJ ORF were amplified from *Legionella* gDNA and cloned using Gibson assembly into pET-21a(+), then subcloned into pSR47s (incompatible with DH5α host) to generate pSR47s-Δ*sidJ*, which was transformed by electroporation into S17-1 λpir *E. coli*. A heavy patch of Lp02 ThyA- was grown on CYE+Thy for

two days, then the whole patch of *Legionella* was collected and washed three times in sterile H_2O then resuspended in 10% glycerol. pSR47s- Δ*sidJ* was then introduced by electroporation into Lp02 ThyA-. Following a 6 h recovery in AYE broth at 37ºC, the *Legionella* were plated on CYE +Thy +Kan (20 µg/mL) to select for colonies having undergone homologous recombination. Metrodiploids were resolved by heavily streaking Kan^R clones on CYE +Thy + 10% sucrose. Resulting colonies were re-struck to CYE +Thy and screened for loss of SidJ by PCR and protein immunoblotting. Single colonies appeared in four days at each step. The Δ*sidJ*, Δ*sdjA*, Δ*sdeA* (ΔΔΔ) strain was generated by successive deletions of SidJ, then SdjA, then SdeA. Complementing strains were generated using the RSF1010 cloning vector pJB908 (Amp^R, ThyA*, tdΔi*), a generous gift from Dr. Ralph Isberg. SidJ, SdjA, and mutants were amplified with primers to ligate a shine-delgardo sequence upstream of the start codon (TCTAGATAAATATTTGAATTT) and cloned into pJB908 via BamHI and SalI sites. Complementation vectors were transformed by electroporation into Lp02 ThyA- and Lp02 Δ*sidJ* ThyA-. Transformants were selected on CYE agar without thymidine and SidJ complementation was verified by PCR and protein immunoblotting. Transformations with pJB908 are very efficient and 10^{-3} dilutions of the transformation mixture were usually required to obtain single colonies.

Culture of *Acanthamoeba castellani*

Acanthamoeba castellanii was maintained as a monolayer culture in PYG medium (20 g/L protease peptone, 1 g/L yeast extract, 150 mM glucose, 4 mM Mg_2SO_4, 0.4 mM $CaCl_2$, 0.1% (w/v) sodium citrate dihydrate, 0.05 mM $Fe(NH_4)^2(SO_4)^2$ x $6H_2O$, 2.5 mM NaH_2PO_3, 2.5 mM K_2HPO_3 pH 6.5) in 175 cm^2 tissue culture flasks at 23°C (atmospheric CO2), in a humidified chamber. Cultures were split 1:5 every four days into fresh flasks. To collect amoeba trophozoites, PYG media was aspirated and the flask was incubated on ice for 5 min. Amoeba were dislodged by pipetting in *A. castellani* buffer (ACB, which is PYG with peptone, yeast extract, and glucose omitted). Cells were pelleted by centrifugation and resuspended in fresh PYG.

Challenge of mammalian cells with *Legionella*

To allow non-phagocytic HEK293A cells to take up *Legionella*, cells were transfected with the IgG receptor FcγRIIa. FcγRIIa was amplified from the Ultimate™ ORF Lite human cDNA collection (Life Technologies) and cloned into pCDNA with a C-terminal flag tag. Dishes (15 cm) of HEK239A cells at ~70% confluency were transfected with pCDNA- FcγRIIa and pCDNA-Myc-SdeA with PolyJet transfection reagent (10 μg each plasmid, 80 μL PolyJet, 750 μL serum-free DMEM). In separate experiments to monitor pR-ubiquitination, 10 cm dishes of HEK239A cells were transfected with pCDNAFcγRIIa and pcDNA-HA-$Ub^{GG/AA}$ with PolyJet transfection reagent (4.5 μg each plasmid, 35μL 12 PolyJet, 200μL serum-

free DMEM). The following day, the medium was changed to serum-free, antibiotic-free DMEM.

To obtain virulent *Legionella* in post-exponential phase, a loopful of bacteria were scraped from a 2-day heavy patch on CYE to inoculate 2 mL AYE broth. Serial dilutions (four 1:3 dilutions) of the inoculum were grown 18-24 h at 37°C with shaking to an OD_{600} of 3.7 to 4.2, at which time the cultures acquired a brown pigmentation and the bacteria were judged to be >70% motile by brightfield microscopy. To estimate CFU, one OD_{600} unit was assumed to correspond to $1x10^9$ cfu/mL. All MOIs were confirmed by plating serial dilutions of the inoculum at the time of the challenge. For challenges, the bacteria were opsonized with rb anti-*Legionella* antibody (ab20943, Abcam). Legionella were diluted into 2.5 mL serum-free DMEM with 1:1000 Rb anti-legionella to $6x10^8$ CFU/mL. Bacteria were opsonized for 30 min at 37°C on a rotator. Opsonized bacteria (approximately $1.5x10^9$ total) were added dropwise to the confluent HEK293 cells for a final MOI of approximately 100. The infection was allowed to proceed for 1 h at 37°C in a tissue culture incubator, then the monolayer was carefully washed 3x in PBS to remove bacteria, and the cells were processed.

Intracellular replication of *Legionella* in amoeba

Eight h prior to infection, confluent amoeba monolayers were collected, resuspended in fresh PYG, counted, and $6×10^5$ cells were seeded into individual wells of 24-well tissue culture plates. 1 h prior to infection, amoeba were carefully

washed twice, the medium was replaced with ACB, and the plates were equilibrated to 37°C. All subsequent incubations were performed at 37°C. *Legionella* cultures at post-exponential phase were grown as described, then diluted in ACB. ~6×10^4 bacteria were added to each well for a MOI of 0.1. Infections were synchronized by centrifugation at 880 xg for 5 min (time = 0). Infections were allowed to proceed for 1 h, then extracellular bacteria were removed by washing each well 3 times in ACB before adding ACB to a final volume of 0.5 mL/well. At timepoints 1 h, 24 h, and 48 h, infected Amoeba cells were lysed in 0.05% saponin in H_2O. Serial dilutions of the infectious inoculum and the amoeba lysate were plated on CYE plates to confirm the MOI and assess bacterial growth.

Cloning eukaryotic and *Legionella* open reading frames for expression in *E. coli*

SidJ, full-length and residues 59-851 (SidJ$^{\Delta NC}$), SdjA residues 36-786, yeast CaM (cmd1), human CaM (calm2), and the SidE effectors (and indicated truncations) were cloned into a modified pet28a bacterial expression vector (ppSUMO), containing an N-terminal 6X-His tag followed by the yeast SUMO (smt3) CDS. The truncation SdeA 178-1100 (SdeA$^{\Delta NC}$) retains pR-Ub activity and was used for most experiments. HA-Ubiquitin B, CaM, Troponin C (TNNC1), SidJ K367A (334-C) and SdeA 518-1100 (R766A) were also cloned into pProEX2 containing an N-terminal 6X-His tag followed by a TEV protease cleavage site. For co-expression with ppSUMO-SdeA H407A, full-length untagged SidJ (or mutants) and human and

yeast CaM were also cloned into the bicistronic petDuet1 vector (Novagen) for coexpression experiments in *E. coli.*

Purification of recombinant protein from *E. coli*

For purification from *E. coli,* plasmids were transformed into Rosetta DE3 cells. A swath of bacteria was collected and used to inoculate starter cultures (10 mL / L final) in LB broth under appropriate selection at 37°C in an orbital shaker set to 250 rpm. After 3-4 h, large volumes of LB were inoculated and grown for ~5 h to an OD_{600} of 0.6-1.1. Bacteria were cooled to room temperature before protein expression was induced with 0.4 mM IPTG for 16-18 h at room temperature with orbital shaking at 250 rpm. Cells were harvested by centrifugation at 4,000 xg for 15 min and resuspended in 50 mM Tris-HCl pH 8, 300 mM NaCl, 1 mM PMSF, 1 mM DTT at $1/10^{th}$ volume. Cells were lysed by sonication in 40 mL volumes with a microtip set to deliver five 30 s pulses at 40% amplitude, each pulse followed by a 1 min recovery on ice. Cell lysates were cleared by centrifugation at 30,000 -35,000 xg for 30 min. The cleared lysate was incubated with Ni-NTA beads for a minimum of 1 h at 4°C, nutating. Ni-NTA beads were washed once each with water and TBS (in 50 mM Tris-HCl pH 8, 300 mM NaCl) (10 mL each) before lysates were bound. Beads were collected in a 20 mL gravity-flow column and washed with 20 mL of 50 mM Tris-HCl pH 8, 300 mM NaCl, 25 mM imidazole, 1 mM DTT. Proteins were eluted with 50 mM Tris-HCl pH 8, 300 mM NaCl, 300 mM imidazole, 1 mM DTT. SUMO-tagged proteins were cut overnight at 4°C with 6X-His tagged ULP SUMO protease

(~2 μg/mL) followed by concentration in a 50,000 Da-cutoff centrifugal filter. Aggregates were removed by a high-speed spin at 20,000 xg for 10 min. Proteins were further purified by gel filtration chromatography using a Superdex 200 gel filtration column attached to an AKTA Pure FPLC chromatography system (GE Healthcare) in running buffer (50 mM Tris-HCl pH 8, 300 mM NaCl, 1 mM DTT). Samples from each fraction were resolved by SDS-PAGE to asses purity and separation. Peak fractions were pooled and concentrated in a centrifugal filter. Proteins were diluted into 50 mM Tris-HCl pH 8, 300 mM NaCl, 1 mM DTT with 5% glycerol, aliquoted, and snap-frozen in liquid nitrogen for storage at -80°C. For purification of ADPR-HA-Ub, 6X His-HA-Ub (in pPROEX2) was purified as above, then subjected to ADP-ribosylation assays in a 1 mL reaction volume containing 2 µM SdeA 519-1100 and 1 mM NAD+ overnight at room temperature. The products were purified by gel filtration chromatography and the fractions containing ADPR-HA-Ub were collected and concentrated.

Production of SidJ and SdeA antibodies

Production of SidJ and SdeA antibodies L. pneumophila SidJ K367A (334-C) and SdeA R766A (518-1100) were purified as a 6X His tagged fusion proteins as described above. The proteins were further purified by Superdex 200 gel filtration chromatography and used to inoculate rabbits for generation of rabbit anti-SidJ and anti-SdeA anti-serum (Cocalico Biologicals). Total IgG was partially purified by ammonium sulfate precipitation (32) and the a-SidJ and a-SdeA antibodies were

affinity-purified by coupling WT untagged SidJ (334-C) and WT untagged SdeA (518-1100) to a HiTrap NHS-activated HP column essentially as described (33). Antibodies were concentrated, aliquoted and stored at -20°C until use.

Kinase and AMPylation assays of SidJ

Assays with [γ-^{32}P]ATP and [α-^{32}P]ATP were performed with SidJ 334-C or the K367A mutant in a reaction mixture containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 100 µM [γ-^{32}P]ATP or [α-^{32}P]ATP (specific radioactivity = 1000 cpm/pmol), 1 mM DTT, 100 µg/mL bovine CaM (Sigma), and 100 µg/mL SidJ or YdiU protein. Myelin basic protein (MBP) was included as a generic kinase substrate at 100 µg/mL. Reactions were incubated at 37°C for 1 hr and terminated by adding EDTA to 37 mM. SDS loading buffer was then added to the samples and boiled. Reaction products were resolved by SDS-PAGE and stained with Coomassie blue. The gels were then fixed in 30% MeOH, 10% glycerol, and dried and imaged with autoradiography film.

Calmodulination assays

30 μL reactions contained 1 μg SUMO-SidJ (full-length), 5 μg SdeA$^{178\text{-}1100}$, 6 μg Calm2, 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1mM DTT, and were initiated with ATP (100 μM final), $MgCl_2$ (5 mM final) and $CaCl_2$ (1 mM final). In some reactions, sodium pyrophosphate or sodium phosphate was added to the protein mix before the ATP mix. Reactions were incubated at 37ºC for 2.5 h, then terminated by boiling in 5X SDS-PAGE loading buffer.

ATP-Pyrophosphate exchange assay

Back-exchange reactions were performed essentially as described in[67]. Fifty μl reactions contained 2.5 μg SidJ, 2.5 μg calm2, 50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1mM DTT and were initiated with a mix of ATP (2 mM final) [^{32}P] sodium pyrophosphate (0.2 mM, S.A. 500 cpm/pmol final), $MgCl_2$ (5 mM final) and $CaCl_2$ (1 mM final). In some reactions, N-ethylmaleimide was added to the protein mix before nucleotide/PP_i. The reaction was incubated at 37ºC for 25 min before termination with 250 μL stop solution (1.2% activated charcoal, 0.1M sodium pyrophosphate, 0.35M perchloric acid). Samples were mixed thoroughly, then centrifuged to collect the charcoal. Charcoal was washed 3 times in H_2O, then the whole microfuge tube was transferred to scintillation vials for Cherenkov counting.

Protein intact mass analysis

Proteins in solution were analyzed for intact mass by the UT Southwestern Proteomics Core Facility. LC/MS was performd using a Sciex X500B Q-ToF mass spectrometer coupled to an Agilent 1290 Infinity II HPLC. Samples were injected onto a POROS R1 reverse-phase column (2.1 x 30 mm, 20 μm particle size, 4000 Å pore size), desalted, and the amount of buffer B was manually increased 10 stepwise until the protein eluted off the column. Buffer A contained 0.1% formic acid in water and buffer B contained 0.1% formic acid in acetonitrile. The mobile phase flow rate was 300 μL/min. The mass spectrometer was controlled by Sciex OS v.1.3 using the following settings: Ion source gas 1 15 psi, ion source gas 2 30 psi, curtain

gas 35, CAD gas 7, temperature 200 oC, spray voltage 5200 V, declustering potential 80 V, collision energy 15 V. Data was acquired from 1400-3600 Da with a 1 s accumulation time and 80 time bins summed. The acquired mass spectra for the proteins of interest were deconvoluted using BioPharmaView v. 2.1 software (Sciex) in order to obtain the molecular weights. The peak threshold was set to ≥ 5%, reconstruction processing was set to 20 iterations with a signal to noise threshold of ≥ 5 and a resolution of 20000.

Co-expression of SidJ, SdeA, and CaM in *E. coli*

Rosetta (DE3) cells were transformed with ppSUMO-SdeA$^{\Delta NC\ H407A}$ (Kan^{R}) and pET-Duet1 (Novagen) with full-length SidJ cloned into the MCS1 and full-length human CaM (CALM2) cloned into the MCS2, both untagged (Amp^{R}). *E. coli* were grown to OD_{600} ~0.8-1.0 under double antibiotic selection and induced with 0.4 mM IPTG overnight. SdeA was affinity purified with Ni/NTA as described above. The SUMO tag was cleaved with ULP, and the protein was purified over gel filtration. Peak fractions were combined and resolved by 6% SDS-PAGE and Coomassie blue staining to detect mobility shifts.

Glutamylation assay with [^{14}C]Glu

Glutamylation reactions were performed in a reaction mix containing 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$, 1 mM ATP, and 200 μM [U-14C] L-Glu (specific radioactivity 200 cpm/pmol). For standard endpoint reactions, a typical 30 μL reaction contained 2.5 μM SdeA$^{\Delta NC}$, 0.2 μM SidJ$^{\Delta NC}$, and 1 μM CaM

(excess CaM was added in some reactions). Reactions were initiated by adding ATP/Mg^{2+}, allowed to proceed at 30°C for 30 min, then terminated by addition of 3 µL of 500 mM EDTA followed by the addition of SDS-loading buffer with β-ME. Reaction products were resolved by SDS-PAGE and visualized by Coomassie staining. Gels were then soaked in EN3HANCE reagent (Perkin Elmer) for 30 min, followed by a 30 min soak in H_2O, then dried overnight. ^{14}C incorporation was detected by autoradiography. In assays where SdeA was glutamylated to inactivate its ART activity, glutamylation was performed in 50 μL reactions with 1 μM $SdeA^{\Delta NC}$, 1 μM $SidJ^{\Delta NC}$, 5 μM CaM, 1 mM unlabeled Glu (or other amino acids), initiated by adding 5 mM MgCl2 and 1 mM ATP. These reactions were performed for 1 h at 30°C before the mixture was cooled on ice, diluted, and used in ART or pR-Ub laddering assays.

ADP-ribosylation assay with [^{32}P]NAD+

Following glutamylation as described above, $SdeA^{\Delta NC}$ H407A or $SdeA^{\Delta NC, E860A}$ were diluted to a final concentration of 12.5 nM in reactions containing 5 mM EDTA, 10 μM HA-Ub, 100 μM [^{32}P]NAD+ (specific radioactivity 500 cpm/pmol), 50 mM Tris-HCl pH 7.5, 150 mM NaCl, and 1 mM DTT. Reactions were initiated by addition of NAD+ and incubated at 23°C for 30 min, or for the indicated timepoints. Reactions were terminated by addition of SDS loading buffer with β-ME and boiling. HA-Ub was resolved by SDS-PAGE and visualized by Coomassie staining. Gels were fixed, dried, and ^{32}P incorporation determined by autoradiography.

Phosphoribosyl-ubiquitination assay

Following glutamylation, SdeA$^{\Delta NC}$ or the indicated mutants were diluted to a final concentration of 200 nM in reactions containing 5 mM EDTA, 15.0 μM HA-Ub or ADPR-HA-Ub, 100 μM NAD+, 50 mM Tris-HCl pH 7.5, 150 mM NaCl, and 1 mM DTT. Reactions were initiated by addition of SdeA, allowed to proceed for 15 min at 23°C, then terminated by addition of SDS loading buffer with β-ME and boiling. Samples were resolved by SDS-PAGE, transferred to a nitrocellulose membrane, and immunoblotted with the indicated antibodies. Abcam anti-Ub (ab7254) was used to detect unmodified ubiquitin.

Gel filtration chromatography of SidJ-CaM complex

SidJ$^{\Delta NC}$ and human CaM were mixed with excess CaM and resolved on a Superdex 200 gel filtration column using AKTA Pure FPLC chromatography system in buffer containing 50 mM Tris HCl pH = 7.5, 300 mM NaCl, and 1 mM DTT.

Isothermal titration calorimetry

SidJ$^{\Delta NC}$ (and indicated mutants) were purified as described above. Human CaM was purified similarly; however, following cleavage of the 6xHis-SUMO tag with ULP, imidazole was removed by buffer exchange and the samples were passed through Ni-NTA resin to remove 6xHis-SUMO and ULP. The samples were then subjected to gel filtration chromatography on Superdex S200, equilibrated with Gel Filtration buffer (25 mM Tris pH 8.0, 150 mM NaCl, 0.5 mM TCEP). Fractions were collected and the proteins were concentrated in 10 kD-cutoff cellulose spin

concentrators (Amicon) thoroughly washed with Gel Filtration buffer. Human CaM at 500 µM syringe concentration was titrated into SidJ$^{\Delta NC}$ at 50 µM cell concentration, using GE Healthcare MicroCal^{TM} iTC200 System at 20 °C. First injection of 0.5 µL was followed by 20 injections of 1.9 µL. Initial small volume injection was rejected from the analysis. Spacing between injections was adjusted to allow for baseline equilibration (300 s for SidJ$^{\Delta NC}$ D542A and FFE, and 150 s for the SidJ mutants IQ, and IQ+FFE). Resulting thermograms were integrated using NITPIC software[68]. Isotherm fit to a binary interaction model was performed using SEDPHAT software. Analysis was made based on single titrations. Error analysis was conducted using automatic confidence interval search with projection method, a toolkit built into the SEDPHAT software. Results are reported as best fit with boundaries of 68.3% confidence interval. Figures were prepared using GUSSI software.

Purification of the SidJ-CaM complex for crystallography

The SidJ$^{\Delta NC}$-yCaM complex was prepared by co-expression of ppSUMO-SidJ$^{\Delta NC}$ with pET-Duet1 MCS2 CaM (untagged) in Rosetta E coli. Protein expression was induced and collected as described above. Selenomethionine-substituted protein was prepared by growing E. coli in SelenoMet minimal media (21.6 g/L SelenoMet Base, 5.1 g/L SelenoMet nutrient mix, 4mL/L SelenoMethionine solution (Molecular Dimensions). Lysis was performed by the French press method in an Avestin Emulsiflex C3 French press. The components were chilled with ice and the pellet from 5 L of bacterial culture was suspended in 100 mL lysis buffer (50 mM

Tris-HCl pH 8, 300 mM NaCl, 1 mM PMSF, 1 mM DTT) and passed 3 times through the cell with a homogenizing pressure of 15,000 psi. The lysate was cleared and purification of the complex with Ni/NTA resin performed as described above. Gel filtration was performed in low ionic strength buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl, 1mM DTT). Peak fractions were concentrated to ≥10 mg/mL in 10 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM TCEP, 1 mM $MgCl_2$, 1 mM $CaCl_2$, and 1 mM AMP-CPP.

Crystallization and structure refinement of the SidJ-CaM complex

Screening for initial conditions was performed at the UT Southwestern Structural Biology Core using a Gryphon robot dispensing 0.2 μL protein (10 mg/mL) and 0.2 μL reservoir solution in INTELLIPLATE 96-3 plates (Art Robbins Instruments). Needle-like clusters of protein crystals appeared after 72 hr. incubation at 20°C in several conditions containing PEG and organic acid salts. Optimization of initial conditions was performed in 24-well Cryschem M trays. Ultimately, native crystals were grown by the sitting drop vapor diffusion method at 20°C in 24-well Cryschem M sitting drop trays (Hampton) using a 1:1 ratio of protein (8 mg/mL) reservoir solution containing 0.2M Na Malonate pH 8.0 (Hampton HR2-807) and 14-15% w/v PEG 3350. Wells were allowed to equilibrate for 24-48 h and crystal growth was initiated by micro-seeding. Seeds were prepared by harvesting poorly-formed crystals from earlier attempts and beating in reservoir solution with a Teflon bead. Serial dilutions of the seed stock were used to obtain the ideal seeding conditions.

Once single crystals formed (After 1-2 weeks), crystals were harvested and cryo-protected with 0.2 M Na Malonate pH 8.0, 10 mM NaCl, 16-17% (w/v) PEG 3350, 35-40% (w/v) ethylene glycol, 1 mM $MgCl_2$, 1 mM $CaCl_2$, and 1 mM AMP-CPP in robotic loops. The loops were plunged into liquid nitrogen and stored for data collection.

Native crystals diffracted to a minimum Bragg spacing (d_{min}) of 2.10 Å and exhibited the symmetry of space group $P2_1$ with cell dimensions of a = 105.5 Å, b = 104.8 Å, c = 107.8 Å, b = 104.0° and contained two complexes per asymmetric unit. Crystals of selenomethionyl-derivatized SidJ-yCaM were grown under similar conditions and crystallized in the same space group and similar lattice constants. All diffraction data were collected at beamline 19-ID (SBC-CAT) at the Advanced Photon Source (Argonne National Laboratory, Argonne, Illinois, USA) and processed in the program *HKL-3000* [69] with applied corrections for effects resulting from absorption in a crystal and for radiation damage [70,71], the calculation of an optimal error model, and corrections to compensate the phasing signal for a radiation-induced increase of non-isomorphism within the crystal [72,73]. These corrections were crucial for successful phasing. Phases were obtained from a single wavelength anomalous dispersion (SAD) experiment using the selenomethionyl-derivatized SidJ-yCaM complex with data to 2.50 Å. Twenty-nine selenium sites were located in the program *SHELXD*[74], phases calculated in the program *MLPHARE*[75], improved via density modification and 2-fold averaging in the program *PARROT* [76] and an initial model containing 80% of all

SidJ-yCaM residues were automatically generated in the program *Buccaneer* [77], primarily these were SidJ residues. Completion of this model was performed by manual rebuilding in the program *Coot* [78]. Modeling of the N- and C-terminal EF hand motifs of yCaM was aided by manual placement of appropriately edited human CaM coordinates from PDB ID 3UCW and 1XFY, respectively. Positional and isotropic atomic displacement parameter (ADP) as well as TLS ADP refinement was performed to a resolution of 2.10 Å for the Mg^{2+}-Ca^{2+}-PP_i-AMP-bound native SidJ-yCaM using the program *Phenix* with a random 1.8% of all data set aside for an R_{free} calculation. Electron density for the yCaM polypeptides is weaker than for the SidJ polypeptide, due to a relative dearth of lattice contacts for the yCaMs. This is reflected in the higher ADPs for the yCaM polypeptides. Data collection and structure refinement statistics are summarized in **Table 4.**

Glutamylation initial rate assays for SidJ mutants

Triplicate reactions were conducted at room temperature for 15 min in a reaction mix containing 100 mM Sodium acetate, 50 mM Bis-Tris, 50 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$, 0.3 mM ATP, and 200 μM L-[3,4-^{3}H]Glu (specific radioactivity 1500 cpm/pmol). 20 μL reactions contained 0.11 μM $SidJ^{\Delta NC}$ and 2.4 μM $SdeA^{\Delta NC}$. After termination with 7 μL of stop mix (0.14 M EDTA pH 8.0 in x5 SDS-PAGE Loading Dye), reactions were boiled for 5 min and 10 μL samples were run on 8% or 12% SDS-PAGE gels. Coomassie stained SdeA bands were excised, placed in glass scintillation vials and digested in 0.5 mL of 30% H_2O_2

at 60 - 65 °C overnight. After digestion was complete, vials were cooled down to ambient temperature, mixed with 12 mL of Budget-Solve scintillation cocktail (RPI, 111167) and ^{3}H quantified by scintillation counting.

Glutamylation assays with nucleotide analogs

30 μL reactions containing 2.5 μM SdeA$^{\Delta NC}$, 0.2 μM SidJ$^{\Delta NC}$, 1 μM CaM, 200 μM L-[3,4-^{3}H]Glu (specific radioactivity 1000 cpm/pmol) were initiated with 2 mM $MgCl_2$ and 500 μM ATP, ADP, AMP-PNP or AMP-CPP. Reactions were incubated at 37°C for 20 min, then terminated by addition of 3 µL of 500 mM EDTA followed by the addition of SDS-loading buffer with β-ME. Reaction products were resolved by SDS-PAGE on 6% polyacrylamide gels and visualized by Coomassie staining. Coomassie stained SdeA bands were excised, placed in glass scintillation vials and digested in 0.5 mL of 30% H_2O_2 at 60 - 65°C overnight. After digestion was complete, vials were cooled down to ambient temperature, mixed with 12 mL of Budget-Solve scintillation cocktail (RPI, 111167) and ^{3}H quantified by scintillation counting.

