

Supplementary Materials:

Supplemental Materials and Methods:

Cloning and mutagenesis. PCR products for SidC (aa. 1-542), full length SidC (aa. 1-917), and full-length SdcA (aa. 1-908) amplified from *L. pneumophila* genomic DNA were digested with BamHI and XhoI restriction enzymes and inserted into a pET28a-based vector in frame with an N-terminal His-SUMO tag (1). Single amino acid substitution of SidC was introduced by in vitro site-directed mutagenesis using oligonucleotide primer pairs containing the appropriate base changes. For mammalian expression, corresponding fragments of SidC were PCR subcloned into pEGFP-C1 vector. The plasmid pZL199 (2) inserted with wild type or C46A mutant SidC was used to complement the $\Delta sidC-sdcA$ mutant (3). PCR products for all single lysine-containing human ubiquitin mutants were amplified from pET30a-hUb containing the corresponding mutations (from Dr. Shu-bing Qian, Cornell University). PCR products for all single lysine to arginine human ubiquitin mutants were amplified from pcDNA-His-hUb containing the corresponding mutation. All ubiquitin DNA products were digested with NdeI and XhoI restriction enzymes and ligated into pET21a plasmid digested with the same enzymes. All constructs were confirmed by DNA sequencing.

Protein Expression and Purification. Tni insect cell line was used for expression of E1 proteins. The E1 protein was first affinity purified by cobalt resins (Clonetech) followed by gel filtration chromatography. For protein expression, *E. coli* Rosetta strains harboring the expression plasmids were grown in Luria-Bertani medium supplemented with 50 $\mu\text{g/ml}$

kanamycin to mid-log phase. Protein expression was induced for overnight at 18°C with 0.1 mM isopropyl-B-D-thiogalactopyranoside (IPTG). Harvested cells were resuspended in a buffer containing 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, and protease inhibitor cocktail (Roche) and were lysed by sonication. Soluble fractions were collected by centrifugation at 18,000 rpm for 20 min at 4°C and incubated with cobalt resins (Clontech) for 1 h at 4°C. Protein bound resins were extensively washed with lysis buffer. The SUMO-specific protease Ulp1 was then added to the resin slurry to release SidC from the His-SUMO tag. Eluted protein samples were further purified by FPLC size exclusion chromatography. The peak corresponding to SidC was pooled and concentrated to 10 mg/ml in a buffer containing 20 mM Tris, pH 7.7, and 200 mM NaCl, 14 mM β -mercaptoethanol. For ubiquitin preparation, bacterial cells expressing ubiquitin or its derivatives were harvested in a buffer containing 20 mM ammonium acetate (pH 5.1) and were lysed by sonication. Soluble fractions were collected by centrifugation at 18,000 rpm for 20 min at 4°C. The resulting lysates were treated with acetic acid to adjust the pH to 4.7. The solutions were centrifuged at 18,000 rpm for 10 min to remove the white flocculent precipitant. The pH of the final clear lysates was re-adjusted to 5.1 with NaOH. The supernatant was collected and purified by cation exchange HiTrap SP column (GE healthcare) with the buffer gradient from 20 mM ammonium acetate (pH 5.1) to 0.5 M ammonium acetate (pH 5.1). Ubiquitin peaks were pooled and further purified by size-exclusion chromatography in 20 mM Tris pH 8.0 and 50 mM NaCl. The expression plasmids encoding E2-25K (4), hUbcH7, hUbcH5 and hCdc34 were first purified as His-sumo tagged fusion by cobalt resins and the tagged was removed by Ulp1 protease. A size exclusion column was used for the final stage purification of these E2 enzymes.

Crystallization. Crystals were grown at room temperature by the hanging-drop vapor diffusion method by mixing 1 μ l of protein (10 mg/ml) with an equal volume of reservoir solution containing 0.1 M cacodylate pH 5.6, 7.5% PEG6000, and 10 mM DTT. Rod-shaped crystals were formed within 2-3 days. For phase determination, protein crystals were soaked in cryoprotectant (0.1 M cacodylate pH 5.6, 7.5% PEG6000, and 25% (v/v) glycerol) with the addition of 10 mM ethylmercury chloride (kind gift from Dr. Steve Ealick at Cornell University) for 10 min at room temperature.

Data collection and processing. Diffraction data sets for native protein crystals were collected at the Cornell synchrotron light source, MacCHESS beam line A1. Data set for mercury derivative SidC crystals was collected at Brookhaven National Laboratory, X4C beamline. All data sets were indexed, integrated and scaled with HKL-2000 (5). The crystals belong to space group $P2_12_12_1$ with $a = 68.64 \text{ \AA}$; $b = 134.45 \text{ \AA}$; $c = 172.68 \text{ \AA}$; $\alpha = \beta = \gamma = 90^\circ$ (Table S1). The calculated Matthews coefficient $V_m = 3.32$ and with 62.9% of solvent in the crystal and two protein molecules in an asymmetric unit (6).

Structure determination and refinement. Four mercury sites corresponding to residues C17 and C46 in both molecules were identified in the crystal using the program HKL2MAP (7). The initial phase was calculated by single isomorphous replacement with anomalous scattering (SIRAS) method and was improved by solvent flattening in HKL2MAP. The *ab initio* protein model was then built with ARP/wARP program (8). Iterative cycles of model building and refinement were carried out with the program COOT (9). and the *refmac5* program (10) in the CCP4 suite (11) to complete the final model.

In vitro deubiquitination assay. Lys-48- or Lys-63-linked polyubiquitin chains (Ub₁₋₇) were purchased from Boston Biochem. The in vitro de-ubiquitination reactions were carried out in a buffer containing 50mM Tris (pH 8.0), 50mM NaCl, 1mM EDTA, and 5 mM DTT. The reaction mixture has a final volume of 20 µl with a final concentration of 40 ng/µl ubiquitin chains and with 1 µg of the SNL domain of SidC and 1 µg of USP5 (positive control; from Boston Biochem). The reactions were stopped at the indicated time points with SDS loading dye and the samples were separated in a 12% acrylamide gel and western blotted with anti-ubiquitin (Covance).

In vitro E3 ubiquitin ligase assay. Ubiquitination assays were performed at 37°C in the presence of 50 mM Tris-HCl (pH 8.0), 5 mM MgCl₂, 0.5 mM DTT, 50 mM creatine phosphate (Sigma P7396), 3 U/ml of pyrophosphatase (Sigma I1643), 3 U/ml of creatine phosphokinase, 150 nM (or 100 nM) E1, 200 nM (or 100 nM) E2, 0.5 µM (or 200 nM) SidC and 100 µM (or 1 µM) ubiquitin. All reactions were stopped by the addition of 5X SDS-PAGE loading buffer containing 250 mM BME and analyzed by either Coomassie Brilliant blue stain or Western blot with mouse anti-ubiquitin, rabbit anti-UbcH7 (BostonBiochem) and anti-SidC.

Cell culture and transfection. HEK293T cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS (Cellgro), and 0.1% Pen/Strep (Cellgro) at 37°C in 5% CO₂ atmosphere. Cells were co-transfected with pCDNA3-HA-Ubiquitin (Addgene) and pEGFP-C1, pEGFP-C1-SidC-542 or pEGFP-C1-SidC-542 CA plasmids for 24 hours. Transfection was performed using polyethyleneimine (PEI) reagent. Cells were harvested with

1X SDS-PAGE loading buffer containing 100 mM BME. The samples were subsequently probed with mouse anti-HA (Sigma), rabbit anti-GFP or mouse anti-GAPDH.