Detection of SidJ reaction products by LCMS/MS

Glutamylation reactions were performed as described above with the following modifications. Reactions (100 µL) containing 3 µM SdeA$^{\Delta NC}$, 0.15 µM SidJ$^{\Delta NC}$, 100 µM ATP, 250 µM $MgCl_2$, and 1 mM Glu were allowed to proceed for 10 min at room temperature and terminated with EDTA. Control reactions (for background detection) contained all these components, but the protein was boiled before addition of the ATP/

$MgCl_2$ mix. Levels of adenine nucleotides were analyzed using a Prominence HPLC system (Shimadzu Scientific Instruments, Columbia, MD), coupled to a 4000 QTRAP® mass spectrometer (AB Sciex, Framingham, MA) with a turbo-ion spray source. The mobile phase consisted of 30% of acetonitrile and 10 mM NH_4 acetate in water for solvent A (pH 6), and 30% of acetonitrile and 1 mM NH_4 acetate in water for solvent B (pH 10.5). A volume of 40 µL was injected at a flow rate of 0.5 mL/min into a Thermo Scientific BioBasic AX, 2.1 × 50 mm, 5 µm column (Fisher) with guard column. Samples were separated with the following gradient: 0–1 min, 0% solvent B; 1–2.5 min, increase to 35% solvent B; 2.5–5 min, 35% solvent B; 5–7 min, increase to 65% solvent B; 7–10 min, 65% solvent B; 10–10.5 min, increase to 100% solvent B; 10.5–15 min, 100% solvent B; 15–15.5 min, decrease to 0% solvent B; 15.5–20.5 min, 0% solvent B. Ions formed by turbo ion spray in positive mode were used to detect adenine nucleotides and the internal standard (UMP-$^{13}C_9$,$^{15}N_2$; "UMP-IS"). MS/MS detection was used to monitor the fragmentation of AMP from 347.99 to 136.0, of ADP from 427.99 to 136.0, of ATP from 507.95 to136.0, and of UMP-IS from 336.09 to 102.0.

Determination of adenine nucleotide amounts

Reactions were processed for analysis of adenine nucleotide levels by addition of 150 µl of 2.085% perchloric acid to the 100 µl glutamylation reaction mixture. Samples were vortexed for 15 s, incubated for 10 min at RT and then centrifuged for 5 min at 16,100 xg in a refrigerated microfuge. The supernatant was recovered and

centrifuged a second time. This material was diluted 1:40 in solvent A containing 100 ng/ml UMP-IS and analyzed as described above. Relative levels of ADP and ATP were determined by normalization of analyte peak areas in each sample by peak areas of the internal standard UMP-IS. AMP levels were quantitated using a six-point standard curve prepared from AMP diluted in various quantities into the enzyme reaction mixture containing ATP and processed as for samples. Back-calculation of all points on the standard curve were within 15% of theoretical.

Detection of the acyl-adenylate intermediate

Acyl-adenylate formation was followed using TCA precipitation[79]. reactions were carried out in the presence of 150 µM [α^{32}P-ATP]; specific radioactivity = 1000 cpm/pmol. The 20 µL reaction contained 1 mg/mL bovine serum albumin (BSA), 100 mM Sodium acetate, 50 mM Tris, 50 mM Bis-Tris pH 6.5, 150 mM NaCl, 0.5 mM $MgCl_2$, 1 mM DTT, 9 µM human CaM, ~ 5 µM of SidJ$^{\Delta NC}$ and SdeA$^{\Delta NC}$. Reactions were incubated at room temperature for 1 h. Sodium glutamate dissolved in 100 mM sodium acetate, 50 mM Tris, 50 mM Bis-Tris pH 6.5, 150 mM NaCl was added before the reaction, or after 30 min as indicated. Reactions were stopped with 0.5 mL of ice-cold 50 mM ATP in 20% TCA, and incubated on ice for 40 min. Reaction products were pelleted by centrifugation at 4°C for 15 min at ~21,000 xg. The pellet was washed twice with 250 µL of ice-cold 20% TCA and the radioactivity in the TCA insoluble material was quantified by scintillation counting.

CHAPTER FOUR
Results

Identification of a putative kinase fold in SidJ

The T4SS effector repertoire from the genus *Legionella* was analyzed for divergent members of the protein kinase superfamily. Using the sequence profile method Fold and Function Assignment System (FFAS) [57], sequences of known and predicted effector proteins were compared against libraries of protein kinase profiles. SidJ (lpg2155), an 873 aa protein, was found to bear weak sequence similarity to protein kinases. The genome of *Legionella pneumophila* encodes two SidJ paralogues, SidJ and SdjA (lpg2508), but only SidJ is known to be required for full virulence[41]. The similarity of SidJ to known kinases was of borderline statistical significance and the sequence identity did not exceed 14 %. The region of similarity to kinases included only parts of the N-lobe, where the ion pair K72 (PKA nomenclature, SidJ; K367) and E91 (PKA nomenclature, SidJ; E381) could be identified in a relatively straightforward manner. A refined FFAS search extended the kinase-like region of SidJ and the kinase domain could be localized approximately from residues 336 to 593 of SidJ (**Figure 5, upper panel**). However, the position of the catalytic loop and metal-binding residues, as well as much of the C-lobe, remained ambiguous. Nevertheless, some alignments (e.g. to bacterial aminoglycoside kinases) suggest the position of metal-binding residues; a ^{542}DLG motif could be the counterpart of PKA ^{184}DFG, and SidJ N534 may correspond to

N171 of PKA (**Figure 5, lower panel**). Notably, the putative metal-binding D542 lies within a ^{542}DXXD motif, previously shown to be required for successful *Legionella* replication in amoeba [43] and mammalian cells [56].

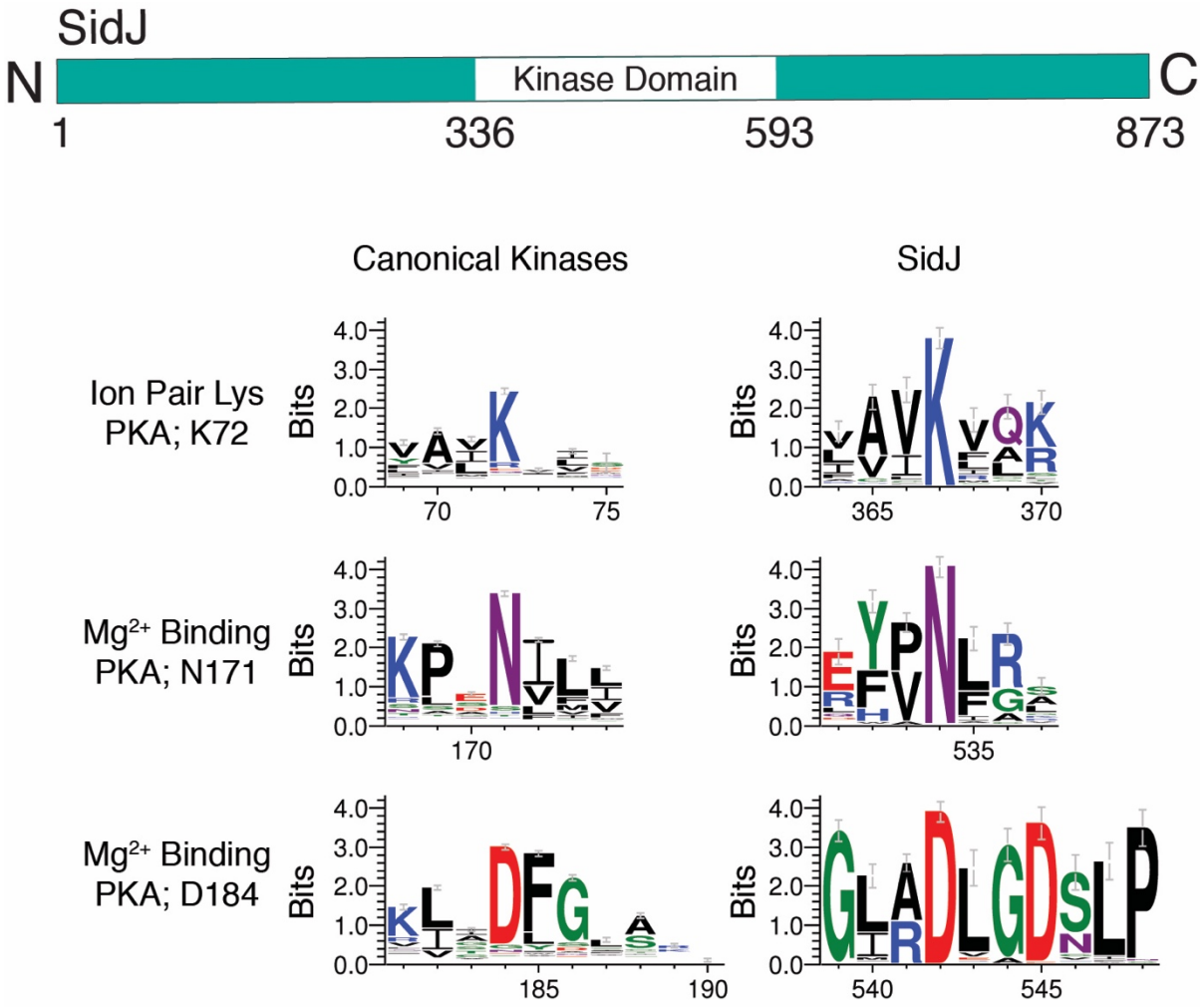


Figure 5. SidJ encodes a putative protein kinase fold. Domain architecture of *L. pneumophila* SidJ (upper panel) depicting the location of the predicted kinase domain. (Lower panel) Sequence logos (weblogos) highlighting conserved kinase active site residues in 106 SidJ homologues and 3,998 homologues of typical protein kinases (Pfam domain PF00069). The height of the amino acid stack is proportional to the sequence conservation at that position.

Suppression of SdeA and SdeC-mediated toxicity by SidJ in yeast

The yeast *Saccharomyces cerevisiae* was chosen as a proxy to study the activity of SidE and SidJ. Expression of the SidE proteins in *Saccharomyces cerevisiae* results in a growth inhibition phenotype that can be suppressed by co-expression of SidJ [43,44]. N-terminally FLAG-tagged SidJ (FLAG-SidJ) and C-terminally GFP tagged SdeA (SdeA-GFP, as a representative of the SidE family) were introduced into yeast under the control of galactose-inducible promoters. To determine if the putative kinase activity of SidJ is required for this suppression, mutations were introduced in the SidJ predicted ion pair Lys (K367A) and the metal-binding Asp (D542A). Expression of SdeA-GFP caused a strong growth inhibition phenotype that was partially suppressed by co-expression of FLAG-SidJ. Ala substitutions of K367A and D542A in SidJ failed to suppress SdeA toxicity **(Figure 6, upper panel).** The SidE effectors are thought to be functionally redundant. Accordingly, FLAG-SidJ, but not the K367A and D542A mutants, also suppressed the toxicity of SdeC-GFP (**Figure 6, lower panel).**

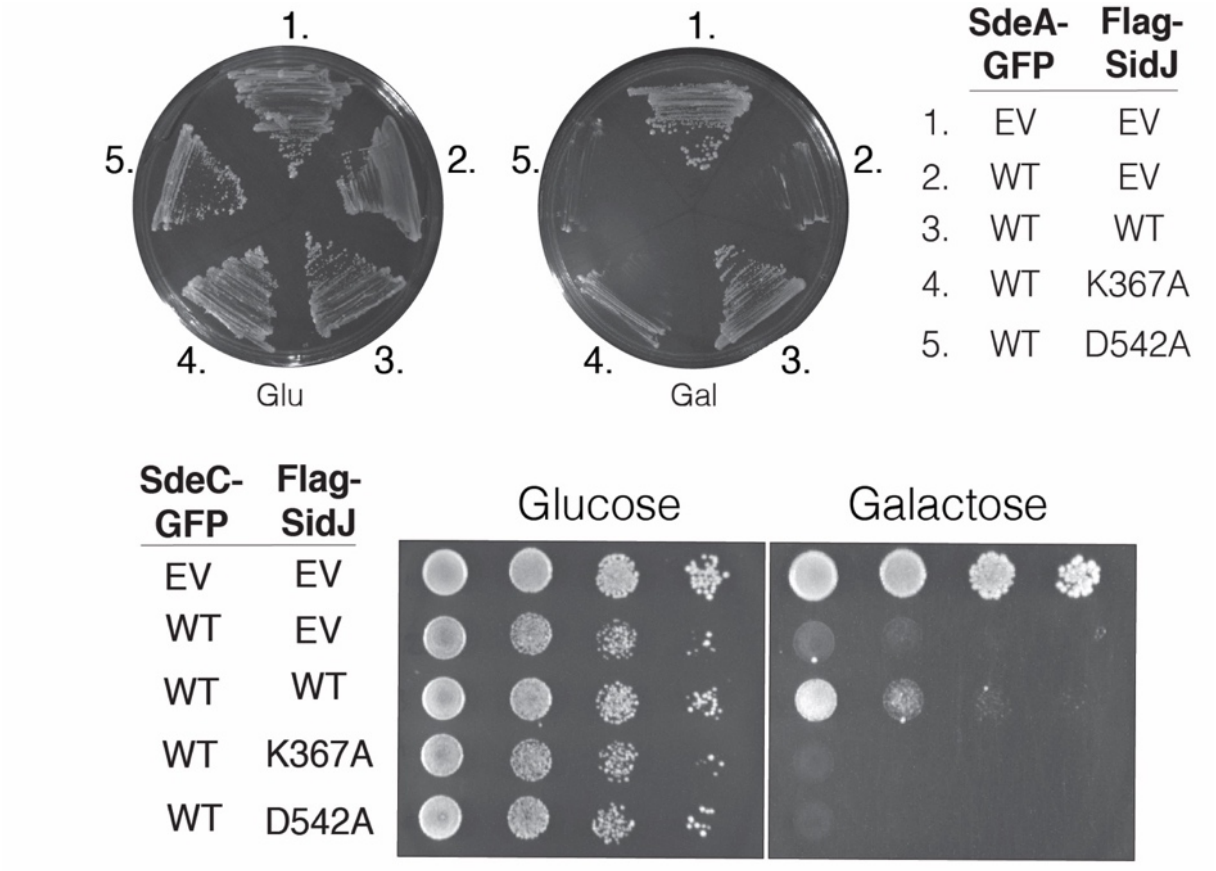


Figure 6. Suppression of SdeA- and SdeC-mediated toxicity by SidJ in yeast. Streaking assay depicting the growth of *S. cerevisiae* expressing SdeA-GFP and Flag-SidJ, or the predicted catalytically inactive K367A and D542A mutants. EV; empty vector. (Lower), Spot-dilution assay depicting growth of *S. cerevisiae* expressing SdeC-GFP and Flag-SidJ, or the K367A and D542A mutants.

Suppression of SidE-mediated phosphoribosyl ubiquitination by SidJ requires predicted kinase residues

Host cell E1 enzymes activate Ub for transfer by adenylating its C-terminal glycine residue [53]; however, this residue is not required for SidE-catalyzed pR-ubiquitination [29]. We mutated the carboxy terminal GG motif of Ub to AA ($Ub^{GG/AA}$) for transient expression in mammalian HEK293A cells to interrogate SidE-catalyzed ubiquitination. Expression of hemagglutinin (HA)-tagged WT Ub but not $Ub^{GG/AA}$ resulted in the modification of several host proteins by the endogenous ubiquitination machinery **(Figure 7, lanes 1 and 2).** To express detectable amounts of SidJ and SdeA in mammalian cells, the open reading frames were codon optimized for mammalian expression and driven by a CMV promoter. When HA-$Ub^{GG/AA}$ was co-expressed with Myc-tagged SdeA, but not the E860A (ART domain) or the H407A (PDE domain) mutants, we detected abundant pR-ubiquitination of host proteins **(Figure 7, lanes 3-5).** Co-expression of SidJ-V5, but not the K367A or D542A mutants, markedly inhibited SdeA-catalyzed ubiquitination **(Figure 7, lanes 6-8).** To test the importance of kinase residues in an infection, strains of *Legionella pneumophila* were generated with a deletion of the entire SidJ open reading frame (**Figure 8**). HEK293A cells were transfected to express HA-$Ub^{GG/AA}$ with FcγRIIa and challenged with antibody-opsonized *Legionella.* This strategy allows the bacteria to deliver effectors into HEK293 cells, and has been used as a model system to

detect *Legionella* effector-catalyzed reactions[26]. Deletion of SidJ in *Legionella* led to an accumulation of pR-Ubiquitination in infected mammalian cells, which was ameliorated by complementation with WT SidJ but not the K367A mutant (**Figure 9**).

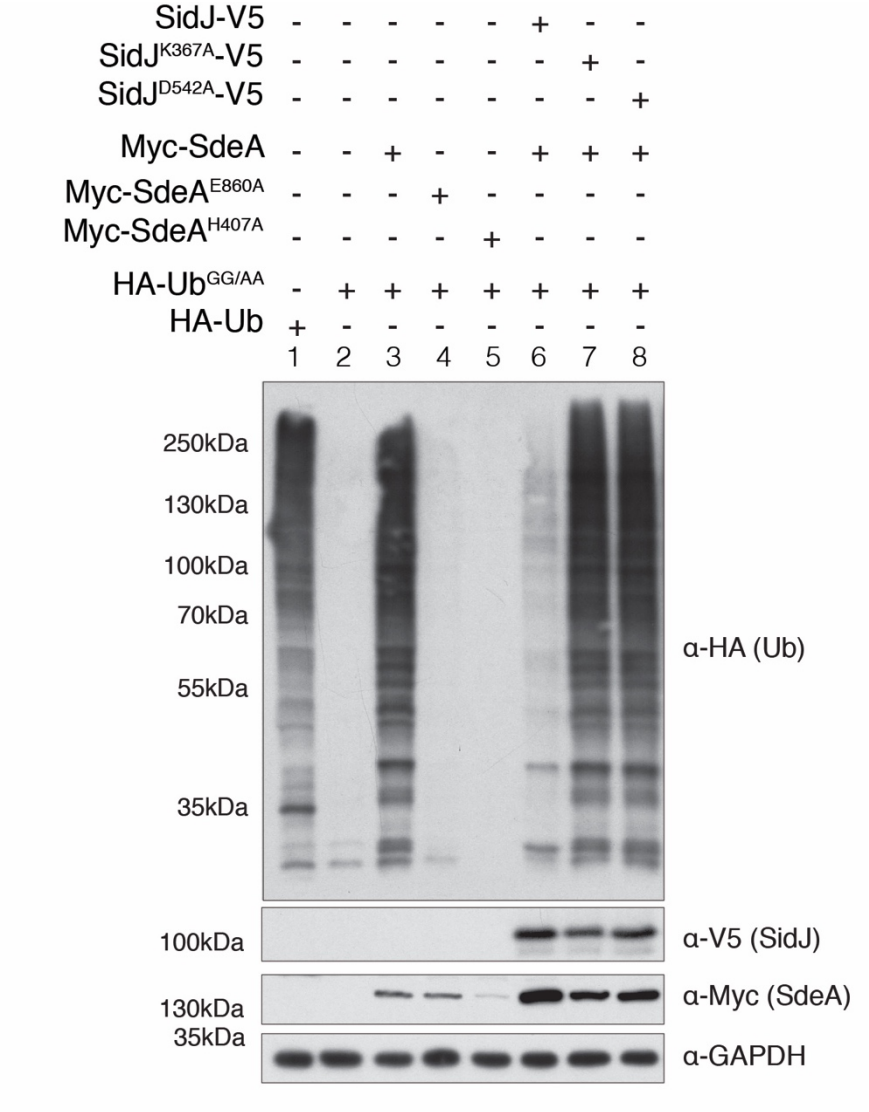


Figure 7. Suppression of SidE-mediated phosphoribosyl ubiquitination by SidJ in mammalian cells. Protein immunoblotting of total extracts from HEK293A cells expressing HA-Ub, HA-Ub$^{GG/AA}$, codon-optimized Myc-SdeA and codon-optimized SidJ-V5, or the indicated mutants. GAPDH is shown as a loading control.

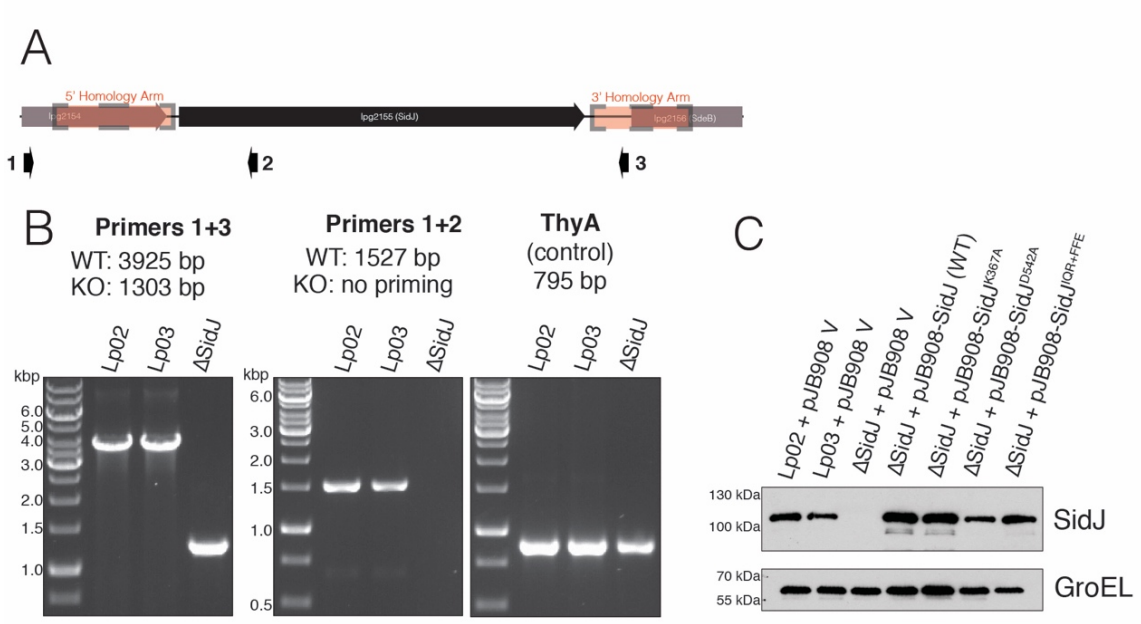


Figure 8. Gene deletion of SidJ in *legionella*. (A) Schematic representation of the *sidJ* locus in *Legionella pneumophila* highlighting the regions of homology used to generate the Δ*sidJ* mutant. The location of the primer annealing sites used to screen for mutants are also shown. (B) Agarose gels depicting PCR products obtained with the indicated primers from A using *Legionella* genomic DNA from the Lp02, Lp03 and the Δ*sidJ* mutant. (C) Protein immunoblotting of *Legionella* extracts from the Lp02, Lp03 and the Δ*sidJ* mutant (or complementing strains) depicting SidJ protein levels. GroEL is shown as a loading control.

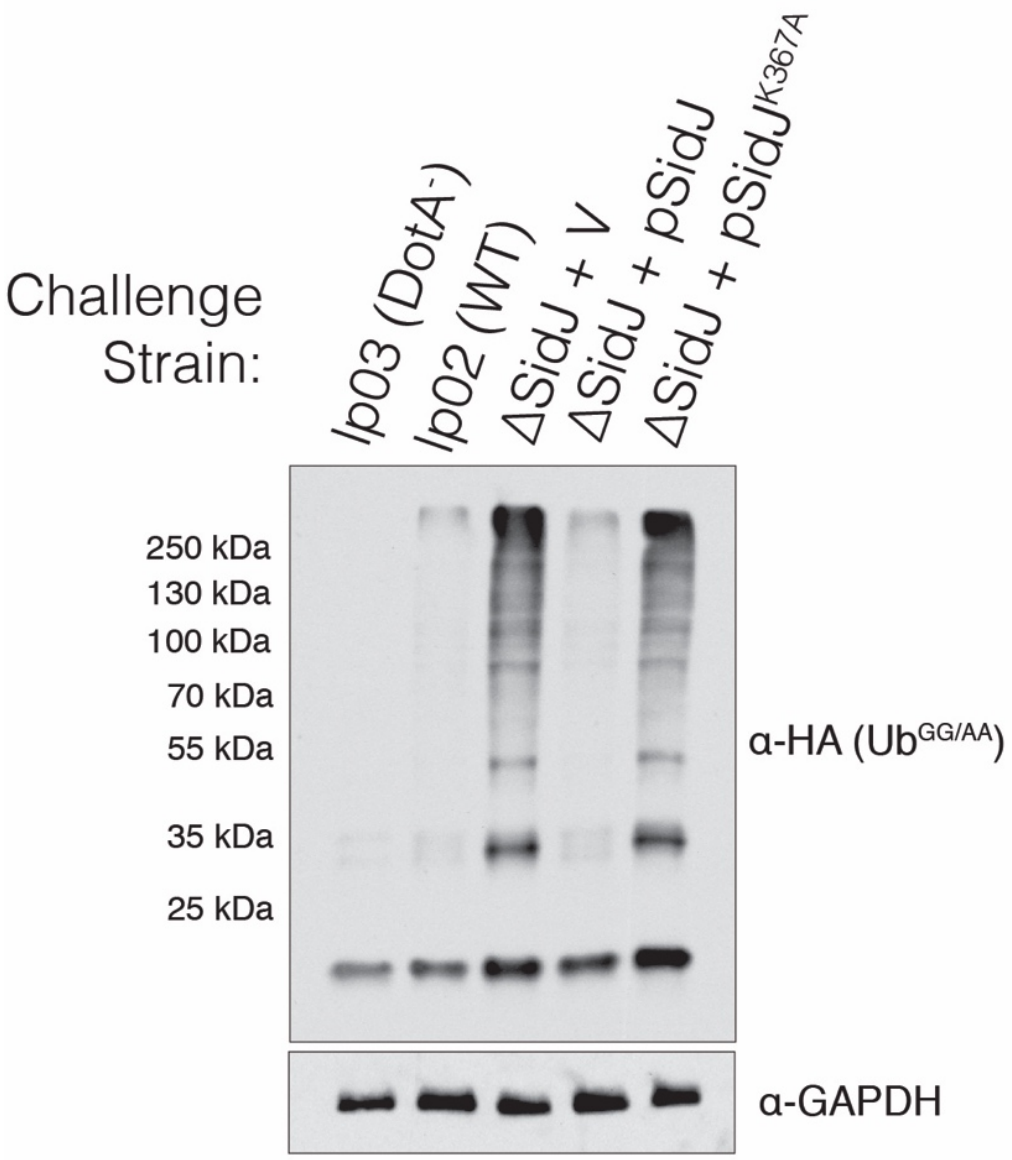


Figure 9. Suppression of SidE-mediated ubiquitination by SidJ during *Legionella* infection. Protein immunoblotting depicting HA-$UB^{GG/AA}$ levels in HEK293A cell extracts following challenge with WT Legionella (lp02), the T4SS mutant (lp03) or SidJ mutants. HEK293A cells were co-transfected with HA-tagged $Ub^{GG/AA}$ and FcγRIIa. The cells were subsequently challenged with antibody opsonized *Legionella*, lysed 1 h post infection, and HA-tagged $Ub^{GG/AA}$ was analyzed by protein immunoblotting. GAPDH is shown as a loading control.

SidJ binds host CaM

Previous studies had shown that overexpression of SidJ alone causes growth inhibition in yeast, raising the possibility that SidJ may have a eukaryotic substrate[43,44]. To identify possible substrates and host factors that interact with SidJ, Flag-SidJ was immunopurified from yeast and anti-Flag immunoprecipitates were analyzed by LC-MS/MS. A strong signal was identified for the eukaryotic calcium sensor CaM (CaM) in SidJ and SidJ^{K367A}, but not in control immunoprecipitates (**Table 1**). A likely "IQ" CaM-binding motif was identified by FFAS at the C-terminus of SidJ (residues 832-853), **(Figure 10).** Mutations to disrupt CaM binding were predicted by substituting conserved I841, Q842, R843 and R846 with acidic residues (SidJ^{IQ}). SidJ-V5 immunoprecipitated endogenous CaM when expressed in mammalian cells, but the SidJ^{IQ} mutant did not **(Figure 11).** When FLAG-SidJ was overexpressed alone in yeast, it caused a mild growth inhibition phenotype, consistent with earlier reports. However, yeast growth could not be rescued by mutations in the predicted kinase catalytic residues of SidJ. Instead, growth was rescued when the SidJ IQ motif was disrupted **(Figure 12),** revealing that SidJ likely intoxicates yeast by binding and sequestering cellular CaM. Co-expression of SidJ-V5^{IQ} and myc-SdeA in mammalian cells with HA-Ub$^{GG/AA}$ revealed that CaM binding is also necessary for SidJ to suppress pR-ubiquitination **(Figure 13).**

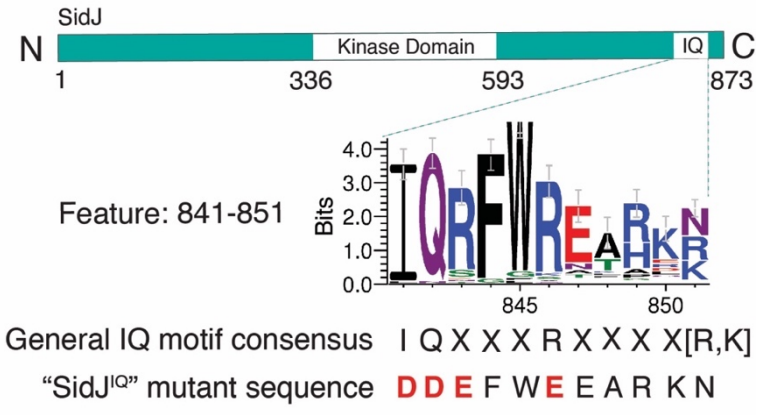


Figure 10. The SidJ IQ motif. Schematic representation of the SidJ protein highlighting the position and conservation (Weblogo) of the predicted IQ CaM binding motif in SidJ homologs. The canonical IQ consensus and SidJ^{IQ} mutant sequences are also shown.

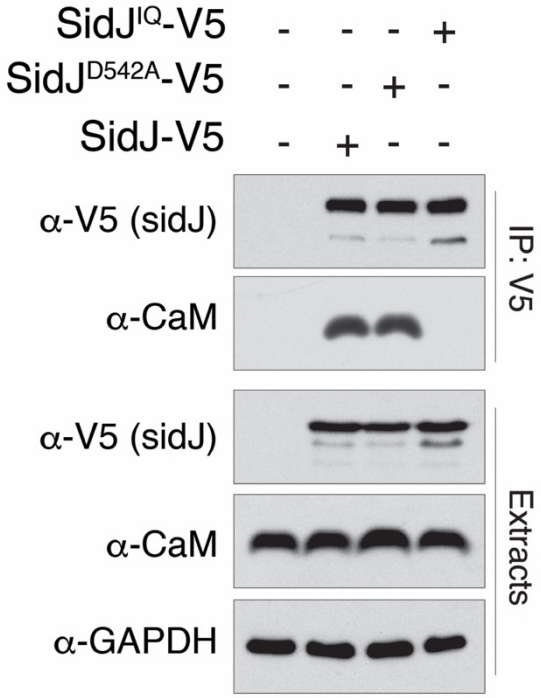


Figure 11. Co-immunoprecipitation of SidJ and CaM. Protein immunoblotting of V5-immunoprecipitates and cell extracts from HEK293A cells expressing codon-optimized, V5 tagged SidJ, SidJ^{D542A}, or the SidJ^{IQ} mutant. Immunoprecipitates and cell extracts were analyzed for CaM and SidJ. GAPDH is shown as a loading control.

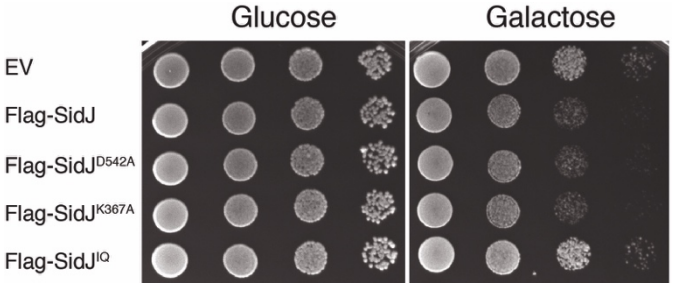


Figure 12. CaM binding accounts for SidJ toxicity in yeast. Spotting assay depicting the growth of *S. cerevisiae* expressing Flag-SidJ, the D542A, K367A, and CaM-binding IQ mutant (I841D, Q842D, R843E, R846E). EV; empty vector.

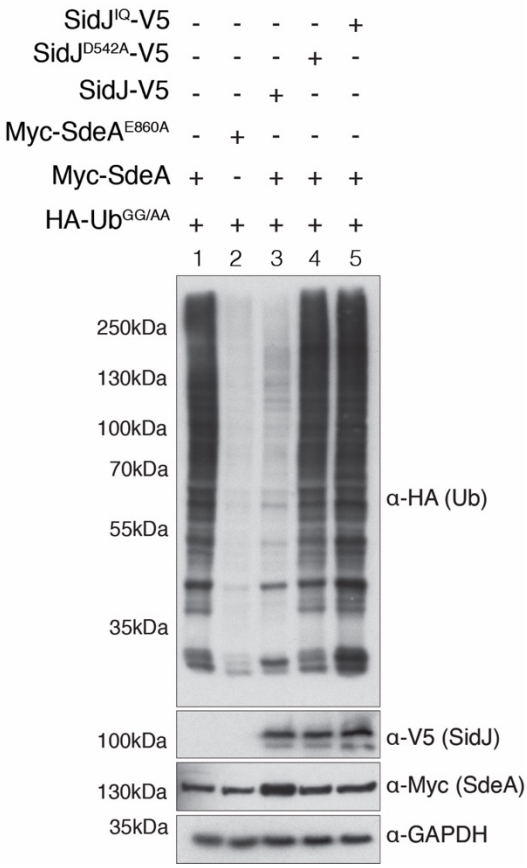


Figure 13. SidJ-CaM binding is required to suppress SidE-mediated phosphoribosyl ubiquitination. Protein immunoblotting of HEK293A total cell extracts expressing HA-tagged $Ub^{GG/AA}$, codon-optimized Myc-SdeA, V5 tagged SidJ, $SidJ^{D542A}$, or the $SidJ^{IQ}$ mutant. GAPDH is shown as a loading control.

SidJ auto-AMPylates

In initial experiments to determine if SidJ possessed kinase activity, recombinant SidJ residues 334-873 ($SidJ^{334\text{-}C}$) was incubated with CaM, high specific activity of [$\boldsymbol{\gamma}$-^{32}P]ATP, metals, and generic kinase substrates or the SidE effectors. However, phosphotransfer was never observed. Prompted by the recent discovery that the SelO family of 'pseudokinases' transfer AMP from ATP instead of the γ-phosphate, separate reactions were performed with [α-^{32}P]ATP and the bacterial SelO homolog YdiU as a positive control. SidJ displayed auto-AMPylation activity that required the catalytic residue K367 **(Figure 14).** In separate experiments, SidJ also required CaM to auto-AMPylate. However, the stoichiometry of auto-AMPylation was calculated at less than 5%, and LC-MS/MS analysis was unable to identify a single AMP acceptor residue. Furthermore, AMPylation of the SidE effectors was never observed (unpublished data).

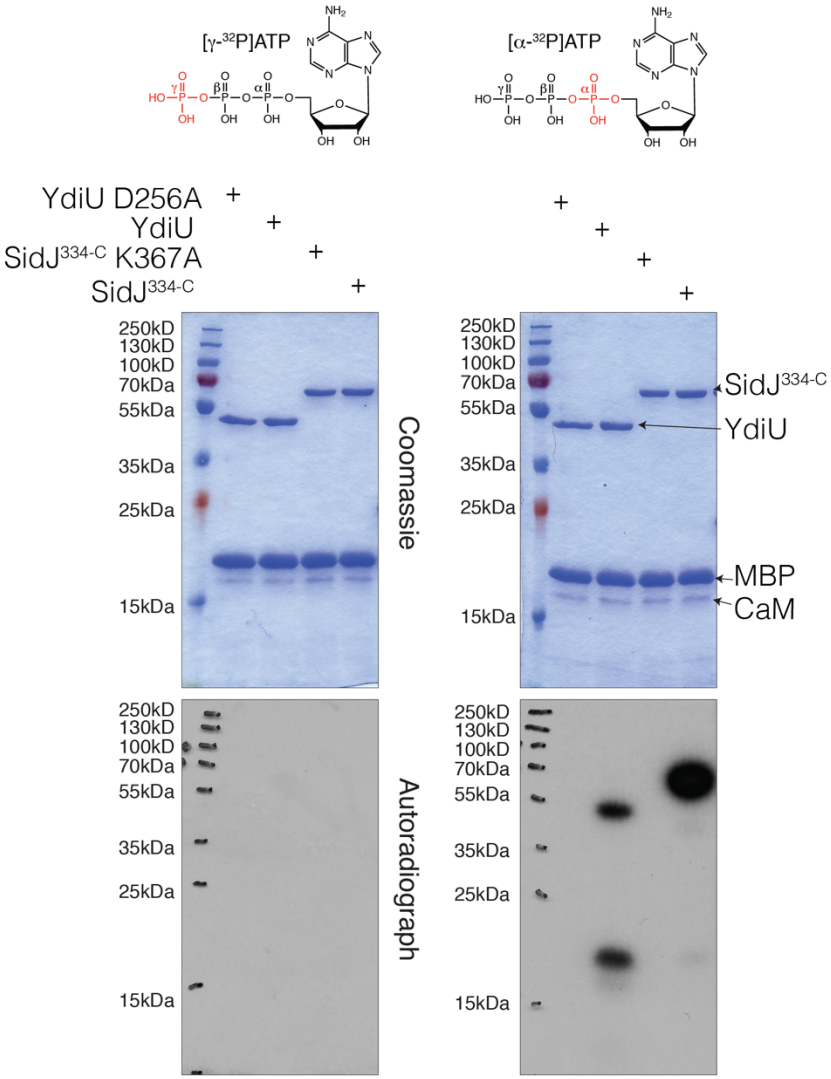


Figure 14. Phosphorylation and AMPylation activity of recombinant SidJ. Autoradiographs depicting the incorporation of γ -^{32}P from [γ-^{32}P]ATP (left) or α-^{32}P from [α-^{32}P]ATP (right) by SidJ$^{334\text{-}C}$ or the K367A mutant in the presence of CaM (CaM). *E coli* YdiU or the inactive D256A mutant were included as controls, and myelin basic protein (MBP) was included as a generic kinase substrate. Reaction products were resolved by SDS-PAGE and visualized by Coomassie blue staining (upper) and autoradiography (lower).

SidJ "Calmodulinates" SdeA

In the course of performing AMPylation experiments, a high-molecular-weight (~130 kDa) species appeared on SDS-PAGE when SidJ, CaM, ATP/Mg^{2+}, and SdeA were incubated together. Immunoblotting confirmed that the new species consisted of CaM covalently bound to SdeA, a modification dubbed 'Calmodulination.' Like AMPylation, Calmodulination required the SidJ catalytic residue K367. Interestingly, Calmodulination was inhibited by the excess pyrophosphate (PP_i), but not inorganic phosphate (P_i) **(Figure 15).** This suggest SidJ hydrolyzes the α-β phosphodiester bond of ATP, like SelO, using this free energy to catalyze a peptide ligation reaction. However, despite overnight incubations, only about half of the SdeA in the reaction became calmodulinated, and calmodulination was never detected in cells (unpublished), suggesting that calmodulination may be an artifact.

SidJ catalyzes back-exchange of ATP and PP_i

Calmodulination resembled peptide-tagging modifications such as ubiquitination, ISGylation, and PUPylation. Enzymes that catalyze these reactions, such as the E1 Ubligase, use ATP form an acyl-adenylate intermediate, releasing PP_i[67,79]. Using [^{32}P]PP_i and isolating ATP with a charcoal extraction, flux through this intermediate can be measured. SidJ back-exchanged PP_i to ATP in the presence of CaM, while the D542A mutant did not **(Figure 16).** SidJ exchanged PP_i after treatment with the thiol-alkylating agent N-ethylmaleimide (NEM), suggesting it does not form a thioester. This result provided evidence that SidJ may use the free energy

from hydrolysis of ATP to perform a peptide ligation reaction, likely involving the SidE effectors and an unknown protein or small molecule.

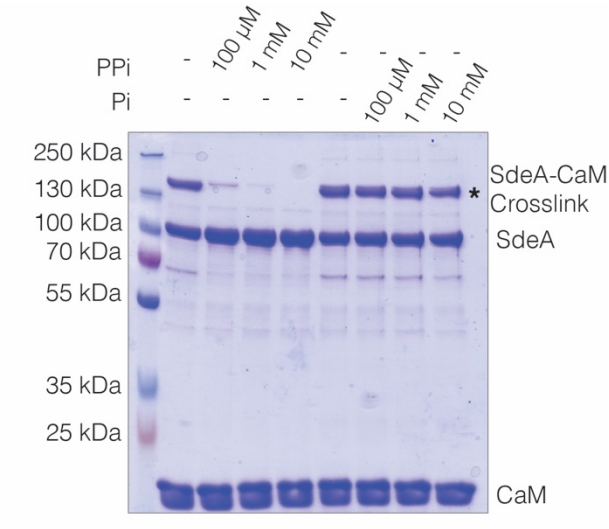


Figure 15. SidJ catalyzes Calmodulination of SdeA. SDS-PAGE demonstrating the SdeA-CaM crosslink catalyzed by SidJ and the effect of excess PP_i or P_i.

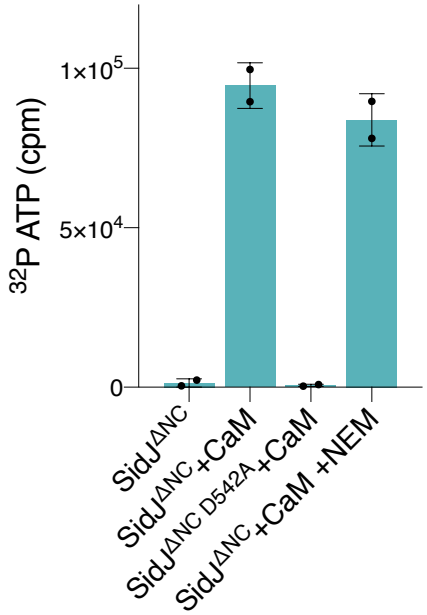


Figure 16. SidJ catalyzes ATP-PP_i exchange. Quantification of [^{32}P]PP_i incorporated into ATP by SidJ$^{\Delta NC}$ or SidJ$^{\Delta NC\ D542A}$ in reactions with CaM or N-ethylmaleimide (NEM) as indicated.

SidJ is a CaM-dependent protein polyglutamylase

To determine if SidJ adopts a kinase fold, a stable truncation of SidJ 59-851 ($SidJ^{\Delta NC}$), was identified which was suitable for purification in *E. coli* and suppressed SdeA-catalyzed ubiquitination in cells **(Figure 17).** Native protein crystals with CaM were obtained (described later), but when SelenoMet derivatized crystals were prepared in an attempt to gain phasing information, a failure to observe anomalous dispersion prompted a careful look at the protein sample. Intact mass analysis of the crystallization sample confirmed failure to incorporate SelenoMet, likely due to an expired reagent. However, intact mass also revealed an unexpected series of mass increases of 129 Da on SidJ **(Figure 18).** Referencing an online database of mass tags (abrf.org/deltamass), 129 Da was noted to correspond well to the predicted mass increase imparted by glutamylation, a post-translational modification of the γ-carboxyl of a glutamate side chain modified with the amino group of free glutamate forming an isopeptide bond **(Figure 18, inset.)**

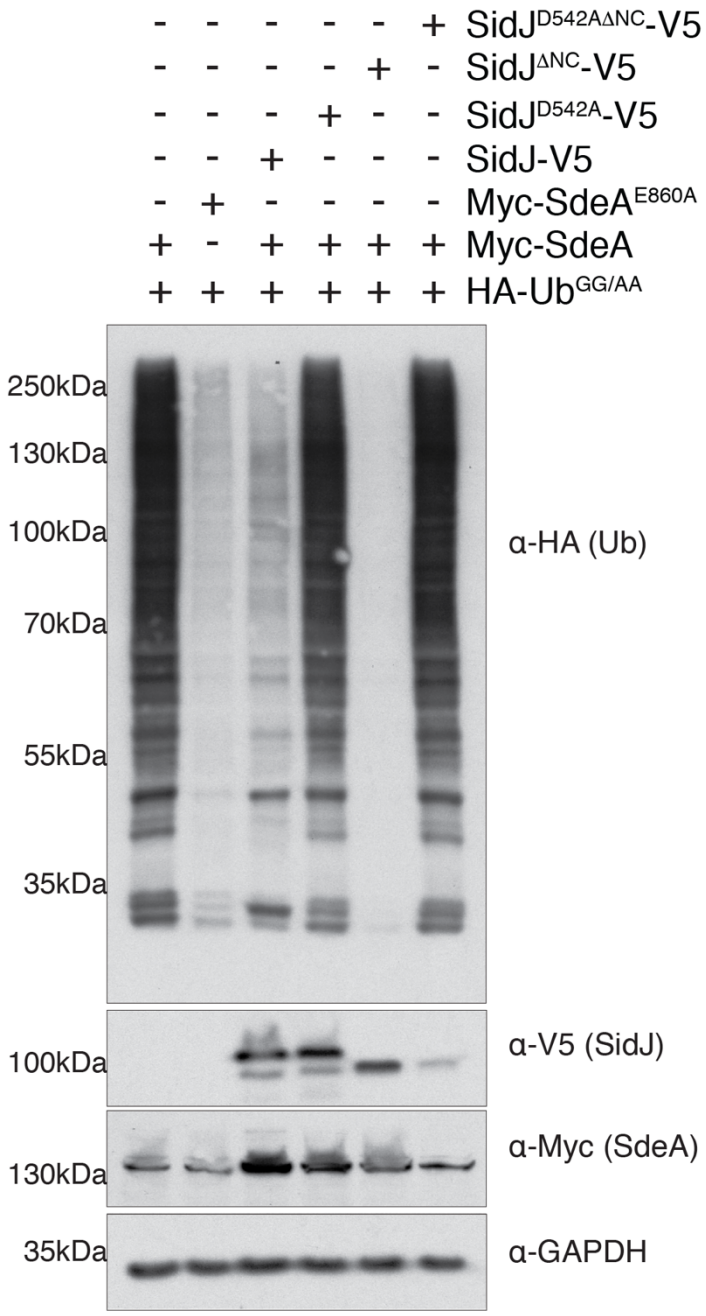


Figure 17. Truncation of SidJ from residues 59-851. Protein immunoblotting of total extracts from HEK293A cells expressing HA-Ub$^{GG/AA}$, codon-optimized Myc-SdeA, codon-optimized SidJ-V5 and codon-optimized SidJ$^{\Delta NC}$, or the indicated mutants. GAPDH is shown as a loading control. Total cell extracts were separated by SDS-PAGE and proteins were detected by immunoblotting using the indicated antibodies.

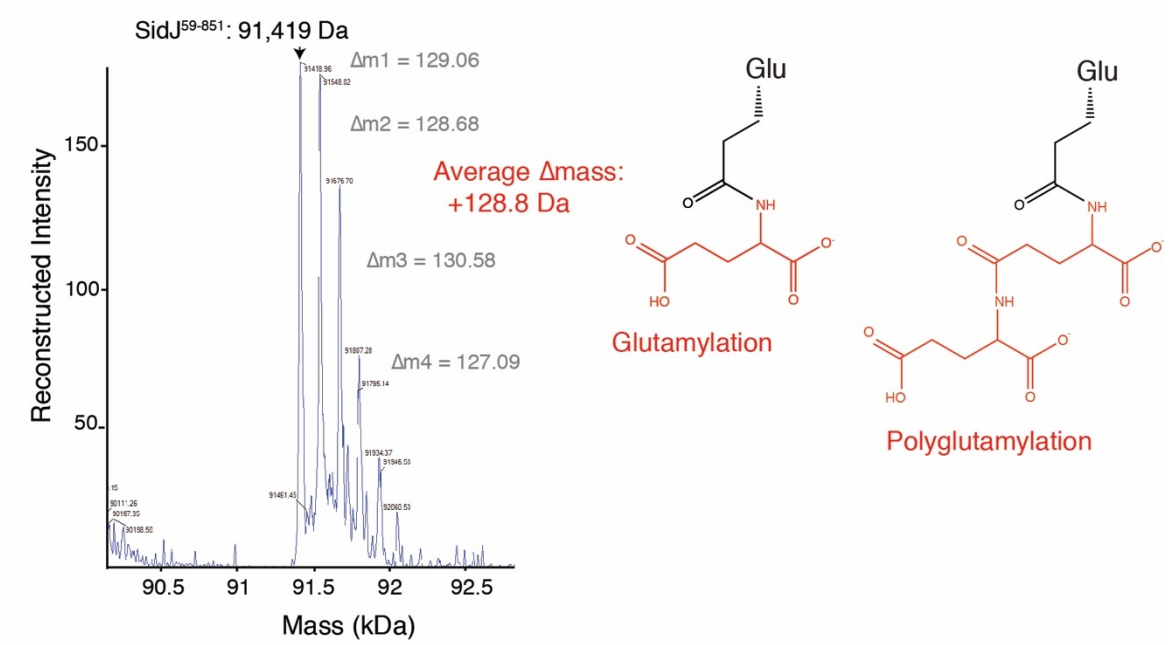


Figure 18. Intact mass analysis of SidJ. LC/MS spectra of SidJΔNC used for crystallization trials (left). Chemical structures depicting glutamylation and polyglutamylation are depicted (right).

Following co-expression of SdeA 178-1100 ($SdeA^{\Delta NC}$, containing PDE and ART domains) with full-length untagged SidJ and CaM in *E. coli,* SdeA was purified and analyzed by SDS PAGE. An electrophoretic mobility shift of SdeA appeared that formed only when WT SidJ and CaM were co-expressed **(Figure. 19, left panel).** Intact mass analysis by ESI-MS revealed a series of mass increases in increments of ~128.99 Da **(Figure. 19, right panel).** LC-MS/MS analysis identified peptides of SdeA modified by one or two glutamates (**Figure. 20**). Using [U-^{14}C] Glu, we reconstituted glutamylation *in vitro.* The reaction required CaM, ATP/Mg^{2+} and the kinase residues in SidJ (**Fig. 21A**). Glutamylation of $SdeA^{\Delta NC}$ was time-dependent (**Figure. 21B**) and while both yeast and human CaM activated SidJ to glutamylate SdeA, a related protein, Troponin C, did not (**Figure. 21C**). The ART domain of SdeA (residues 519-1100) was not glutamylated by SidJ (**Figure 22A**). This is consistent with previous results showing SdeA 551-1100 is toxic to yeast, but not suppressed by SidJ[44]. Only truncations of SdeA that included both the PDE and ART domains were glutamylated by SidJ *in vitro,* implying that SidJ requires the PDE domain for substrate recognition. SidJ also glutamylated SdeB, SdeC and SidE **(Figure 22B).** To test if SidJ might transfer other amino acids, we performed glutamylation with [U-^{14}C] Glu in the presence of excess unlabeled, "cold" Gln, Asp, Lys, or Gly. Only Gln partially competed with Glu at 10x concentrations **(Figure 23).** In separate experiments, SidJ failed to incorporate [U-^{14}C] Gln into SdeA. Thus, SidJ is a CaM-dependent protein polyglutamylase.

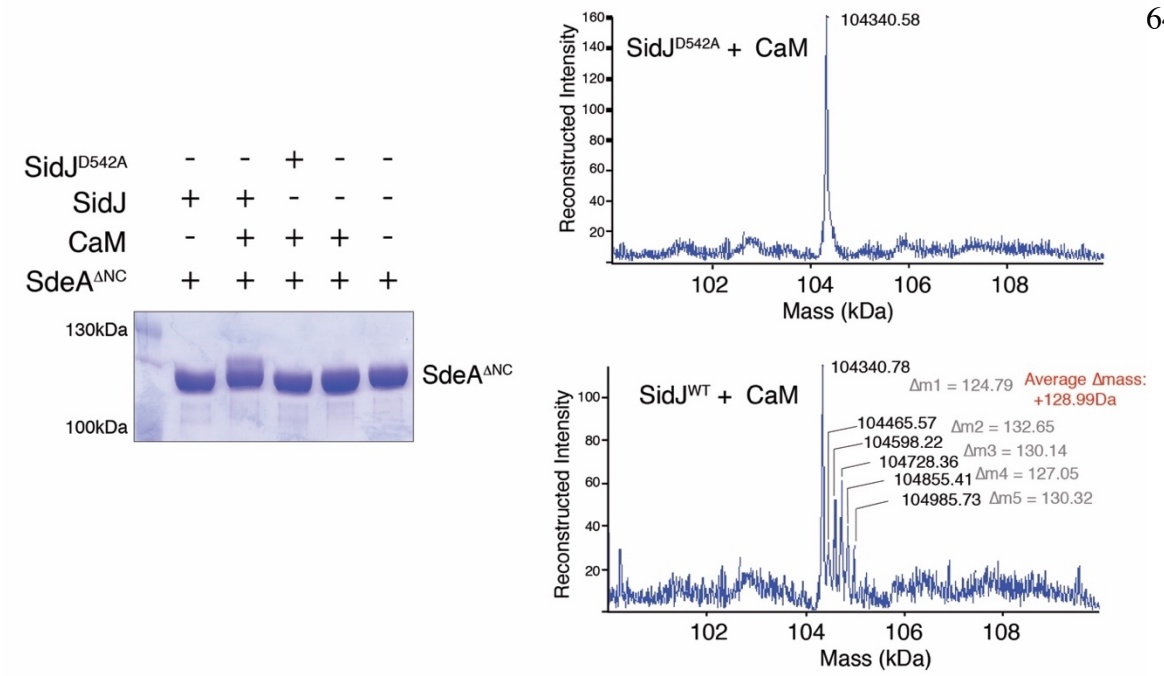


Figure 19. Co-expression of SidJ, CaM, and SidE in *E. Coli.* (Left) SDS-PAGE analysis and Coomassie staining of SdeA$^{\Delta NC}$ and (right) intact mass LC/MS spectra of SdeA$^{\Delta NC}$isolated from *E. coli* following co-expression with CaM, SidJ, or the indicated mutants.

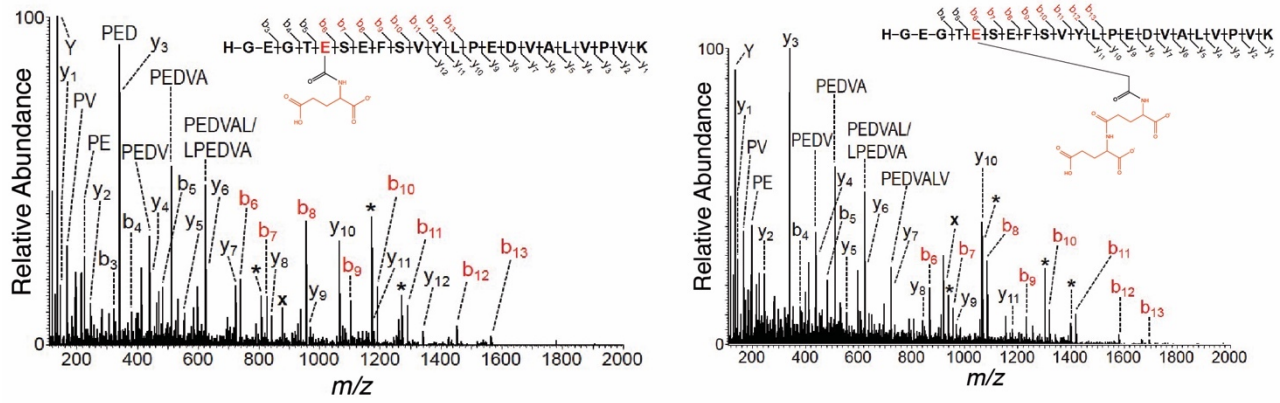


Figure 20. MS/MS spectra of glutamylated and polyglutamylated SdeA. MS/MS spectrum of monoglutamylated (Glu, left) or diglutamylated (GluGlu, right) SdeA$^{\Delta NC}$ peptide ion HGEGTE(Glu)SEFSVYLPEDVALVPVK. The precursor ion, *m/z* 878.10 (3+) and labeled with (x), was subjected to HCD fragmentation to generate the MS/MS spectrum shown. Fragment b-ions containing the modified glutamate residue show a mass shift consistent with the addition of one Glu group (+129.043 Da) or GluGlu (+258.085 Da, lower) (red labels). Peaks labeled with a single asterisk (*) correspond to loss of water (-18 Da) from fragment ions.

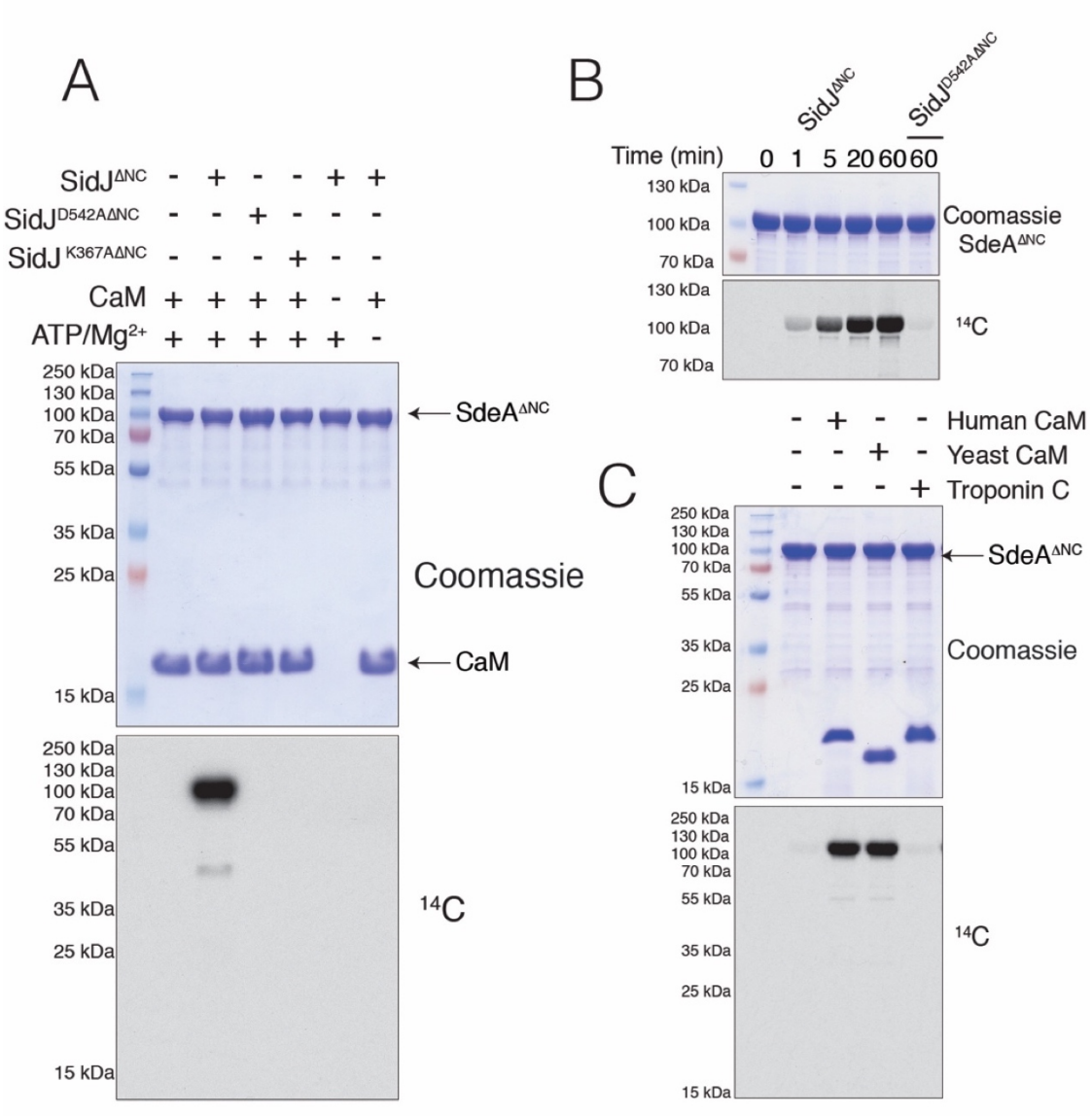


Figure 21. Incorporation of ^{14}C-Glu into SdeA by SidJ. Reaction products were separated by SDS PAGE and the ^{14}C visualized by autoradiography. **(A)** Glutamylation of SdeA$^{\Delta NC}$ required ATP/Mg^{2+}, CaM and the conserved kinase active site residues K367 and D542 in SidJ. The Coomassie stained gel (upper) and autoradiograph (lower) are shown. **(B)** Time-dependent incorporation of ^{14}C-Glu into SdeA$^{\Delta NC}$ by SidJ$^{\Delta NC}$. **(C)** Glutamylation using human CaM, yeast CaM or human troponin C as activators.

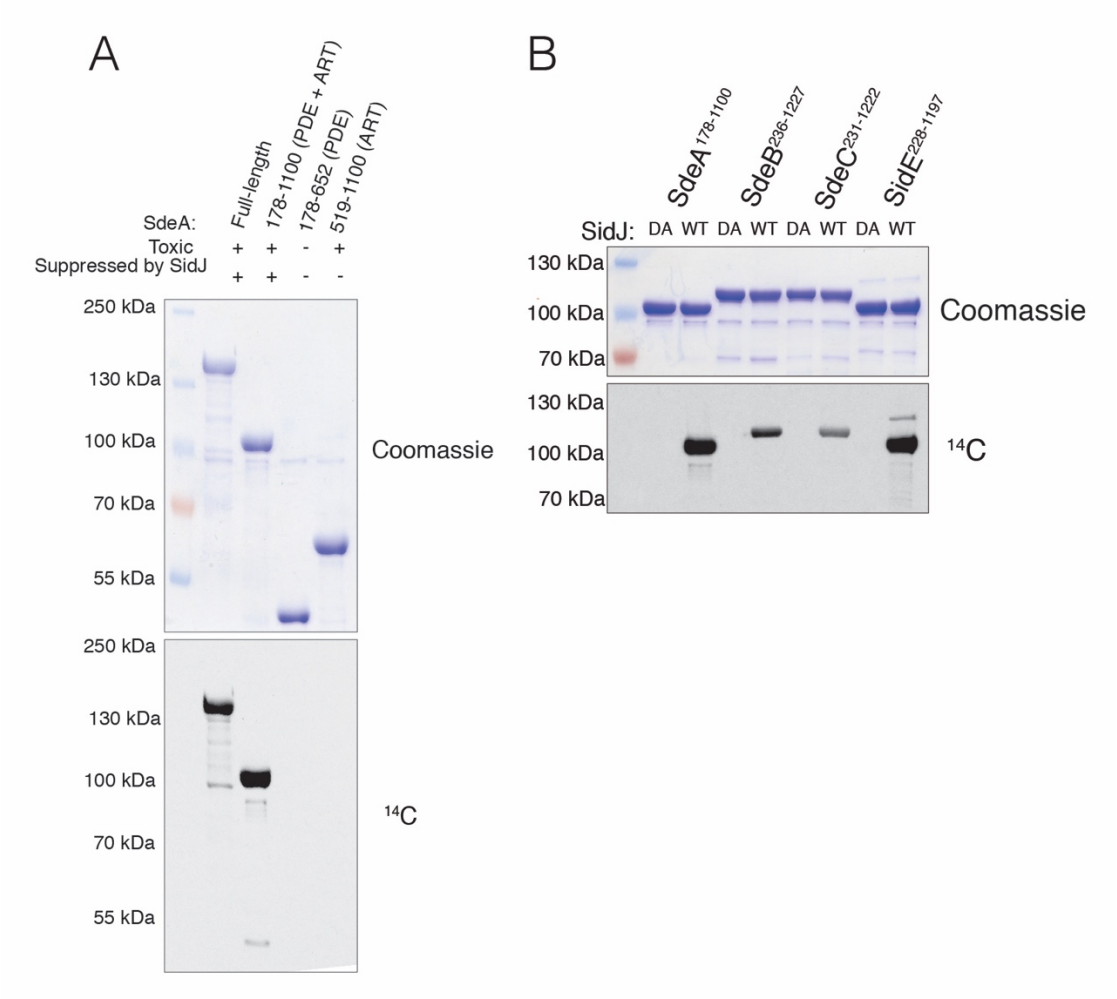


Figure 22. Glutamylation of the SidE effector family. (A) Different truncations of SdeA were tested as substrates of SidJ, including full-length, the ART domain (519-1100), the PDE domain (179-652) and ART+PDE (178-1100). **(B)** Incorporation of 14C-Glu into SdeA$^{\Delta NC}$, SdeB$^{236-1227}$ SdeC$^{231-1222}$, and SidE$^{228-1197}$ by SidJ$^{\Delta NC}$ but not the D542A mutant.

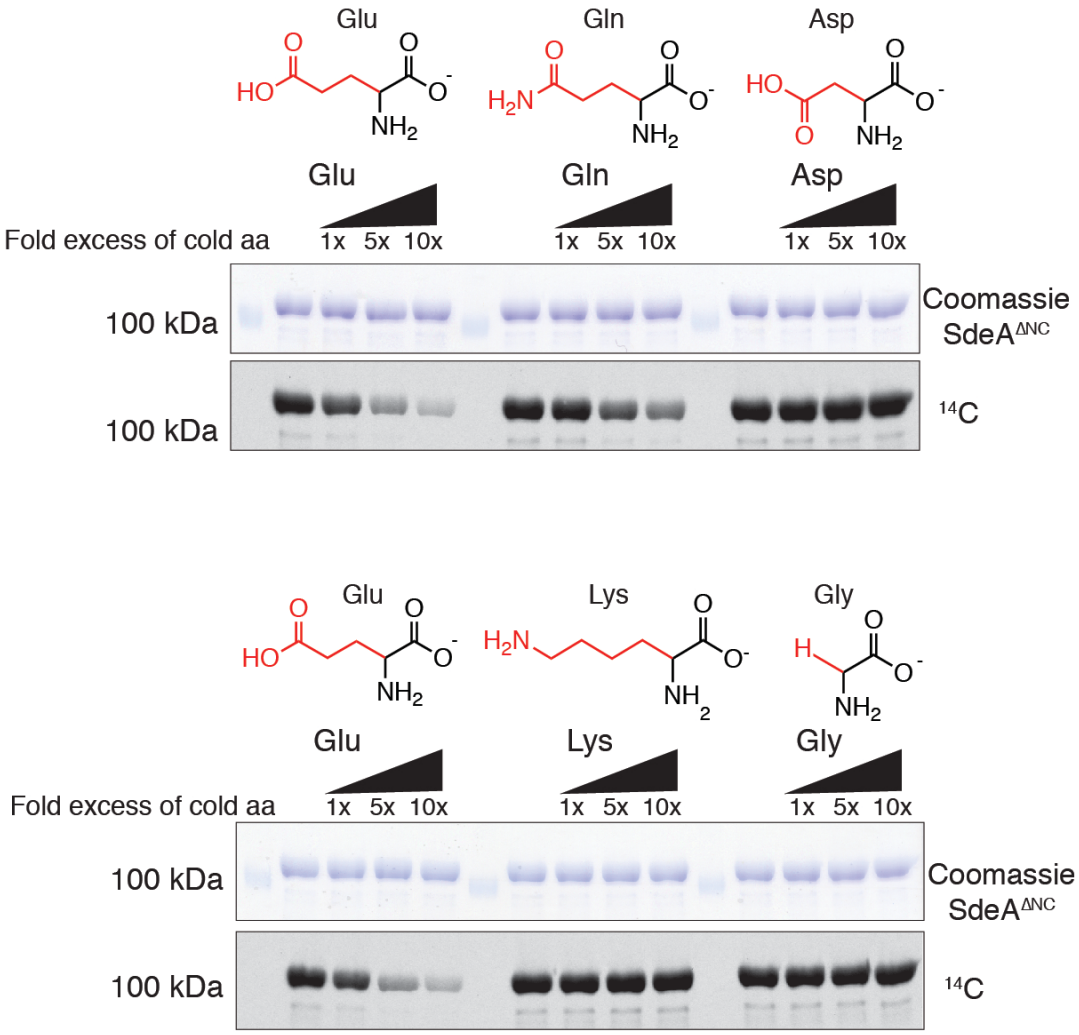


Figure 23. Specificity of SidJ for glutamate. Incorporation of ^{14}C-Glu into SdeA$^{\Delta NC}$ by SidJ$^{\Delta NC}$ in the presence of unlabeled amino acids in 1x, 5x or 10x the concentration of [U-^{14}C] Glu.

SidJ glutamylates a catalytic Glu in the SidE ART active site *in vitro,* in cells and during *Legionella* infection

The majority of the glutamylated tryptic fragments identified by LC-MS/MS corresponded to the ART toxin turn-turn loop of the SdeA ART domain, implicating the conserved catalytic residues E860 or E862 as potential sites of glutamylation. Mutagenesis of each residue to Ala, as well as R766/S820, as a control to inactive SdeA ART activity, revealed that Glu860 is the major site of glutamylation by SidJ **(Figure 24, right panel).** Glu860 is part of the conserved catalytic R-S-E^{860}XE motif present in Arg-specific ARTs. Structures of SdeA and SidE with Ub revealed that E860 plays a key role in catalysis[48-50,52] **(Fig. 24, left panel).** Importantly, LC-MS/MS analysis of full length Myc-tagged SdeA isolated from mammalian cells confirmed that E860 is polyglutamylated when co-expressed with WT SidJ but not the D542A mutant **(Figure 25).** To determine if SdeA is glutamylated on E860 during *Legionella* infection, we co-expressed Myc-tagged SdeA with FcγRIIa and challenged HEK293 cells with antibody-opsonized *Legionella.* SdeA isolated from cells challenged with the WT strain Lp02 yielded tryptic fragments corresponding to glutamylated peptides containing Glu860, while challenge with a Δ*sidJ Legionella* strain or the DotA- strain, Lp03, which lacks a functional T4SS, yielded only unmodified peptides from the ARTT loop **(Table 2).** Thus, SidJ covalently glutamylates a highly-conserved, catalytic glutamate residue in the ART active site of the SidE effectors.

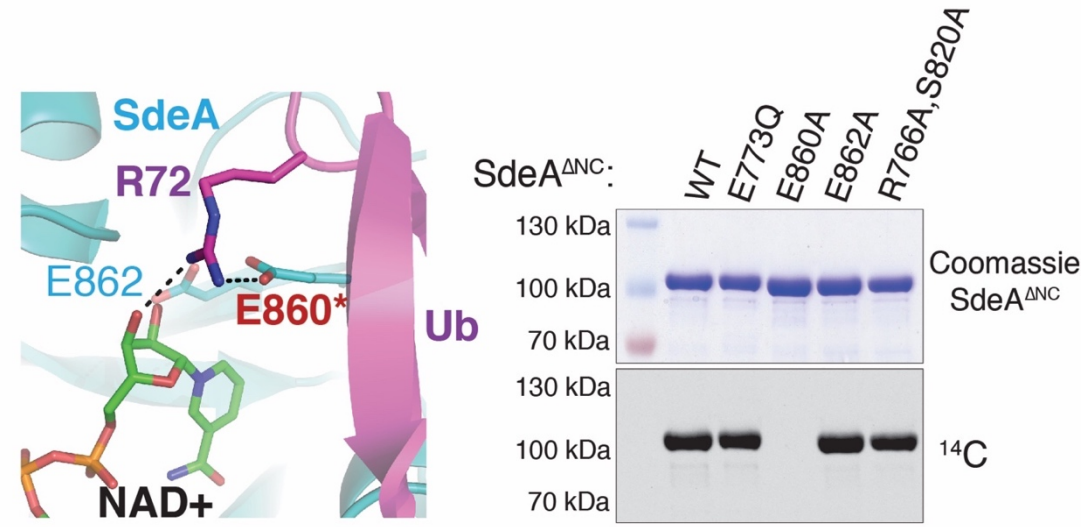


Figure 24. SdeA is glutamylated on a catalytic glutamate in the ART active site. (Left) Close-up of the SdeA ART active site (PDB 5YIJ). NAD+ is rendered as sticks (green). The Ub substrate is shown in magenta, and the critical residue R72 is rendered in stick representation. SdeA is in cyan and the catalytic glutamates are shown as sticks. (Right) Incorporation of ^{14}C-Glu into WT SdeA$^{\Delta NC}$, but not the E860A mutant by SidJ$^{\Delta NC}$. The E773Q, E862A, and R766A/S820A mutants were also analyzed as controls.

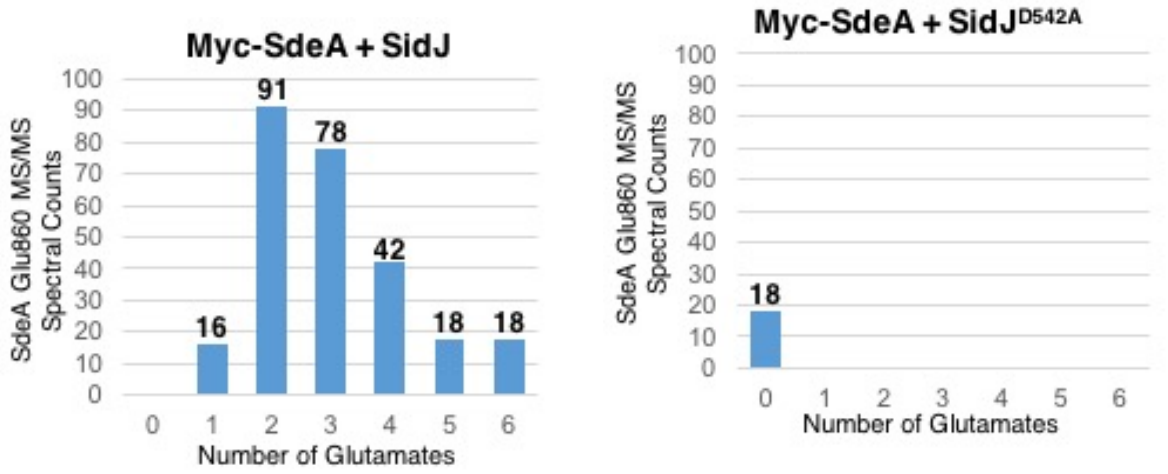


Figure 25. Glutamylation of SdeA co-expressed with SidJ in Mammalian cells. Histograms depicting the MS/MS spectral matches to unmodified, glutamylated, and polyglutamylated E860-containing tryptic peptides of Myc-tagged SdeA immunopurified from HEK293 cells expressing SidJ-V5 or the D542A mutant.

Glutamylation suppresses SidE ubiquitination by blocking ART activity

The target residue of SidJ (E860 in SdeA) is part of the conserved catalytic R-S-EXE motif present in all Arg-specific ARTs, and is strictly required for ADP-ribosylation of Ub. Structures of SdeA and SidE with their substrates revealed that E860 plays a key role in coordinating first R72, then likely engaging R42 of Ub to position it for nucleophilic attack on the NAD^+ oxocarbenium cation intermediate[50,52]. Therefore, glutamylation of this residue is predicted to inhibit SidE ART activity. To assay the activity of the SidE effectors, the SdeA H407A and SdeC H416A mutants were used, which inactivate the PDE domain and allow for specific analysis of ART activity[31]. Glutamylation of $SdeA^{H407A\ \Delta NC}$ inhibited its ability to ADP-ribosylate Ub **(Figure 26, left panel).** Inactivation of $SdeA^{H407A\ \Delta NC}$ required CaM, ATP/Mg^{2+}, Glu and K367 and D542 of SidJ. Substituting Glu with Gln, Asp, Lys, or Gly had no effect on $SdeA^{H407A\ \Delta NC}$ ART activity **(Figure 26, right panel).**

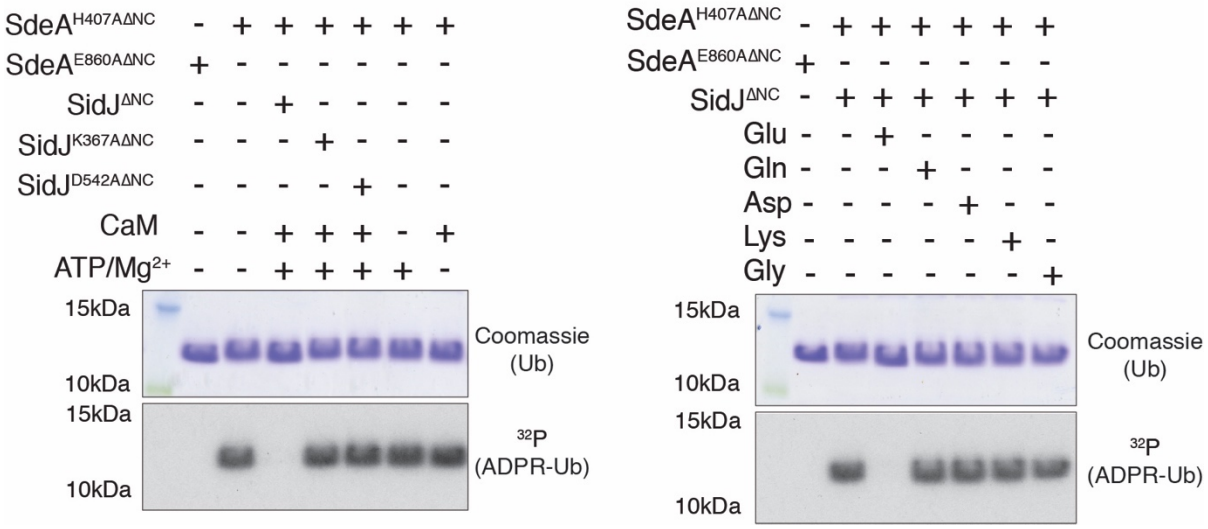


Figure 26. Polyglutamylation inactivates SdeA ADP-ribosylation activity. (Left) SDS PAGE and autoradiography depicting the incorporation of ^{32}P-ADPR into HA-Ub from [α-^{32}P]NAD^{+} by SdeA$^{H407A\Delta NC}$. SdeA$^{H407A\Delta NC}$ or the E860A mutant were pretreated in glutamylation assays with SidJ$^{\Delta NC}$ or the indicated mutants (-/+ ATP/Mg^{2+} or CaM) and SdeA$^{H407A\Delta NC}$ activity was subsequently analyzed. (Right) SdeA$^{H407A\Delta NC}$ or the E860A mutant were pretreated with SidJ$^{\Delta NC}$, CaM, ATP/Mg^{2+} or Glu, Gln, Asp, Lys or Gly and SdeA$^{H407A\Delta NC}$ activity was analyzed.

Consistent with previous reports, expression of SdeA^{H407A}-GFP or SdeC^{H416A}-GFP in yeast induced a potent growth defect, suggesting that ART activity is both necessary and sufficient for toxicity[31]. Indeed, FLAG-SidJ also suppressed the toxicity of both histidine mutants (**Figure 27**), indicating that SidJ targets the SidE effectors at the point of ART activity, rather than detoxifying the ubiquitinated reaction products.

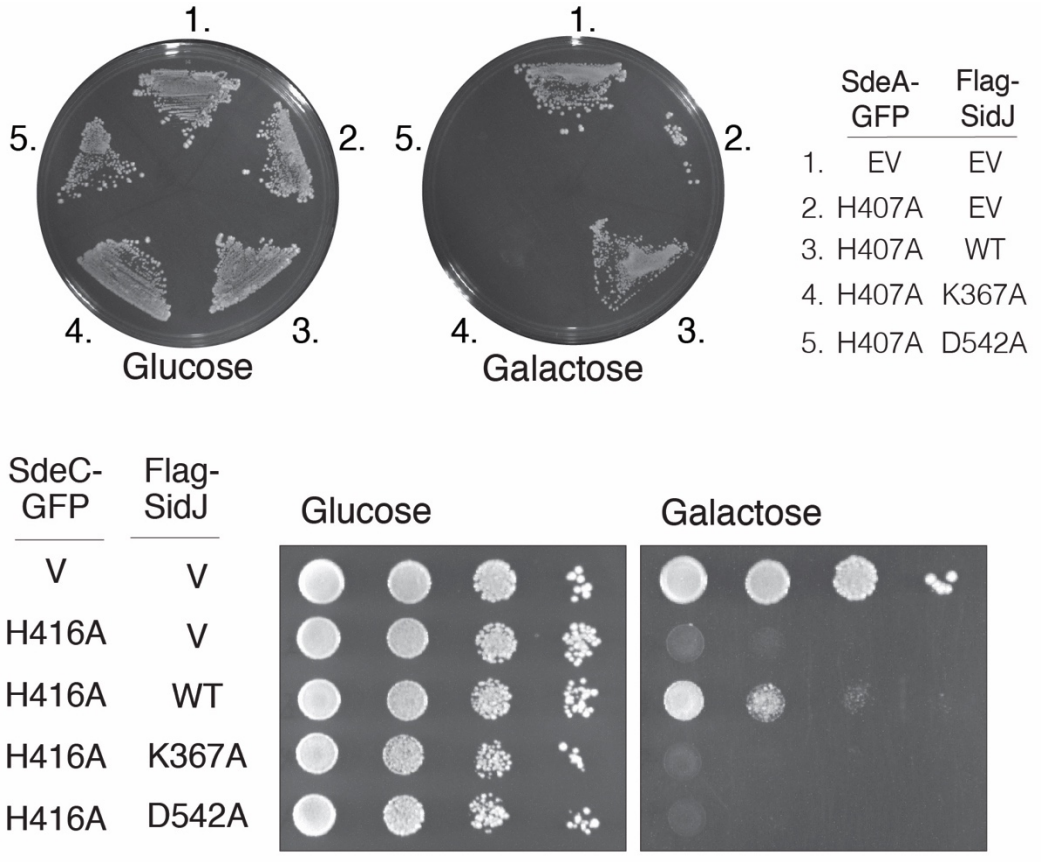


Figure 27. SidJ suppresses the ADP-ribosylation activity of SdeA and SdeC in yeast. Streaking assay depicting the growth of *S. cerevisiae* expressing SdeA^{H407A}-GFP and Flag-SidJ, or the catalytically inactive K367A and D542A mutants. (Lower), Spot-dilution assay depicting growth of *S. cerevisiae* expressing SdeC^{H416A}-GFP and Flag-SidJ, or the K367A and D542A mutants.

Because ART activity is the first step in the SidE-catalyzed ubiquitination reaction, assays were performed with glutamylated SdeA$^{\Delta NC}$ or mutants, and HA-Ub laddering monitored. The reaction products were also immunoblotted using an anti-Ub antibody (Abcam ab7254) whose target epitope matches the sequence surrounding Arg42 and therefore loses immunoreactivity when ADP-ribosylated, or phosphoribosylated, by the SidE family[31]. As expected, mutation of SdeA H407A abolished auto ubiquitination but not ART activity as judged by immunoreactivity to the Ub antibody, while the E860A mutation abolished both (**Figure 28, lanes 1-4**). Substituting Ub for ADP-ribosylated Ub in the reaction bypassed the ART domain and restored Ub laddering by the E860A mutant (**Figure 28, lane 5**). Glutamylated SdeA$^{\Delta NC}$ behaved identically to SdeA$^{\Delta NC}$ E860A in these experiments, losing both its ability to ladder and to phosphoribosylate Ub, yet was unaffected when ADPR-Ub was used as a substrate (**Figure 28, lanes 6-8**). In contrast, when SidJ$^{\Delta NC,\ D542A}$ was used in glutamylation reactions, no inhibition of SdeA was observed at any step of the ubiquitination reaction (**Figure 28, lanes 9-11**). Collectively, these results indicate that SidJ-catalyzed glutamylation of SidE effectors inactivates ART activity and the subsequent ubiquitination of target host proteins.

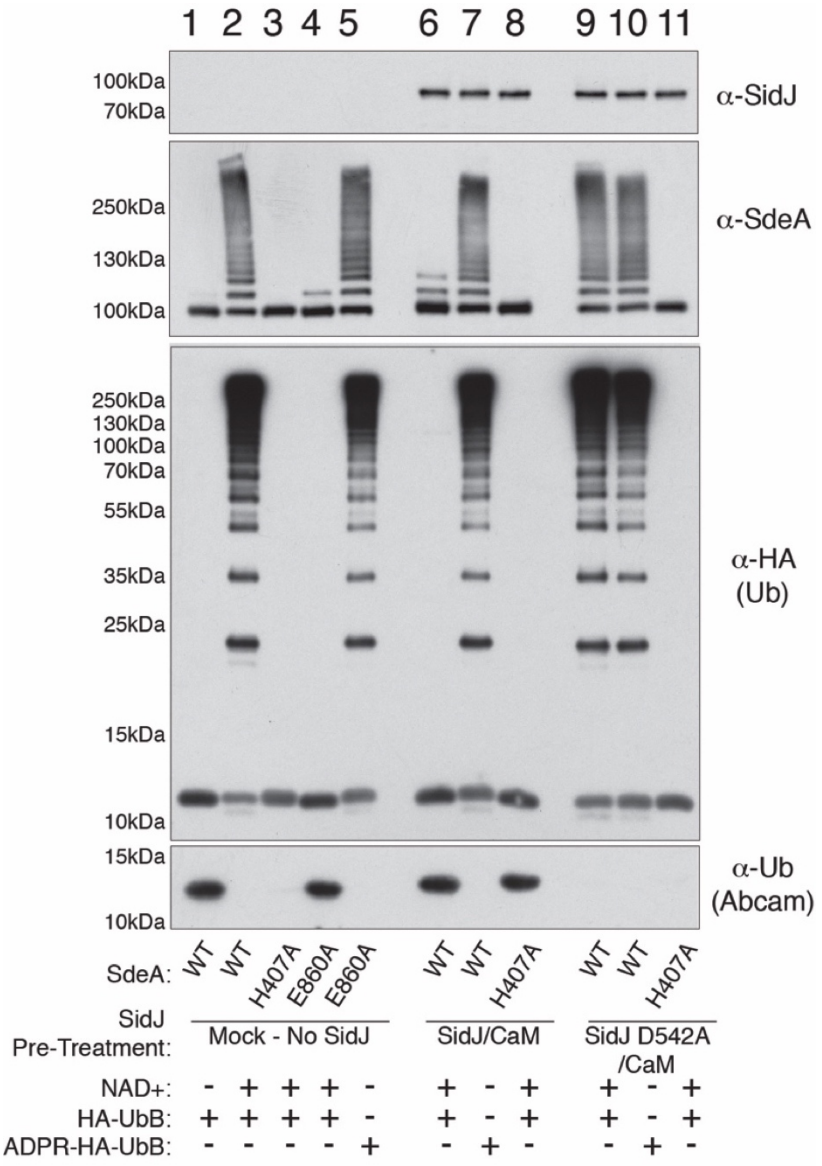


Figure 28. Polyglutamylation inactivates SdeA phosphoribosyl-ubiquitination. Protein immunoblotting following *in vitro* ubiquitination assays with the indicated mutants of SdeA$^{\Delta NC}$. SdeA$^{\Delta NC}$ and mutants were pretreated in glutamylation reactions in the absence (**lanes 1-5**) or presence (**lanes 6-8**) of SidJ$^{\Delta NC}$-CaM, or SidJ$^{D542A\Delta NC}$-CaM (**lanes 9-11**). Ubiquitination reactions were started by the addition of NAD^{+} and HA-Ub or HA-ADPR Ub. The reaction components were resolved by SDS-PAGE and SidJ$^{\Delta NC}$, SdeA$^{\Delta NC}$ and Ub (using anti-HA or Abcam antibodies) were detected by immunoblotting.

Structure of the SidJ-CaM complex reveals a kinase fold with a migrated nucleotide-binding site

Early attempts to crystalize the apo form of SidJ were unsuccessful, as the protein formed phase-separated droplets. CaM was added to stabilize the conformation of SidJ. Recombinant SidJ$^{\Delta NC}$ bound to CaM with a K_D of ~30 nM, and when SidJ was mixed with CaM, the two proteins formed a stable complex on gel filtration chromatography **(Figure 29).** To prepare the crystallization sample, untagged CaM was co-expressed with SidJ$^{\Delta NC}$. Both vertebrate and yeast CaM were screened in extensive crystallization trials, but conditions with yeast CaM (yCaM) produced the most promising hits, which were optimized to obtain singular, diffracting crystals **(Figure 30).**

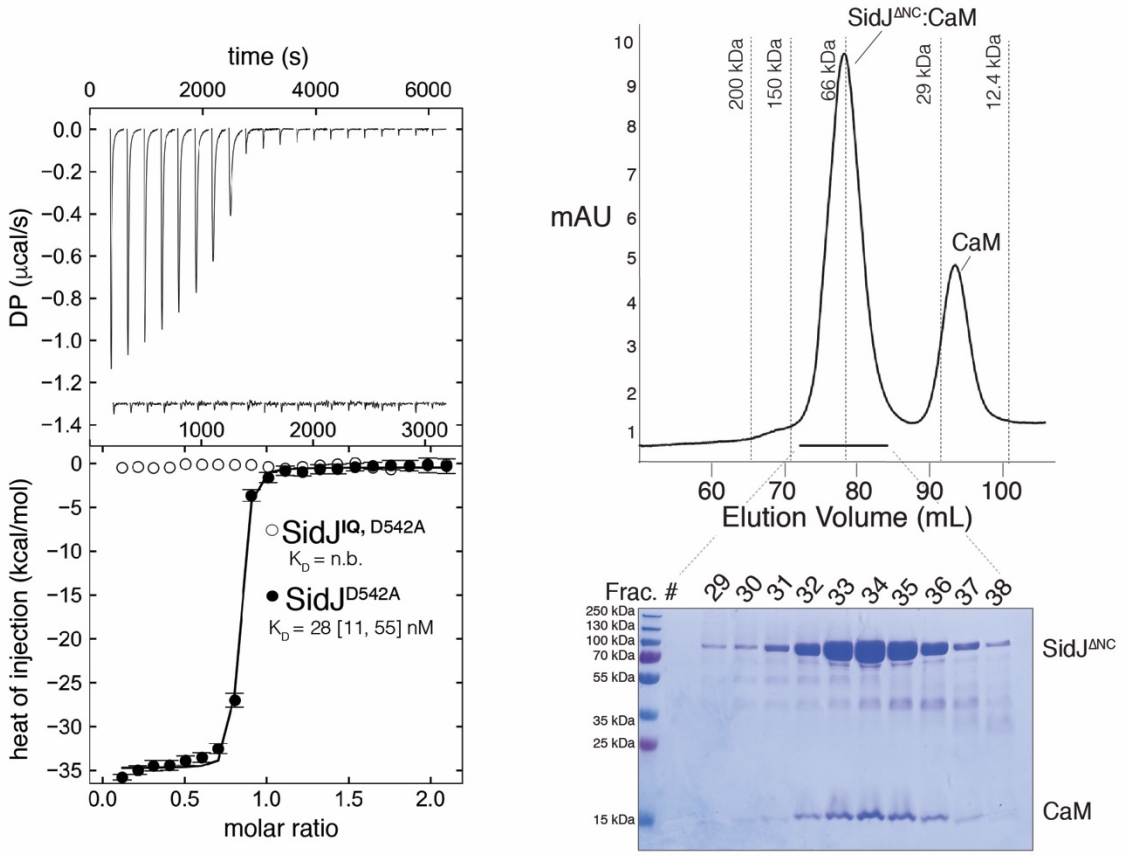


Figure 29. Recombinant SidJ and CaM bind to form a stable complex. (Left) Isotherm depicting the binding of human CaM to SidJ$^{\Delta NC,D542A}$ or SidJ $^{\Delta NC,IQ,D542A}$. The top panel shows the SVD-reconstructed thermograms provided by NITPIC (DP=differential power), the lower panel shows the isotherms. Results are reported as best fit with boundaries of 68.3% confidence interval. (Right) Size exclusion chromatography trace of SidJ$^{\Delta NC}$ and CaM (upper). The indicated fractions were separated by SDS-PAGE and visualized with Coomassie Blue (lower).

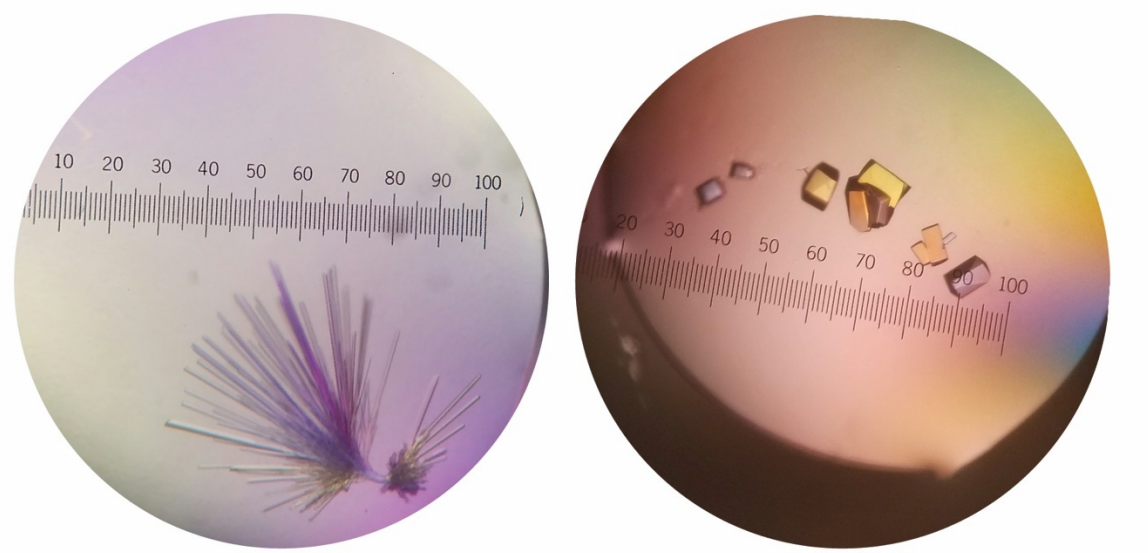


Figure 30. Crystallization of the SidJ-yCaM complex. (Left) Photo of protein crystals of SidJ$^{\Delta NC}$-yCaM obtained 72 h after initial screening with Hampton Index Screen, condition H3 (0.2M Sodium Malonate, pH 7.0, 20% w/v PEG 3350. (Right) Photo of crystals after optimization, initiated by micro-seeding in 0.2M Na Malonate pH 8.0 and 14% w/v PEG 3350.

The structure of SidJ$^{\Delta NC}$ bound to yeast CaM (yCaM) was determined to a resolution of 2.1Å. The crystal belongs to the $P2_1$ space group and two SidJ-yCaM molecules were found in the asymmetric unit **(Appendix A)**. The structure resolves an N-terminal domain (NTD), the kinase domain (KD) and a C-terminal domain (CTD,) of SidJ (**Figures 31 and 32**). Structural homology searches [80] on the SidJ KD identify the Histone H3 kinase Haspin [81] and the AMPylating selenoprotein-O (SelO) [82] as the closest structural homologs. The α-helical NTD and CTD are nestled beneath the C-lobe of the KD, and yCaM is bound to the 'back' of SidJ opposite the nucleotide-binding pocket. As found in all members of the protein kinase superfamily, the SidJ KD exhibits a β-strand-rich N-lobe containing the regulatory αC helix, and an α-helical-rich C-lobe. Although we added the ATP analog α,β-methyleneadenosine 5′-triphosphate (AMP-CPP) to the crystallization buffer, the SidJ KD is bound to AMP, Mg^{2+} and PP_i. We suspect that SidJ bound and hydrolyzed ATP when co-expressed with yCaM in *E. coli.* In typical kinases, the nucleotide is positioned in a pocket between the N- and C-lobes. Surprisingly, in the SidJ structure, AMP, PP_i, and Mg^{2+} are bound in a migrated pocket formed by an insertion in the KD catalytic loop (484-530). This insertion contains highly-conserved residues which bury the adenosine ring in a tight cleft near the base of the SidJ C-lobe (**Figure 33**).

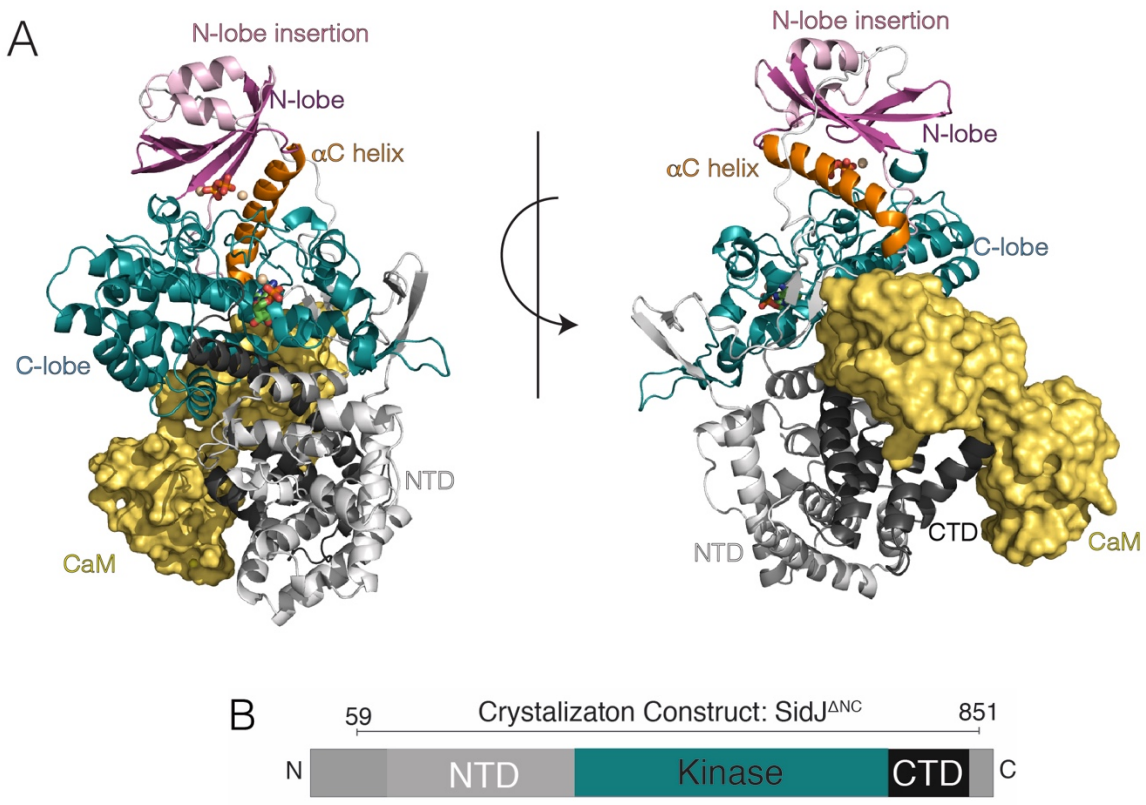


Figure 31. Overall structure of the SidJ-yCaM complex. (A) Overall structure of SidJ$^{\Delta NC}$-yCaM. SidJ and yCaM are in cartoon and surface representation, respectively. The kinase N-lobe of SidJ is colored magenta, with insertions colored light pink. The kinase C-lobe is in teal, the N-terminal domain (NTD) is white, and the C-terminal domain (CTD) is dark grey. yCaM is yellow. Bound nucleotide and pyrophosphate are pictured in stick. (B) Domain architecture of SidJ. Residues corresponding to SidJ$^{\Delta NC}$ are indicated above the schematic.

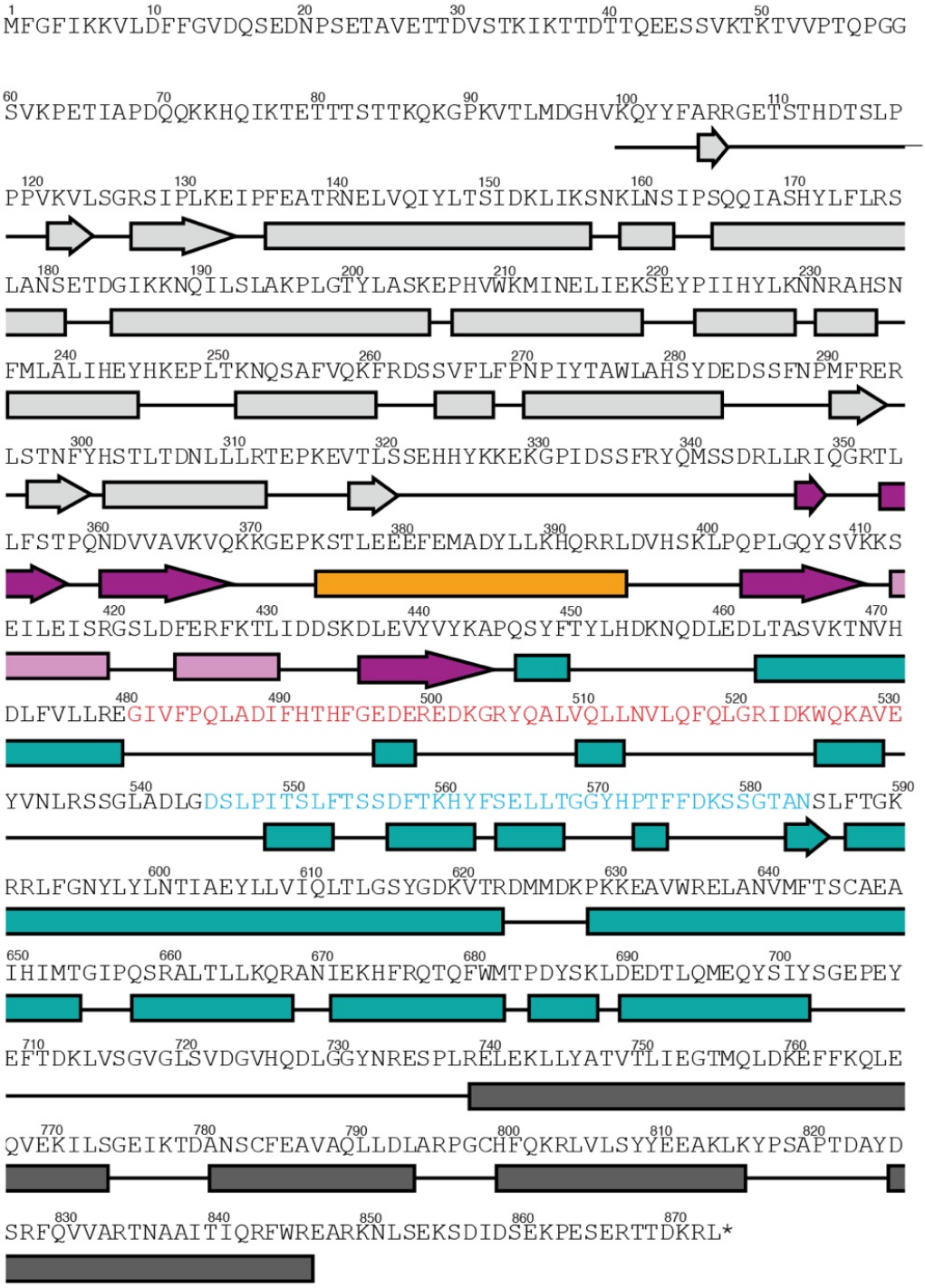


Figure 32. Secondary structure of SidJ. Amino acid sequence of SidJ depicting the secondary structural elements. Coloring is as in Figure 31. The catalytic loop insertion is in red, and the activation loop insertion is in blue.

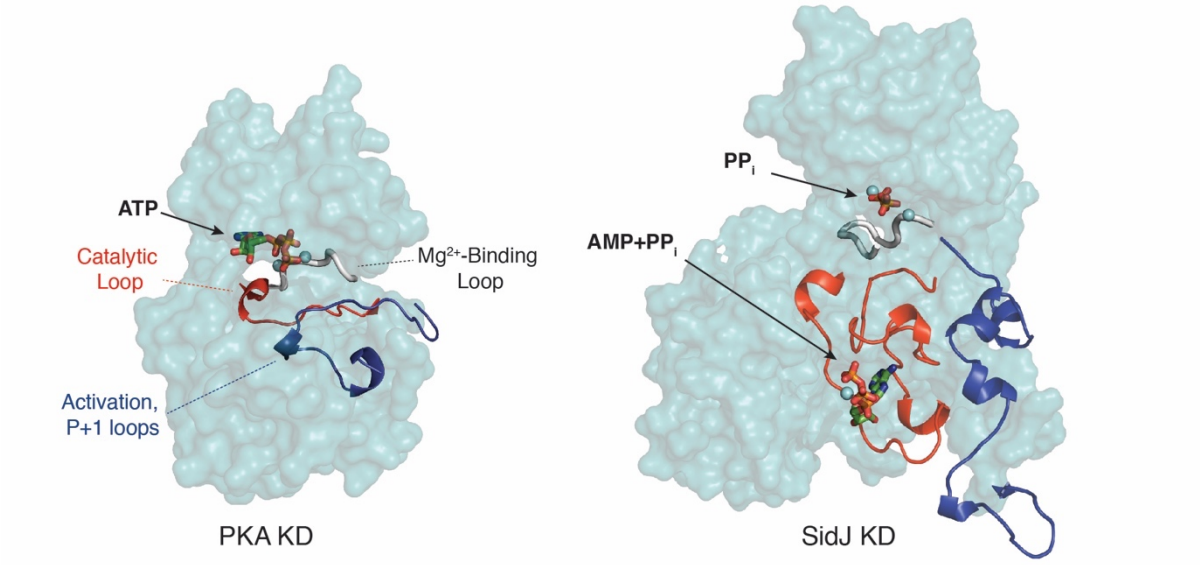


Figure 33. Comparison of nucleotide binding by PKA and SidJ. Surface representations of the kinase domains of PKA (PDB 1ATP) and SidJ highlighting the catalytic loop (red), activation loop (blue), Mg^{2+}- binding loop (white) and bound nucleotides and pyrophosphate (PP_i). Note the AMP/PP_i bound in the migrated catalytic loop insertion in the SidJ structure.

The SidJ canonical kinase active site contains pyrophosphate (PP_i) and 2 Mg^{2+} ions, which are coordinated by standard kinase catalytic residues **(Figure 34A).** In addition to the K367A and D542A mutants, Ala substitutions of R352 and N534 inactivated SidJ glutamylation (**Figure 34B**). Within the migrated nucleotide-binding pocket **(Figure 35A),** Ala substitutions of the invariant H492 and Y506 also inactivated SidJ. R500 and N733 also coordinate the phosphate group of AMP and face into solvent; their mutagenesis to Ala markedly inhibited glutamylation activity **(Figure 35B).** The SidJ mutants which lost glutamylation activity *in vitro* also lost the ability to suppress SdeA-catalyzed ubiquitination of host proteins in mammalian cells **(Figure 36).**

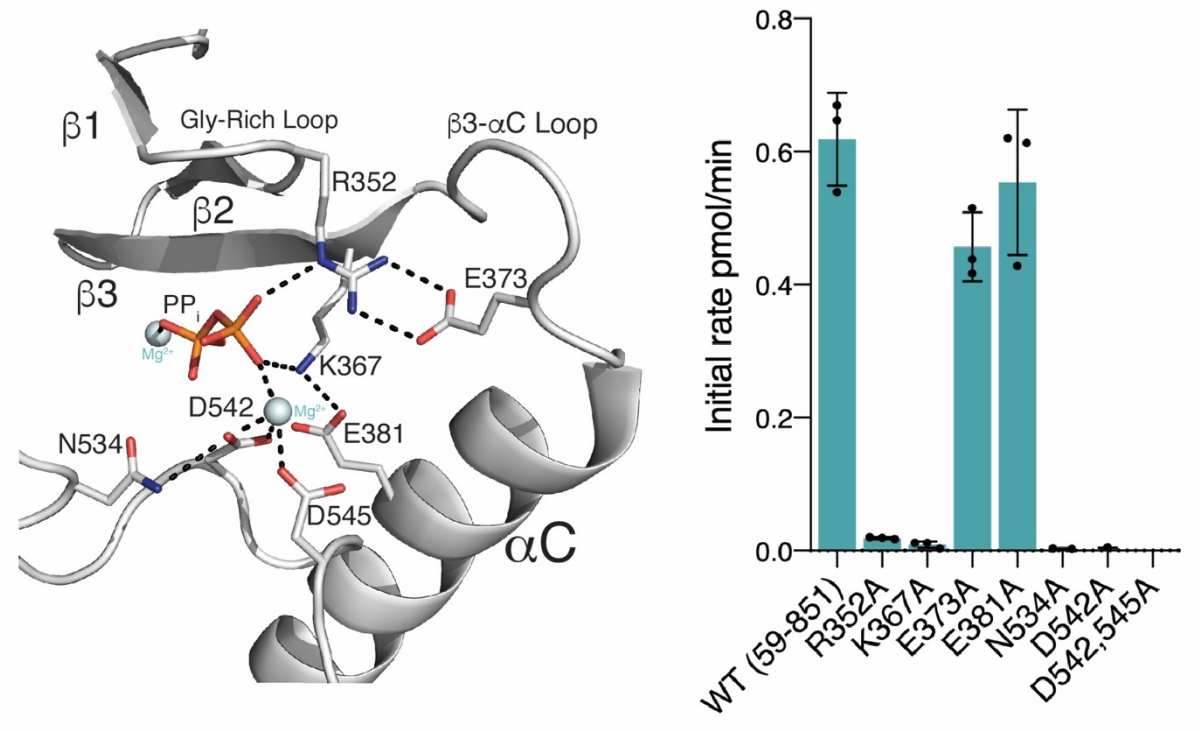


Figure 34. The canonical kinase cleft of SidJ. (Left) Magnified view of the canonical kinase active site of SidJ showing the interactions (dashed lines) involved in PPi binding. The PPi is shown as sticks and the Mg^{2+} ions as light blue spheres. (Right) Glutamylation activity of SidJ and mutants using $SdeA^{\Delta NC}$ and [^{3}H]Glu as substrates. Reaction products were resolved by SDS-PAGE, radioactive gel bands were excised, and ^{3}H incorporation into $SdeA^{\Delta NC}$ was quantified by scintillation counting.

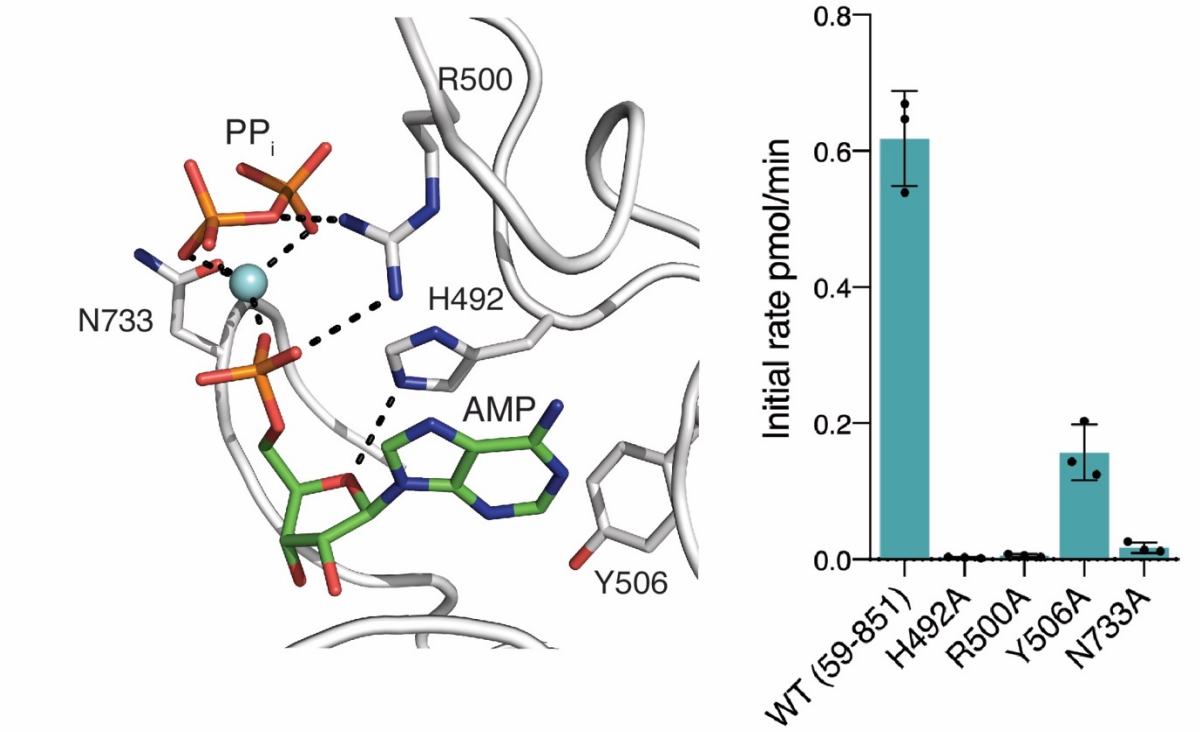


Figure 35. The migrated nucleotide-binding pocket of SidJ. (Left) Magnified view of the migrated SidJ nucleotide-binding site depicting the interactions involved in AMP and PP_i binding. The AMP and PP_i are shown as sticks and the Mg^{2+} ion as a light blue sphere. (Right) Glutamylation activity of SidJ and kinase active-site mutants using $SdeA^{\Delta NC}$ and [^{3}H]Glu as substrates.

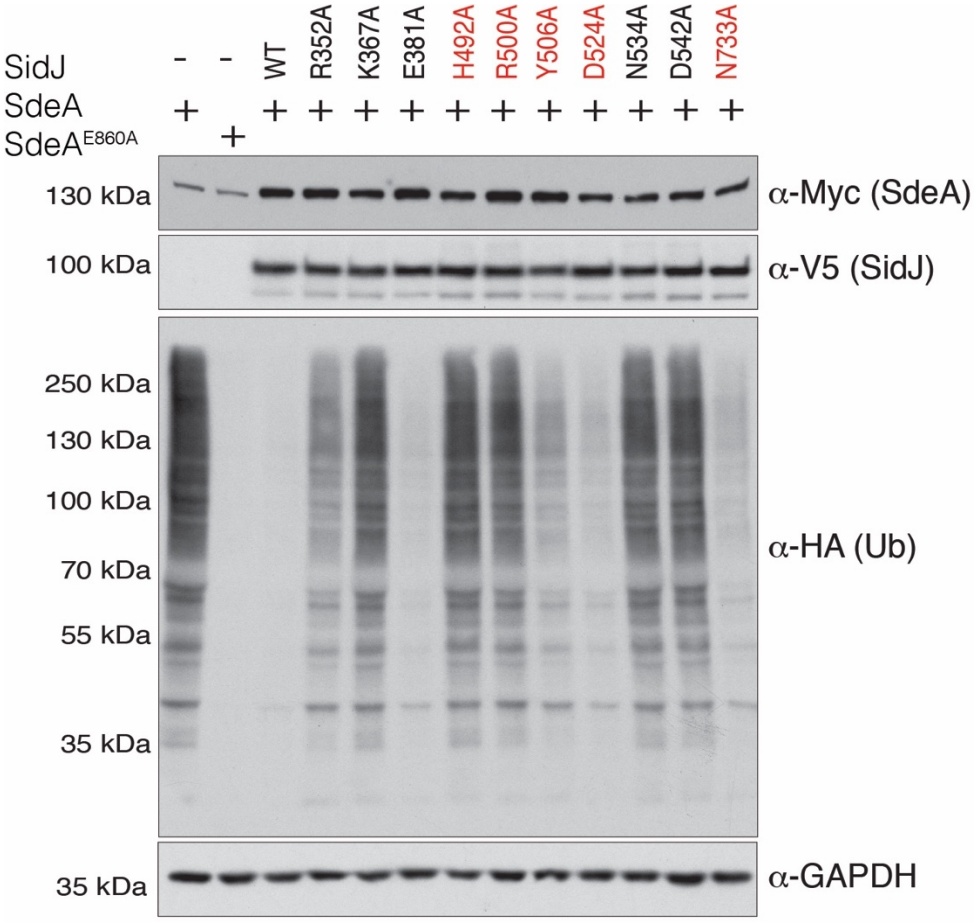


Figure 36. Suppression of SidE-mediated phosphoribosyl ubiquitination by SidJ structure-guided mutants in mammalian cells. Protein immunoblotting of HEK293A cell total extracts expressing HA-Ub$^{GG/AA}$, codon-optimized, Myc-SdeA, SidJ-V5, or structure-guided mutants of SidJ. Mutations colored in black are in the canonical kinase active site while those in red are in the migrated nucleotide-binding pocket.

Structural features of the SidJ-CaM interaction

yCaM adopts a closed conformation with Ca^{2+} ions coordinated in EF1 and EF3 **(Figure 37).** The interactions between SidJ and yCaM span more than 2000 $Å^2$ of surface area and involve both lobes of yCaM. Within the IQ helix, I841 is buried in a hydrophobic cleft of the yCaM C-lobe. R843 and R846 contact Glu's, pinching the yCaM molecule around the IQ helix **(Figure 38, right panel).** The αA and αB helices of the yCaM N-lobe are engaged in a groove formed by the CTD of SidJ **(Figure. 38, left panel)**. To disrupt the binding of the N-lobe of CaM to the CTD of SidJ, we mutated F763, F801 and E812 to Ala, Glu, Ala, respectively, to generate the mutant "$SidJ^{FFE}$". $SidJ^{FFE}$ bound to CaM with 100-fold lower affinity than WT SidJ and, $SidJ^{IQ+FFE}$ failed to bind CaM entirely **(Figure 39A-D).** The CaM-binding mutants were also significantly impaired in their ability to glutamylate $SdeA^{\Delta NC}$ *in vitro* **(Figure 39E)**.

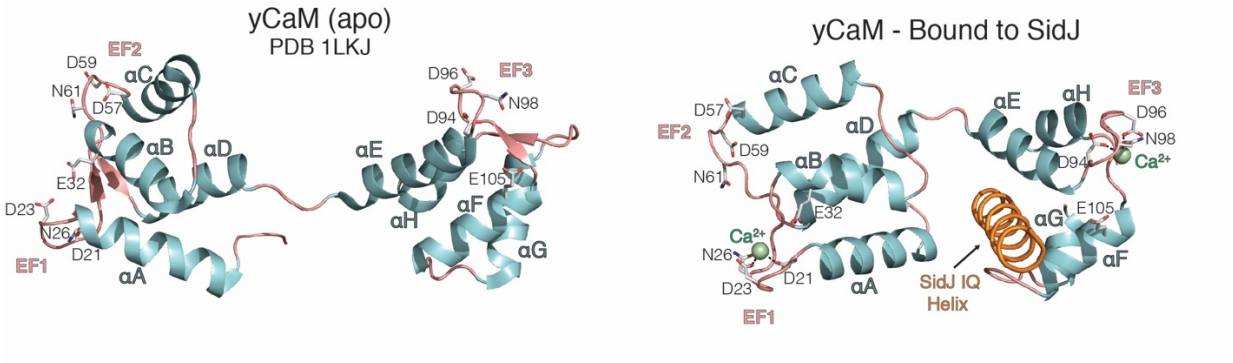


Figure 37. Conformation of yCaM bound to SidJ. Ribbon representations of the structures of apo yCaM (left) (PDB 1LKJ) and SidJ-bound CaM (right). Helices are colored teal, while strands and loops are in salmon. The SidJ IQ helix is also shown in orange.

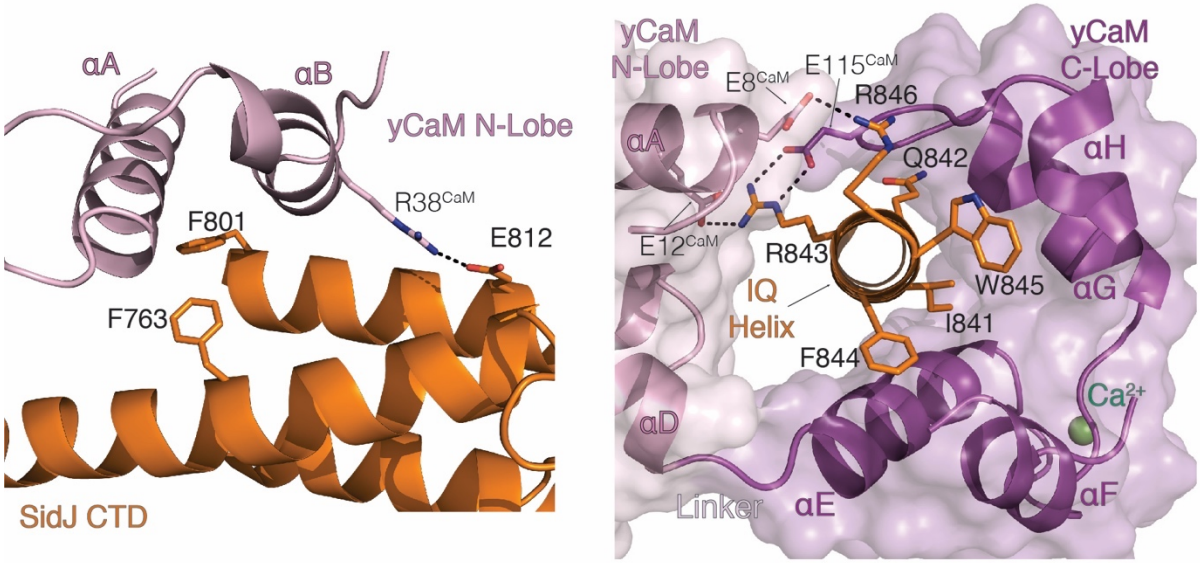
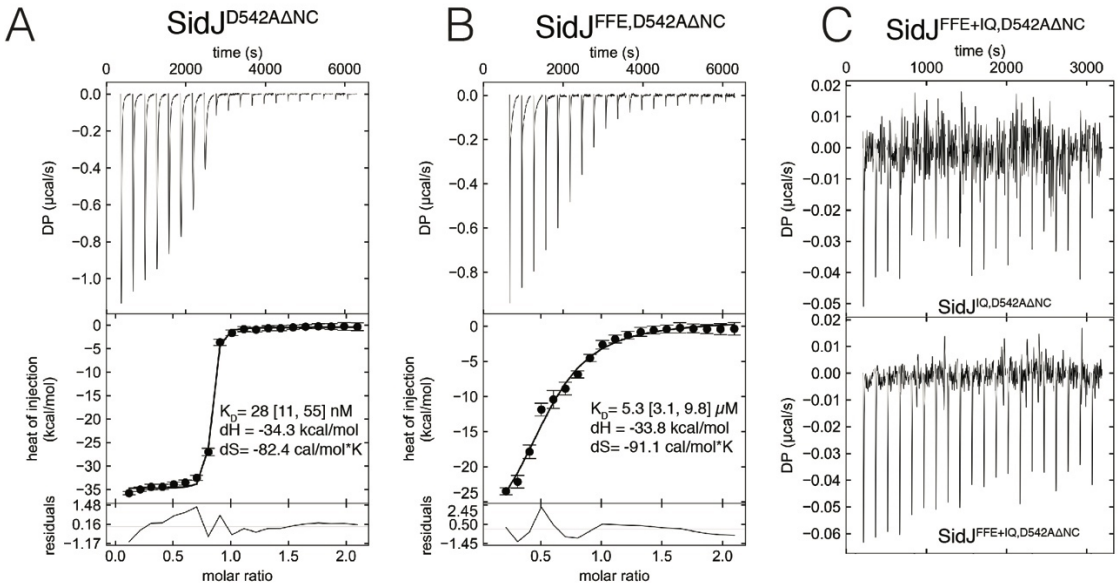


Figure 38. The SidJ-yCaM interface. (Left) Magnified view of the interaction between the N lobe of CaM and residues in the CTD of SidJ. (Right) Magnified view of the IQ helix of SidJ and its interactions with the N and C lobes of yCaM. SidJ is colored orange, and the yCaM N lobe and C lobes are in pink and magenta, respectively.



D

	K_D (nM)
SidJ^{D542AΔNC}	28 [11, 55]
SidJ^{IQ,D542AΔNC}	n.b
SidJ^{FFE,D542AΔNC}	5300 [3100, 9800]
SidJ$^{FFE+IQ,D542AΔNC}$	n.b

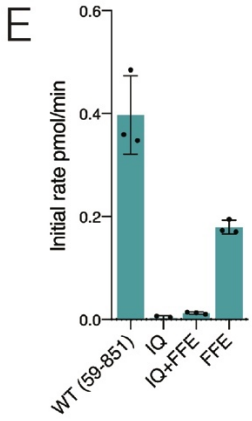


Figure 39. Structure-guided mutagenesis of the SidJ-yCaM interface. (A-C) Isotherms depicting the binding of human CaM to SidJ^{ΔNC,D542A} (A), SidJ^{ΔNC,FFE,D542A} (B), and SidJ$^{ΔNC,FFE+IQ,D542A}$ (C). In panels (A) and (B), the top panel shows the SVD-reconstructed thermograms provided by NITPIC (DP=differential power), the lower panel shows the isotherms. Results are reported as best fit with boundaries of 68.3% confidence interval. Panel (C) depicts SVD-reconstructed thermograms showing no detectable binding with the CaM-binding SidJ mutant. (D) Table depicting the binding affinities (K_D) between the SidJ mutants and human CaM as quantified by ITC. (E) Glutamylation activity of SidJ- and CaM-binding mutants.

SidJ mutations impair *Legionella* replication in amoeba

SidJ is one of only a handful of T4SS effectors which produce a phenotype when ablated from the *L. pneumophila* genome [35,41]. The Δ*sidJ Legionella* strain was complemented with SidJ mutants and growth was monitored within the environmental host *Acanthamoeba castellanii*. Deletion of SidJ resulted in a marked growth defect, consistent with previous reports[43]. Complementation of the Δ*sidJ Legionella* strain with WT SidJ, but not the K367A, D542A or the SidJ^{IQ+FFE} mutants, restored replication (**Figure 40**). Thus, SidJ requires kinase and CaM binding residues to facilitate *Legionella* infection by inactivating SidE.

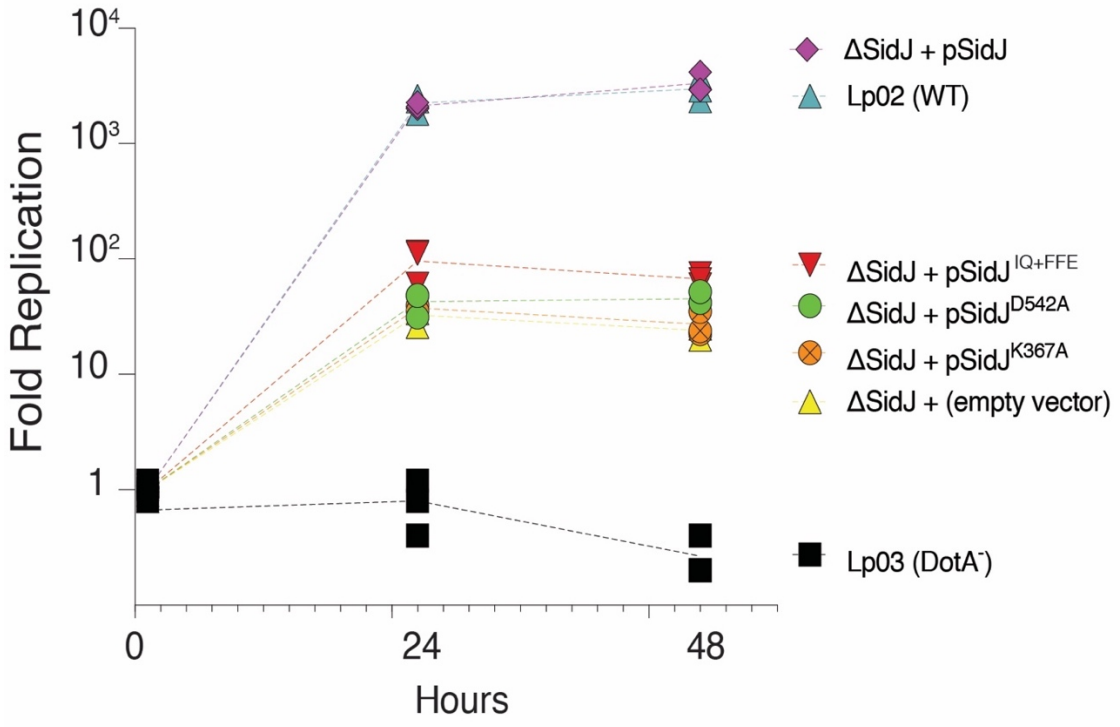
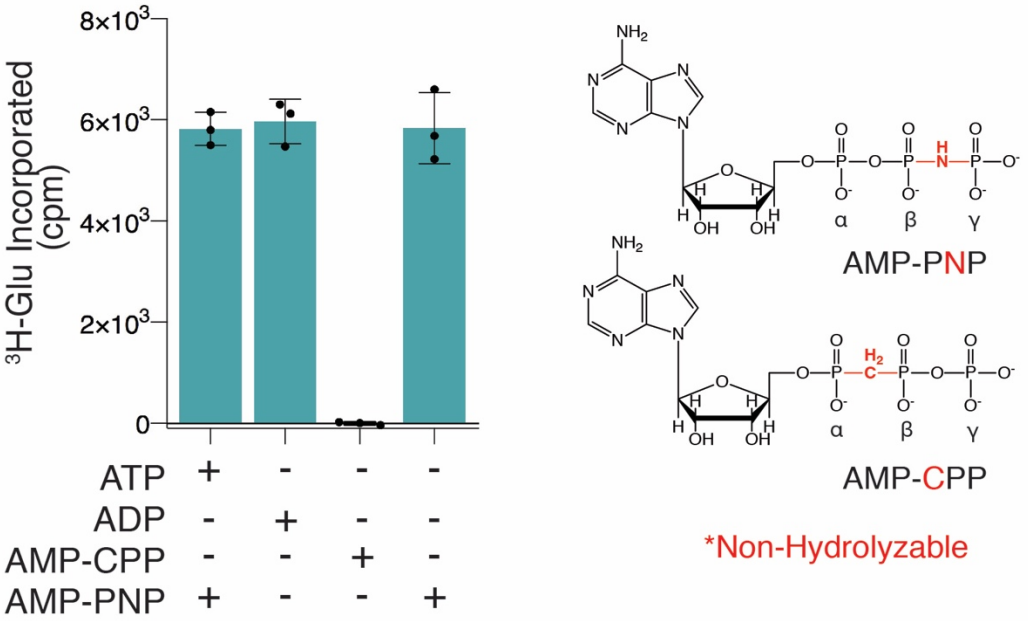


Figure 40. Replication of *L. pneumophila* strains in *A. castellanii*. Infected amoeba cells were lysed at the indicated time points and bacterial replication was quantified by plating serial dilutions of lysates. Results are representative of three independent experiments.

SidJ activates SidE for glutamylation through an acyl-AMP intermediate

Protein glutamylation is an amidoligase reaction which typically proceeds through an acyl-phosphate intermediate[83,84]. However, both acyl-phosphate and acyl-AMP introduce plausible leaving groups and might precede the formation of a Glu-Glu isopeptide bond in the SidJ reaction. To distinguish between these possibilities, glutamylation reactions were performed with ATP, ADP, AMP-PNP (non-hydrolysable bond between the β- and γ-phosphates) and AMP-CPP (non-hydrolysable bond between α- and β-phosphates). Although SidJ prefers ATP, it was also able to use ADP and AMP-PNP, but not AMP-CPP, as co-substrates in glutamylation reactions **(Figure 41).** These results suggest that SidJ-catalyzed glutamylation uses the energy of the ATP α-β-phosphate linkage. Thus, the reaction likely forms an acyl-adenylate between Glu860 of SdeA and AMP. This high-energy intermediate may be subsequently attacked by the amino group of free Glu to form an isopeptide bond **(Figure 42).** In support of this mechanism, we detected AMP, but not ADP, during the SidJ glutamylation reaction only when Glu was added **(Figure 43 and Table 2).** To detect the intermediate, [^{32}P]AMP incorporation from [α-^{32}P]ATP was monitored during glutamylation reactions in the absence of Glu. Because the intermediate contains an alkali unstable acyl-adenylate, reactions were terminated with trichloroacetic acid (TCA) and the acyl-adenylate was detected as ^{32}P in TCA-insoluble material. SidJ adenylated SdeA but not the SdeA^{E860A} mutant (**Figure 44**). Addition of Glu decreased acid-insoluble label, suggesting that AMP is

liberated by the formation of the Glu-Glu isopeptide bond. We propose a catalytic mechanism for SidJ glutamylation of SdeA whereby CaM binds SidJ to stabilize an active conformation, which allows the canonical kinase-like active site of SidJ to bind ATP and transfer AMP to E860 on SdeA. Adenylated SdeA then binds the migrated nucleotide binding pocket in SidJ, which positions the acyl adenylate and glutamate for glutamylation and inactivation of the SidE-family of effectors.



K_m(ATP) = 4.03 ± 0.24 μM, V_{max}(ATP) = 0.89 ± 0.01 pmol/min

K_m(ADP) = 37.29 ± 5.19 μM, V_{max}(ADP) = 0.27 ± 0.01 pmol/min

K_m(AMP-PNP) = 166.90 ± 21.25 μM, V_{max}(AMP-PNP) = 0.45 ± 0.02 pmol/min

Figure 41. Glutamylation activity of SidJ with different nucleotide analogs. SdeA$^{\Delta NC}$ was glutamylated using [^{3}H]Glu. Reactions went to completion. The chemical structures of AMP-PNP and AMP-CPP are also shown with the nonhydrolyzable bonds in red. K_m and V_{max} were calculated for ATP, ADP and AMP-PNP and shown below the figure.

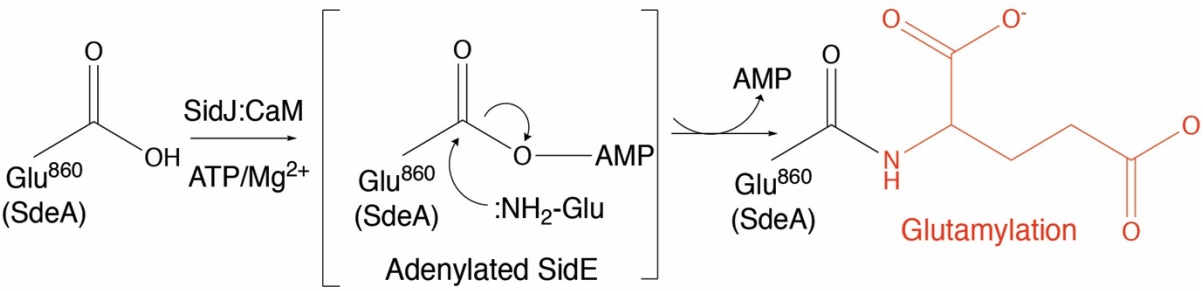


Figure 42. Schematic representation of the proposed SidJ-catalyzed glutamylation reaction. The acyl-AMP SidE intermediate in brackets.

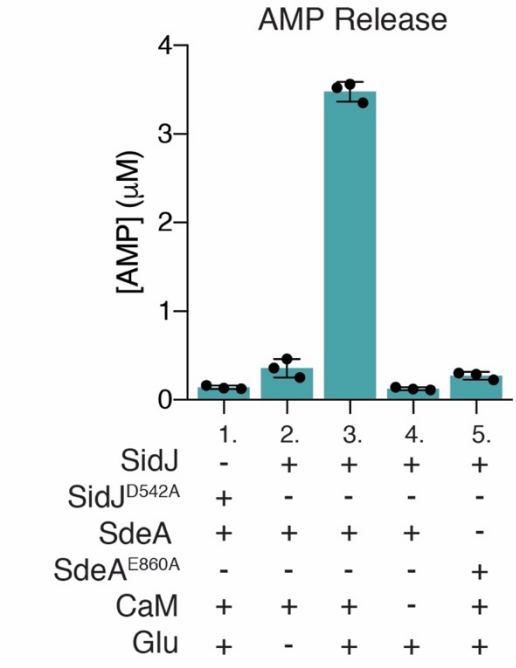


Figure 43. Quantification of AMP release by SidJ glutamylation. Levels of adenine nucleotides were analyzed using HPLC-MS/MS. AMP levels were quantitated using a six-point standard curve. Analyte/internal standard ratios are given in **Table 2**.

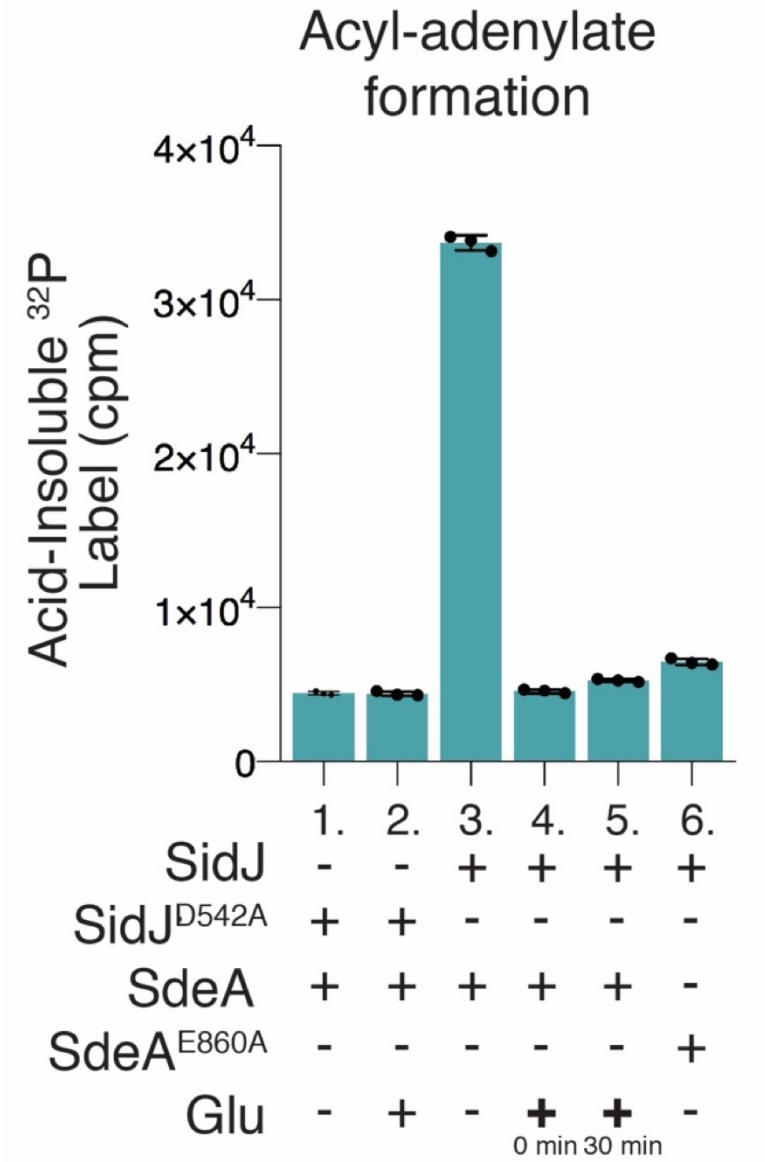


Figure 44. Detection of acyl-adenylate intermediate. After reactions with SidJ$^{\Delta NC}$, SdeA$^{\Delta NC}$, and [α-32P]ATP, reactions were terminated by the addition of trichloroacetic acid (TCA) and the SdeA acyl-adenylate was detected by scintillation counting of the acid-insoluble material. In lanes 4 and 5, Glu was added at time 0 or after 30 min.

SdjA is a protein polyglutamylase

Legionella pnuemophila contains a paralog of SidJ, SdjA (Lpg2508). SdjA is shorter than SidJ (807 aa), but retains 52% identity with SidJ and all the catalytic residues, as well as the IQ motif. However, the fact that endogenous SdjA cannot compensate for SidJ activity in the *ΔsidJ* strain indicates that it is not redundant with SidJ during infection. SdjA was found to potently suppress SdeB, SdeC, and SidE in yeast[47], but deletion of SdjA in *Legionella* caused only a mild replication defect in *Dictyostelium* amoeba, and a double deletion of SidJ and SdjA did not cause a more drastic phenotype than loss of SidJ alone[32]. Recombinant SdjA, residues 36-789 ($SdjA^{\Delta NC}$) was also found to possess CaM-dependent, ATP-dependent glutamylation activity. However, it only glutamylated SdeB, SdeC, and SidE **(Figure 45, upper panel).** 'Partial redundancy' could account for the weak SdjA phenotypes. To test this, a *Legionella* strain was generated with deletions of *sidJ*, *sdjA*, and *sdeA* (ΔΔΔ). This strain displays a replication defect similar to the Δ*sidJ* strain in *Acanthamoeba*, and was nearly completely rescued by complementation with SidJ. Complementation with SdjA was expected to rescue the replication to a similar extent, since SdjA can glutamylate the remaining SidE effectors in this strain. However, SdjA failed to rescue the ΔΔΔ phenotype **(Figure 45, lower panel).** The reasons for this are not clear but may reflect differences in SdjA expression, secretion, or activity. Ongoing work aims to determine structural elements that confer substrate specificity to SdjA and its role during infection.

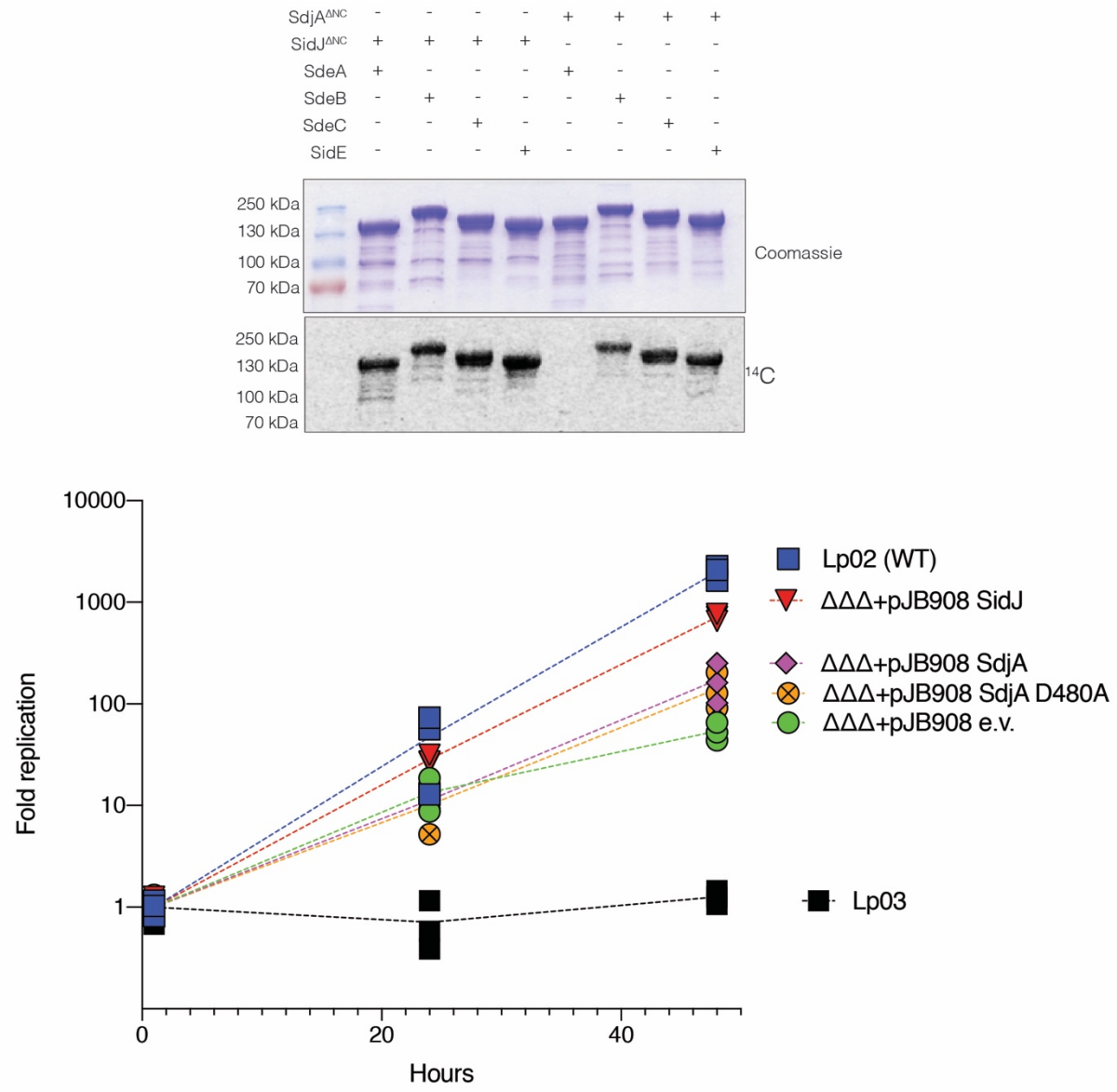


Figure 45. SdjA is a protein polyglutamylase. Upper panel: [^{14}C] glutamylation assays with either SidJ$^{\Delta NC}$ or SdjA 36-789 performed with full-length SdeA, SdeB, SdeC and SidE as substrates. Lower panel: Replication of L. pneumophila Δ*sdjA*/Δ*sidJ*/Δ*sdeA* (ΔΔΔ) strains in *A. castellanii*. Infected amoeba cells were lysed at the indicated time points and bacterial replication was quantified by plating serial dilutions of lysates.

CHAPTER FIVE
Conclusions

Three independent studies have now confirmed the polyglutamylation activity of SidJ and provided several additional high-resolution structures of the SidJ-CaM complex [85-87]. These studies have overwhelmingly supported the conclusions in this work, but also shed light on important outstanding questions and nuances concerning SidJ activity and its role in *Legionella* infection.

The SidJ catalytic cycle

One unresolved question is the role of the two 'active sites' of SidJ – the canonical kinase active site and the migrated nucleotide-binding pocket – during catalysis. It seems likely that the acyl-adenylate intermediate is catalyzed by the canonical kinase active site. This is reminiscent of SelO, which also uses a canonical kinase active site to transfer AMP to its substrates. Furthermore, the SidJ ion-pair lysine and metal binding motifs are required for AMPylation and PP_i exchange in addition to glutamylation. The SidJ canonical kinase cleft might bind ATP in a 'flipped' orientation, like SelO, with the adenosine moiety positioned into the solvent for transfer to E860 of SdeA. However, despite adding AMP-CPP to the crystallization conditions of several SidJ mutants, most of the crystal structures to date have revealed only pyrophosphate, water, and metals bound to this site. Only one group has interpreted electron density in this pocket to correspond to AMP in a 'flipped' orientation[87], however, the density is somewhat hard to interpret.

The role of the migrated nucleotide-binding pocket of SidJ is more ambiguous. An interpretation of the data presented here is that this pocket has evolved to 'grip' the AMP moiety of the acyl-adenylate intermediate, then position glutamate for nucleophilic attack on the δ-carbon of SdeA E860. This conclusion is supported by the presence of two strongly conserved residues in the migrated pocket which do not coordinate the nucleotide, but face into the solvent: R500 an N733. Mutagenesis of these residues abolished glutamylation, but not PPi exchange (unpublished data). These residues may therefore coordinate the free glutamic acid or stabilize a transition state when SdeA is bound.

Sulpizio, et. al.[86] offer an alternative explanation that the migrated pocket is an allosteric site which does not participate in catalysis, but binds adenosine nucleotides to order the kinase activation loop. In this scenario, both adenylation and glutamylation would be catalyzed by the canonical kinase active site. This conclusion is based on mutagenesis of nucleotide-binding residue H492 in the migrated pocket; Alanine substitution of H492 abolishes glutamylation but also auto-AMPylation.

To determine the catalytic role of each site, it will be necessary to determine how SidJ binds the SidE substrate and glutamate. Structures of the SidJ-CaM complex, auto-AMPylation, and PP_i exchange assays, which occur in absence of SidE and Glu, detect futile side reactions that do not necessarily provide useful

information about the genuine catalytic cycle. It is indeed possible that a structural reorganization brings the two pockets of SidJ together into a single active site when adenylated SidE is bound.

Mode of glutamate binding

Determining the residues which endow SidJ specificity for glutamic acid would be an important step. However, this is challenging because the formation of the acyl-adenylate appears to precede glutamate binding, and adding glutamate to the intermediate will result in rapid resolution of the intermediate. In an attempt to identify a non-transferable glutamate analog, the analog 2-methylglutaric acid (glutamate with the amino group replaced with a methyl group) was introduced in glutamylation reactions, but it failed to compete with [^{14}C]Glu. One approach to identify the mode of glutamate binding would be to mutate candidate residues (like R500 and N733) and monitor incorporation of [^{14}C]Gln or other amino acids to change the specificity of SidJ.

Mode of SidE substrate recognition and other substrates

These experiments indicate that SidJ recognizes both the PDE and ART domains of SdeA, and will not glutamylate the minimal ART domain. This implies a multivalent mode of recognition and that the SidJ-SidE interface may be extensive. Future structural studies are warranted to pinpoint this interface. This result also implies that the SidE effectors may be the only substrates of SidE. In support of this conclusion, SidJ glutamylation activity does not cause toxicity to yeast. Nonetheless,

Bhogaraju *et. al.*[85] identified host substrates of SidJ by performing affinity purification/MS with the polyglutamylation antibody GT-335. A number of proteins were enriched from cells infected with WT legionella, but not a Δ*sidJ* strain. However, these substrates have yet to be confirmed. In my experiments the GT-335 antibody fails to detect *in vitro* SidJ-catalyzed polyglutamylation by immunoblot, but immunopurification was not attempted.

Activation by CaM and calcium

CaM is a eukaryotic specific protein, so its requirement for SidJ activity prevents premature inactivation of the SidE effectors in the bacterium. Because an 'apo' structure of SidJ has remained elusive, it is difficult to pinpoint the rearrangements that occur upon CaM binding to impart catalytic competency. One study noted the N-terminal extension of SidJ (the NRD, or N-terminal regulatory domain, aa 1-133) meanders across the kinase domain and contacts the CaM N-lobe[86]. This region may be an important part of the network that stabilizes the kinase activation loop.

Because CaM is a calcium sensor, SidJ may be regulated by calcium concentrations. However, the SidJ-CaM interaction was initially identified in the presence of EGTA, but is also stable in the presence of $Ca^{2+,}$ as in the crystal structure. I have found that SidJ activity is unaffected whether CaM is prepared with an EGTA dialysis step or without, suggesting calcium binding is not required for glutamylation. In agreement with this conclusion, other studies have determined

structures of SidJ-CaM with one[86,87] or no[85] calcium ions bound, and one group found that SidJ lost only 15% of its glutamylation activity when Ca^{2+} was increased to 3mM[86]. Nonetheless, another study found that chelating cellular calcium increased SidJ activity and have proposed that local Ca^{2+} fluctuations at the LCV may regulate SidJ[85].

Role of SdjA

The role of SdjA during *Legionella* infection remains an ongoing line of investigation. Data presented here indicates SdjA can glutamylate and suppress the toxicity of SdeB, SdeC, and SidE, but not SdeA. This would explain why SdjA fails to rescue the replication in the *ΔsidJ* strain, as SdjA could not counteract SdeA toxicity. However, our genetic experiments have failed to demonstrate a role for SdjA even when SdeA and SidJ are removed. *Legionella* have a very broad host range, and the redundancy of their effector repertoire may reflect survival strategies tailored to specific host species[34]. SdeB, SdeC and SidE may simply not exert as toxic an effect on the *Acanthamoeba* host, therefore SdjA would not appear to influence bacterial replication. Another possibility is that each of the SidE members targets an overlapping but distinct repertoire of host proteins for pR-Ubiquitination, some not present in *Acanthamoeba*. In this context, SdjA may appear unnecessary. While several studies have so far identified almost 200 different host proteins as SidE substrates[29,30,48,49,54,55] none yet have distinguished between the four SidE members.

Regulation of SidE and SidJ activity during infection

Because SidE generates ubiquitinated species that are required for proper LCV biogenesis and full replication[49], SidJ must be carefully controlled so that SidE activity is tuned to appropriate levels. The mechanisms for this regulation are not entirely clear. SidE effectors and pR-Ub species are noted to localize on the bacterial phagosome minutes after infection, suggesting they are translocated and target host proteins very quickly[42,43]. Over the next four hours, SidE is removed from the developing LCV, but only when SidJ is present[43]. Others have disputed this finding[56]. Although it is now clear that SidE is inactivated by glutamylation, it is not apparent why glutamylation would cause SidE to change its subcellular localization. Other legionella or host factors may be involved, similar to the way microtubule-associated proteins are recruited to polyglutamylated tubulin to regulate cytoskeleton dynamics[88]. Levels of translocated SidE protein appears to peak 2 h after infection, while SidJ levels continue to rise for at least 5 h[43]. This suggests that temporal control of secretion or membrane localization could be a means of regulating SidJ and SidE.

The cell biological basis for the replication defect of the *ΔsidJ* strain has not been completely described. While overexpression of SidE leads to a general 'poisoning' of the Ub system[31], these effects may not be relevant in the context of an infection, when much lower concentrations of effectors are present. Studies of host

cell viability or innate immune mechanisms during infection with *ΔsidJ* bacteria are warranted.

Action of additional effector proteins to antagonize SidJ or SidE

Vertebrate polyglutamylation is reversed by a group of cytosolic carboxypeptidase enzymes (CCPs) which possess a metallopeptidase fold[89]. Although we added the carboxypeptidase inhibitor 1,10-phenanthroline to our *in vivo* glutamylation experiments, appreciable deglutamylation activity was never noted even when it was excluded. Nonetheless, *Legionella* may possess a yet-unidentified deglutamylating enzyme among its effectors to counteract SidJ.

SidJ was previously thought to function as a DUB to deconjugate both SidE-catalyzed and canonical Ub chains[56]. Since the discovery that SidJ is a glutamylase, two reports have identified effectors which appear to be the *bona fide* pR-Ub DUBs[54,90]. They are Lpg2154 (DupA) and Lpg2507 (DupB), the two small open reading frames immediately upstream of SidJ and SdjA, respectively **(Figure 3).** These effectors contain a single SidE-like PDE domain. Unlike the SidE PDE domain, DupA and DupB bind with low-nanomolar affinity to Ub, ADPR-Ub, and pR-Ub-linked peptides. Instead of transferring pR-Ub to substrates, they catalyze the removal of the pR-Ub, regenerating the serine residue of the target protein. These effectors influence the presence of pR-ubiquitinated proteins on the bacterial vacuole and add an additional layer of regulation to the *Legionella* ubiquitination cycle.

Distant SidJ homologs

SidJ homologs exist in a taxonomically diverse groups of organisms, including several bacteria and viruses which have a host-associated lifestyle: *Waddlia chondrophila* (*Chlamydiae*), an intracellular pathogen of humans and cattle, the gammaproteobacterium *Endozoicomonas*, a symbiont of marine sponges, a *Spirochaete* from termite gut, and Nile crocodilepox viruses, which inflict significant economic losses on the crocodilian industry[91] **(Figure 46).** It is tempting to speculate that these homologs may possess polyglutamylation activity and contribute to signaling at the host-pathogen or host-symbiont interface.

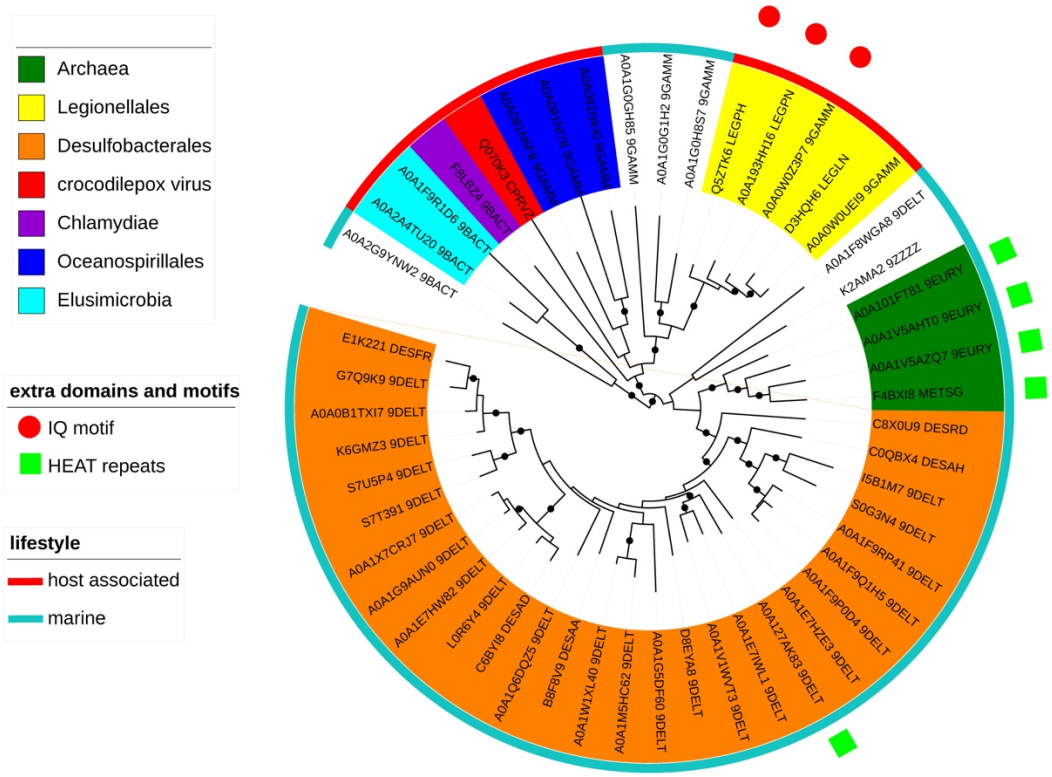


Figure 46. SidJ-like kinase domains are found in taxonomically diverse groups of organisms. A Phylogenetic tree (Maximum Likelihood) for selected SidJ homologs built from a manually corrected MAFFT multiple sequence alignment of SidJ type kinase domains, collected using JackHMMer. Coloring is by taxonomy. Red circles denote proteins containing IQ CaM-binding motifs, green squares denote proteins with HEAT repeats. Outer circle denotes lifestyles: red – host associated, cyan – free living, mostly aquatic. Black circles denote branches with bootstrap support above 0.9.

Catalytic versatility of the kinase fold

SidJ is the first example of a kinase-fold enzyme that uses the free energy of ATP to catalyze polyglutamylation. Only two protein folds are known to catalyze similar ligations within a single polypeptide: the glutamine synthetase fold, found in bacterial PUP ligases[83,92], and the ATP-grasp/glutathione synthetase fold, present in the TTL ligases which catalyze glutamylation, glycinylation, and tyrosylation[83]. Interestingly, the protein kinase fold may share an evolutionary origin with the ATP grasp fold[93]. However, the mechanism of SidJ catalysis, which generates an acyl-AMP intermediate, is distinct from the ATP-grasp enzymes, which use an acyl-phosphate intermediate[79]. SidJ and SelO illustrate that the catalytic versatility of the kinase fold extends beyond phosphotransfer. Both these enzymes lacked certain kinase catalytic motifs, technically classifying them as pseudokinases. There are over 500 protein kinases in humans, and a vast diversity of eukaryotic protein kinase-like enzymes found in nature[94]. These results suggest they should be examined for alternative activities, including ATP-dependent ligations.

TABLES

Rank (WT IP)	prot_acc	prot_desc	prot_score (Bait = SidJ WT)	prot_score (Bait = SidJ K367A)	prot_score (Bait = V)
1	P07264	LEUC (3-isopropylmalate dehydratas	1244	1703	1265
2	**P06787**	**CALM (Calmodulin)**	**1071**	**1005**	**0**
3	P40433	6P21 (6-phosphofructo-2-kinase 1)	660	1179	858
4	P07259	PYR1 (URA2)	504	814	120
5	P11484	HSP75 (Heat shock protein SSB1)	502	788	410
6	P10591	HSP71 (Heat shock protein SSA1)	437	691	449
7	P05750	RS3 (40S ribosomal protein S3)	409	633	611
8	P00560	PGK (Phosphoglycerate kinase)	368	704	152
9	P26783	RS5 (40S ribosomal protein S5)	333	633	262
10	P17076	RL8A (60S ribosomal protein L8-A)	311	393	365

Table 1. Identification of SidJ-interacting proteins in yeast. Top proteins identified by MS/MS using the MASCOT search engine from anti-Flag immunoprecipitates from yeast expressing Flag-tagged SidJ, SidJ^{K367A} and control (V). The top ten yeast proteins identified from the WT SidJ immunoprecipitates are listed, with its Protein Mascot Score (prot_Score) in each IP denoted in the last three columns.

Strain	E860 Glutamylated (PSM's)	E860 unmodifed (PSM's)
Lp02 (WT)	57 (1X Glu) 3 (2X Glu)	17
Lp02 (ΔsidJ)	0	45
Lp03 (ΔdotA)	0	21

Table 2. MS/MS analysis of Myc-SdeA co-expressed with SidJ in mammalian cells. Each row indicates a peptide containing SdeA E860 (column 1) modified by the indicated number of glutamates (column). The number of spectral counts corresponding to each peptide are given in (column).

	AMP				ADP					ATP				
Sample ID	Conc. (uM)	Avg (uM)	Stdev			Analyte/IS Ratio	Avg Ratio	Stdev			Analyte/IS Ratio	Avg Ratio	Stdev	UMP-IS
Blank	0.05	1.12	1.511441		32000	0.39	0.53	0.201753		4160000	50.49	48.17	3.271325	82400
1	0.18				37000	0.54				3710000	54.00			68700
1	0.21				38300	0.51				3870000	51.67			74900
1	0.18	0.19	0.019655		37300	0.51	0.52	0.015067		3940000	54.27	53.31	1.430885	72600
4	0.16				36800	0.49				3800000	50.80			74800
4	0.19				37800	0.53				3850000	53.77			71600
4	0.17	0.17	0.015822		41000	0.56	0.53	0.036381		4010000	55.23	53.27	2.25823	72600
2	0.41				36400	0.50				3710000	50.54			73400
2	0.51				41100	0.57				4000000	55.87			71600
2	0.30	0.41	0.105014		42100	0.55	0.54	0.040451		3890000	51.12	52.51	2.920974	76100
5	0.27				44000	0.62				3690000	51.90			71100
5	0.35				44100	0.59				3930000	52.33			75100
5	0.34	0.32	0.04236		43200	0.56	0.59	0.027894		3850000	50.20	51.47	1.128696	76700
3	3.57				44500	0.62				3670000	51.19			71700
3	3.61				40400	0.56				3680000	51.04			72100
3	3.40	3.53	0.111505		46500	0.65	0.61	0.046265		3790000	53.08	51.77	1.13876	71400

Table 3. LCMS/MS analysis of adenine nucleotides in glutamylation reactions. Sample ID corresponds to the labeling in Figure 43. UMP was added to each sample as an internal standard (last column), and Analyte/IS ratios are given for ADP and ATP. AMP concentration

Data collection			
Crystal	SeMet peak[c]	SidJ$^{\Delta_{Nt}}$/yCaM + AMP	
Space group	$P2_1$	$P2_1$	
Cell constants (Å, º)	a = 105.36, b = 104.56, c = 107.70, β = 104.02º	a = 105.49, b = 104.75, c = 107.77, β = 103.97º	
Wavelength (Å)	0.97926	0.97927	
Resolution range (Å)	46.75 – 2.50 (2.54 – 2.50)	46.84 – 2.10 (2.14 – 2.10)	
Unique reflections	75,598 (2,743)	130,441 (5,807)	
Multiplicity	20.3 (10.1)	11.5 (4.8)	
Data completeness (%)	96.3 (70.4)	98.8 (88.8)	
R_{merge} (%)[b]	11.2 (92.1)	7.5 (81.4)	
R_{pim} (%)[c]	2.4 (25.4)	2.1 (37.7)	
I/σ(I)	29.8 (1.3)	32.5 (1.5)	
$CC_{1/2}$ (highest resolution shell)	0.68	0.77	
Wilson B-value (Å^2)	31	23	
Phase determination			
Anomalous scatterers	selenium, 29 out of 32 possible sites		
Figure of merit (47.75 – 2.50 Å)	0.24		
Refinement statistics			
Crystal		SidJ$^{\Delta_{Nt}}$/yCaM + AMP	
Resolution range (Å)		46.84 – 2.10 (2.15 – 2.10)	
No. of reflections R_{work}/R_{free}		112,310/1,992 (2,703/50)	
Data completeness (%)		84.9 (29.0)	
Atoms (non-H protein/metals/nucleotide/solvent)		14,199/13/73/743	
R_{work} (%)		20.8 (21.6)	
R_{free} (%)		23.5 (23.2)	
R.m.s.d. bond length (Å)		0.005	

R.m.s.d. bond angle (°)	0.65	
Mean B-value (Å^2) (chain A/chain B/chain C/chain D/metals/nucleotide/solvent)	30/59/41/88/56/40/37	
Ramachandran plot (%) (favored/additional/disallowed)[d]	96.62/3.32/0.06	
Clashscore[d]	7.3	
Maximum likelihood coordinate error	0.23	
Missing residues	A: 59-98, 847-851. B: 1-3, 125-136, 147. C: 59-99, 843-851. D: 1-4, 125-136, 147.	.

Data for the outermost shell are given in parentheses.

[a]Bijvoet-pairs were kept separate for data processing.

[b]$R_{merge} = 100\ \Sigma_h\Sigma_i|I_{h,i} - \langle I_h\rangle|/\Sigma_h\Sigma_i\langle I_{h,i}\rangle$, where the outer sum (h) is over the unique reflections and the inner sum (i) is over the set of independent observations of each unique reflection.

[c]$R_{pim} = 100\ \Sigma_h\Sigma_i[1/(n_h - 1)]^{1/2}|I_{h,i} - \langle I_h\rangle|/\Sigma_h\Sigma_i\langle I_{h,i}\rangle$, where n_h is the number of observations of reflections **h**.

[d]As defined by the validation suite MolProbity (*33*)

Table 4. Data collection and refinement statistics, SidJ$^{\Delta NC}$-yCaM structures.

BIBLIOGRAPHY

1 Walsh, D. A., Perkins, J. P. & Krebs, E. G. An adenosine 3',5'-monophosphate-dependant protein kinase from rabbit skeletal muscle. *J Biol Chem* **243**, 3763-3765 (1968).

2 NobelPrize.org. *The Nobel Prize in Physiology or Medicine 1992*, <<https://www.nobelprize.org/prizes/medicine/1992/press-release/>> (2020).

3 Cohen, P. Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov* **1**, 309-315, doi:10.1038/nrd773 (2002).

4 Roskoski, R., Jr. Properties of FDA-approved small molecule protein kinase inhibitors: A 2020 update. *Pharmacol Res* **152**, 104609, doi:10.1016/j.phrs.2019.104609 (2020).

5 Hanks, S. K. & Hunter, T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* **9**, 576-596 (1995).

6 Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. The protein kinase complement of the human genome. *Science* **298**, 1912-1934, doi:10.1126/science.1075762 (2002).

7 Tagliabracci, V. S. *et al.* Secreted kinase phosphorylates extracellular proteins that regulate biomineralization. *Science* **336**, 1150-1153, doi:10.1126/science.1217817 (2012).

8 Nguyen, K. B. *et al.* Phosphorylation of spore coat proteins by a family of atypical protein kinases. *Proc Natl Acad Sci U S A* **113**, E3482-3491, doi:10.1073/pnas.1605917113 (2016).

9 Lopez, V. A. *et al.* A Bacterial Effector Mimics a Host HSP90 Client to Undermine Immunity. *Cell* **179**, 205-218 e221, doi:10.1016/j.cell.2019.08.020 (2019).

10 Sreelatha, A. *et al.* A Legionella effector kinase is activated by host inositol hexakisphosphate. *Journal of Biological Chemistry* **RA120.013067** (2020).

11 Sreelatha, A. *et al.* Protein AMPylation by an Evolutionarily Conserved Pseudokinase. *Cell*, doi:10.1016/j.cell.2018.08.046 (2018).

12 Eyers, P. A. & Murphy, J. M. Dawn of the dead: protein pseudokinases signal new adventures in cell biology. *Biochem Soc Trans* **41**, 969-974, doi:10.1042/BST20130115 (2013).

13 Bardwell, L. Pseudokinases: Flipping the ATP for AMPylation. *Curr Biol* **29**, R23-R25, doi:10.1016/j.cub.2018.11.018 (2019).

14 Fraser, D. W. *et al.* Legionnaires' disease: description of an epidemic of pneumonia. *N Engl J Med* **297**, 1189-1197, doi:10.1056/NEJM197712012972201 (1977).

15 McDade, J. E. *et al.* Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease. *N Engl J Med* **297**, 1197-1203, doi:10.1056/NEJM197712012972202 (1977).
16 Center for Disease Control and Preventon. *Legionella (Legionnaires' Disease and Pontiac Fever) History, Burden and Trends*, <https://www.cdc.gov/legionella/about/history.html> (2020).
17 Oliva, G., Sahr, T. & Buchrieser, C. The Life Cycle of L. pneumophila: Cellular Differentiation Is Linked to Virulence and Metabolism. *Front Cell Infect Microbiol* **8**, 3, doi:10.3389/fcimb.2018.00003 (2018).
18 Hilbi, H., Hoffmann, C. & Harrison, C. F. Legionella spp. outdoors: colonization, communication and persistence. *Environ Microbiol Rep* **3**, 286-296, doi:10.1111/j.1758-2229.2011.00247.x (2011).
19 Newton, H. J., Ang, D. K., van Driel, I. R. & Hartland, E. L. Molecular pathogenesis of infections caused by Legionella pneumophila. *Clin Microbiol Rev* **23**, 274-298, doi:10.1128/CMR.00052-09 (2010).
20 Cornejo, E., Schlaermann, P. & Mukherjee, S. How to rewire the host cell: A home improvement guide for intracellular bacteria. *J Cell Biol* **216**, 3931-3948, doi:10.1083/jcb.201701095 (2017).
21 Isberg, R. R., O'Connor, T. J. & Heidtman, M. The Legionella pneumophila replication vacuole: making a cosy niche inside host cells. *Nat Rev Microbiol* **7**, 13-24, doi:10.1038/nrmicro1967 (2009).
22 Green, E. R. & Mecsas, J. Bacterial Secretion Systems: An Overview. *Microbiol Spectr* **4**, doi:10.1128/microbiolspec.VMBF-0012-2015 (2016).
23 Ghosal, D. *et al.* Molecular architecture, polar targeting and biogenesis of the Legionella Dot/Icm T4SS. *Nat Microbiol* **4**, 1173-1182, doi:10.1038/s41564-019-0427-4 (2019).
24 Gomez-Valero, L. *et al.* More than 18,000 effectors in the Legionella genus genome provide multiple, independent combinations for replication in human cells. *Proc Natl Acad Sci U S A* **116**, 2265-2273, doi:10.1073/pnas.1808016116 (2019).
25 Burstein, D. *et al.* Genomic analysis of 38 Legionella species identifies large and diverse effector repertoires. *Nat Genet* **48**, 167-175, doi:10.1038/ng.3481 (2016).
26 Mukherjee, S. *et al.* Modulation of Rab GTPase function by a protein phosphocholine transferase. *Nature* **477**, 103-106, doi:10.1038/nature10335 (2011).
27 Chen, Y. *et al.* Structural basis for Rab1 de-AMPylation by the Legionella pneumophila effector SidD. *PLoS Pathog* **9**, e1003382, doi:10.1371/journal.ppat.1003382 (2013).
28 Neunuebel, M. R. *et al.* De-AMPylation of the small GTPase Rab1 by the pathogen Legionella pneumophila. *Science* **333**, 453-456, doi:10.1126/science.1207193 (2011).

29 Qiu, J. *et al.* Ubiquitination independent of E1 and E2 enzymes by bacterial effectors. *Nature* **533**, 120-124, doi:10.1038/nature17657 (2016).
30 Kotewicz, K. M. *et al.* A Single Legionella Effector Catalyzes a Multistep Ubiquitination Pathway to Rearrange Tubular Endoplasmic Reticulum for Replication. *Cell Host Microbe* **21**, 169-181, doi:10.1016/j.chom.2016.12.007 (2017).
31 Bhogaraju, S. *et al.* Phosphoribosylation of Ubiquitin Promotes Serine Ubiquitination and Impairs Conventional Ubiquitination. *Cell* **167**, 1636-1649 e1613, doi:10.1016/j.cell.2016.11.019 (2016).
32 Luo, Z. Q. & Isberg, R. R. Multiple substrates of the Legionella pneumophila Dot/Icm system identified by interbacterial protein transfer. *Proc Natl Acad Sci U S A* **101**, 841-846, doi:10.1073/pnas.0304916101 (2004).
33 O'Connor, T. J., Boyd, D., Dorer, M. S. & Isberg, R. R. Aggravating genetic interactions allow a solution to redundancy in a bacterial pathogen. *Science* **338**, 1440-1444, doi:10.1126/science.1229556 (2012).
34 Park, J. M., Ghosh, S. & O'Connor, T. J. Combinatorial selection in amoebal hosts drives the evolution of the human pathogen Legionella pneumophila. *Nat Microbiol* **5**, 599-609, doi:10.1038/s41564-019-0663-7 (2020).
35 O'Connor, T. J., Adepoju, Y., Boyd, D. & Isberg, R. R. Minimization of the Legionella pneumophila genome reveals chromosomal regions involved in host range expansion. *Proc Natl Acad Sci U S A* **108**, 14733-14740, doi:10.1073/pnas.1111678108 (2011).
36 Park, B. C., Reese, M. & Tagliabracci, V. S. Thinking outside of the cell: Secreted protein kinases in bacteria, parasites, and mammals. *IUBMB Life* **71**, 749-759, doi:10.1002/iub.2040 (2019).
37 Ge, J. *et al.* A Legionella type IV effector activates the NF-kappaB pathway by phosphorylating the IkappaB family of inhibitors. *Proc Natl Acad Sci U S A* **106**, 13725-13730, doi:10.1073/pnas.0907200106 (2009).
38 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* **6**, e00354-00315, doi:10.1128/mBio.00354-15 (2015).
39 Lee, P. C. & Machner, M. P. The Legionella Effector Kinase LegK7 Hijacks the Host Hippo Pathway to Promote Infection. *Cell Host Microbe* **24**, 429-438 e426, doi:10.1016/j.chom.2018.08.004 (2018).
40 Dong, N. *et al.* Modulation of membrane phosphoinositide dynamics by the phosphatidylinositide 4-kinase activity of the Legionella LepB effector. *Nat Microbiol* **2**, 16236, doi:10.1038/nmicrobiol.2016.236 (2016).

41 Liu, Y. & Luo, Z. Q. The Legionella pneumophila effector SidJ is required for efficient recruitment of endoplasmic reticulum proteins to the bacterial phagosome. *Infect Immun* **75**, 592-603, doi:10.1128/IAI.01278-06 (2007).

42 Bardill, J. P., Miller, J. L. & Vogel, J. P. IcmS-dependent translocation of SdeA into macrophages by the Legionella pneumophila type IV secretion system. *Mol Microbiol* **56**, 90-103, doi:10.1111/j.1365-2958.2005.04539.x (2005).

43 Jeong, K. C., Sexton, J. A. & Vogel, J. P. Spatiotemporal regulation of a Legionella pneumophila T4SS substrate by the metaeffector SidJ. *PLoS Pathog* **11**, e1004695, doi:10.1371/journal.ppat.1004695 (2015).

44 Havey, J. C. & Roy, C. R. Toxicity and SidJ-Mediated Suppression of Toxicity Require Distinct Regions in the SidE Family of Legionella pneumophila Effectors. *Infect Immun* **83**, 3506-3514, doi:10.1128/IAI.00497-15 (2015).

45 Tan, Y., Arnold, R. J. & Luo, Z. Q. Legionella pneumophila regulates the small GTPase Rab1 activity by reversible phosphorylcholination. *Proc Natl Acad Sci U S A* **108**, 21212-21217, doi:10.1073/pnas.1114023109 (2011).

46 Kubori, T., Shinzawa, N., Kanuka, H. & Nagai, H. Legionella metaeffector exploits host proteasome to temporally regulate cognate effector. *PLoS Pathog* **6**, e1001216, doi:10.1371/journal.ppat.1001216 (2010).

47 Urbanus, M. L. *et al.* Diverse mechanisms of metaeffector activity in an intracellular bacterial pathogen, Legionella pneumophila. *Mol Syst Biol* **12**, 893, doi:10.15252/msb.20167381 (2016).

48 Akturk, A. *et al.* Mechanism of phosphoribosyl-ubiquitination mediated by a single Legionella effector. *Nature* **557**, 729-733, doi:10.1038/s41586-018-0147-6 (2018).

49 Kalayil, S. *et al.* Insights into catalysis and function of phosphoribosyl-linked serine ubiquitination. *Nature* **557**, 734-738, doi:10.1038/s41586-018-0145-8 (2018).

50 Dong, Y. *et al.* Structural basis of ubiquitin modification by the Legionella effector SdeA. *Nature* **557**, 674-678, doi:10.1038/s41586-018-0146-7 (2018).

51 Kim, L. *et al.* Structural and Biochemical Study of the Mono-ADP-Ribosyltransferase Domain of SdeA, a Ubiquitylating/Deubiquitylating Enzyme from Legionella pneumophila. *J Mol Biol* **430**, 2843-2856, doi:10.1016/j.jmb.2018.05.043 (2018).

52 Wang, Y. *et al.* Structural Insights into Non-canonical Ubiquitination Catalyzed by SidE. *Cell* **173**, 1231-1243 e1216, doi:10.1016/j.cell.2018.04.023 (2018).

53 Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu Rev Biochem* **67**, 425-479, doi:10.1146/annurev.biochem.67.1.425 (1998).

54 Shin, D. *et al.* Regulation of Phosphoribosyl-Linked Serine Ubiquitination by Deubiquitinases DupA and DupB. *Mol Cell* **77**, 164-179 e166, doi:10.1016/j.molcel.2019.10.019 (2020).

55 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* **21**, 2031-2038, doi:10.1016/j.celrep.2017.10.088 (2017).

56 Qiu, J. *et al.* A unique deubiquitinase that deconjugates phosphoribosyl-linked protein ubiquitination. *Cell Res* **27**, 865-881, doi:10.1038/cr.2017.66 (2017).

57 Xu, D., Jaroszewski, L., Li, Z. & Godzik, A. FFAS-3D: improving fold recognition by including optimized structural features and template re-ranking. *Bioinformatics* **30**, 660-667, doi:10.1093/bioinformatics/btt578 (2014).

58 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858, doi:10.1038/nprot.2015.053 (2015).

59 Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* **430**, 2237-2243, doi:10.1016/j.jmb.2017.12.007 (2018).

60 Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389-394, doi:10.1093/nar/gkv332 (2015).

61 Finn, R. D. *et al.* HMMER web server: 2015 update. *Nucleic Acids Res* **43**, W30-38, doi:10.1093/nar/gkv397 (2015).

62 Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780, doi:10.1093/molbev/mst010 (2013).

63 Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190, doi:10.1101/gr.849004 (2004).

64 Chien, M. *et al.* The genomic sequence of the accidental pathogen Legionella pneumophila. *Science* **305**, 1966-1968, doi:10.1126/science.1099776 (2004).

65 Salomon, D. *et al.* Effectors of animal and plant pathogens use a common domain to bind host phosphoinositides. *Nat Commun* **4**, 2973, doi:10.1038/ncomms3973 (2013).

66 Salomon, D. & Sessa, G. Identification of growth inhibition phenotypes induced by expression of bacterial type III effectors in yeast. *J Vis Exp*, doi:10.3791/1865 (2010).

67 Haas, A. L., Warms, J. V. & Rose, I. A. Ubiquitin adenylate: structure and role in ubiquitin activation. *Biochemistry* **22**, 4388-4394, doi:10.1021/bi00288a007 (1983).

68 Keller, S. *et al.* High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal Chem* **84**, 5066-5073, doi:10.1021/ac3007522 (2012).

69 Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* **62**, 859-866, doi:10.1107/S0907444906019949 (2006).

70 Borek, D., Minor, W. & Otwinowski, Z. Measurement errors and their consequences in protein crystallography. *Acta Crystallogr D Biol Crystallogr* **59**, 2031-2038 (2003).

71 Otwinowski, Z., Borek, D., Majewski, W. & Minor, W. Multiparametric scaling of diffraction intensities. *Acta Crystallogr A* **59**, 228-234 (2003).

72 Borek, D., Cymborowski, M., Machius, M., Minor, W. & Otwinowski, Z. Diffraction data analysis in the presence of radiation damage. *Acta Crystallogr D Biol Crystallogr* **66**, 426-436, doi:10.1107/S0907444909040177 (2010).

73 Borek, D., Dauter, Z. & Otwinowski, Z. Identification of patterns in diffraction intensities affected by radiation exposure. *J Synchrotron Radiat* **20**, 37-48, doi:10.1107/S0909049512048807 (2013).

74 Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* **58**, 1772-1779 (2002).

75 Otwinowski, Z. in *CCP4 Study Weekend.* (eds M. Wolf, P. R. Evans, & A.G.W. Leslie) 80-86 (Daresbury Laboratory: Science & Engineering Research Council).

76 Cowtan, K. Recent developments in classical density modification. *Acta Crystallogr D Biol Crystallogr* **66**, 470-478, doi:10.1107/S090744490903947X (2010).

77 Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* **62**, 1002-1011, doi:10.1107/S0907444906022116 (2006).

78 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

79 Haas, A. L., Warms, J. V., Hershko, A. & Rose, I. A. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *J Biol Chem* **257**, 2543-2548 (1982).

80 Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545-549, doi:10.1093/nar/gkq366 (2010).

81 Eswaran, J. *et al.* Structure and functional characterization of the atypical human kinase haspin. *Proc Natl Acad Sci U S A* **106**, 20198-20203, doi:10.1073/pnas.0901989106 (2009).
82 Sreelatha, A. *et al.* Protein AMPylation by an Evolutionarily Conserved Pseudokinase. *Cell* **175**, 809-821 e819, doi:10.1016/j.cell.2018.08.046 (2018).
83 Iyer, L. M., Abhiman, S., Maxwell Burroughs, A. & Aravind, L. Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol Biosyst* **5**, 1636-1660, doi:10.1039/b917682a (2009).
84 Janke, C. *et al.* Tubulin polyglutamylase enzymes are members of the TTL domain protein family. *Science* **308**, 1758-1762, doi:10.1126/science.1113010 (2005).
85 Bhogaraju, S. *et al.* Inhibition of bacterial ubiquitin ligases by SidJ-calmodulin catalysed glutamylation. *Nature* **572**, 382-386, doi:10.1038/s41586-019-1440-8 (2019).
86 Sulpizio, A. *et al.* Protein polyglutamylation catalyzed by the bacterial calmodulin-dependent pseudokinase SidJ. *Elife* **8**, doi:10.7554/eLife.51162 (2019).
87 Gan, N. *et al.* Regulation of phosphoribosyl ubiquitination by a calmodulin-dependent glutamylase. *Nature* **572**, 387-391, doi:10.1038/s41586-019-1439-1 (2019).
88 Boucher, D., Larcher, J. C., Gros, F. & Denoulet, P. Polyglutamylation of tubulin as a progressive regulator of in vitro interactions between the microtubule-associated protein Tau and tubulin. *Biochemistry* **33**, 12471-12477, doi:10.1021/bi00207a014 (1994).
89 Rogowski, K. *et al.* A family of protein-deglutamylating enzymes associated with neurodegeneration. *Cell* **143**, 564-578, doi:10.1016/j.cell.2010.10.014 (2010).
90 Wan, M. *et al.* Deubiquitination of phosphoribosyl-ubiquitin conjugates by phosphodiesterase-domain-containing Legionella effectors. *Proc Natl Acad Sci U S A* **116**, 23518-23526, doi:10.1073/pnas.1916287116 (2019).
91 Huchzermeyer, F. W. Diseases of farmed crocodiles and ostriches. *Rev Sci Tech* **21**, 265-276, doi:10.20506/rst.21.2.1334 (2002).
92 Ozcelik, D. *et al.* Structures of Pup ligase PafA and depupylase Dop from the prokaryotic ubiquitin-like modification pathway. *Nat Commun* **3**, 1014, doi:10.1038/ncomms2009 (2012).
93 Grishin, N. V. Phosphatidylinositol phosphate kinase: a link between protein kinase and glutathione synthase folds. *J Mol Biol* **291**, 239-247, doi:10.1006/jmbi.1999.2973 (1999).

94 Kannan, N., Taylor, S. S., Zhai, Y., Venter, J. C. & Manning, G. Structural and functional diversity of the microbial kinome. *PLoS Biol* **5**, e17, doi:10.1371/journal.pbio.0050017 (2007).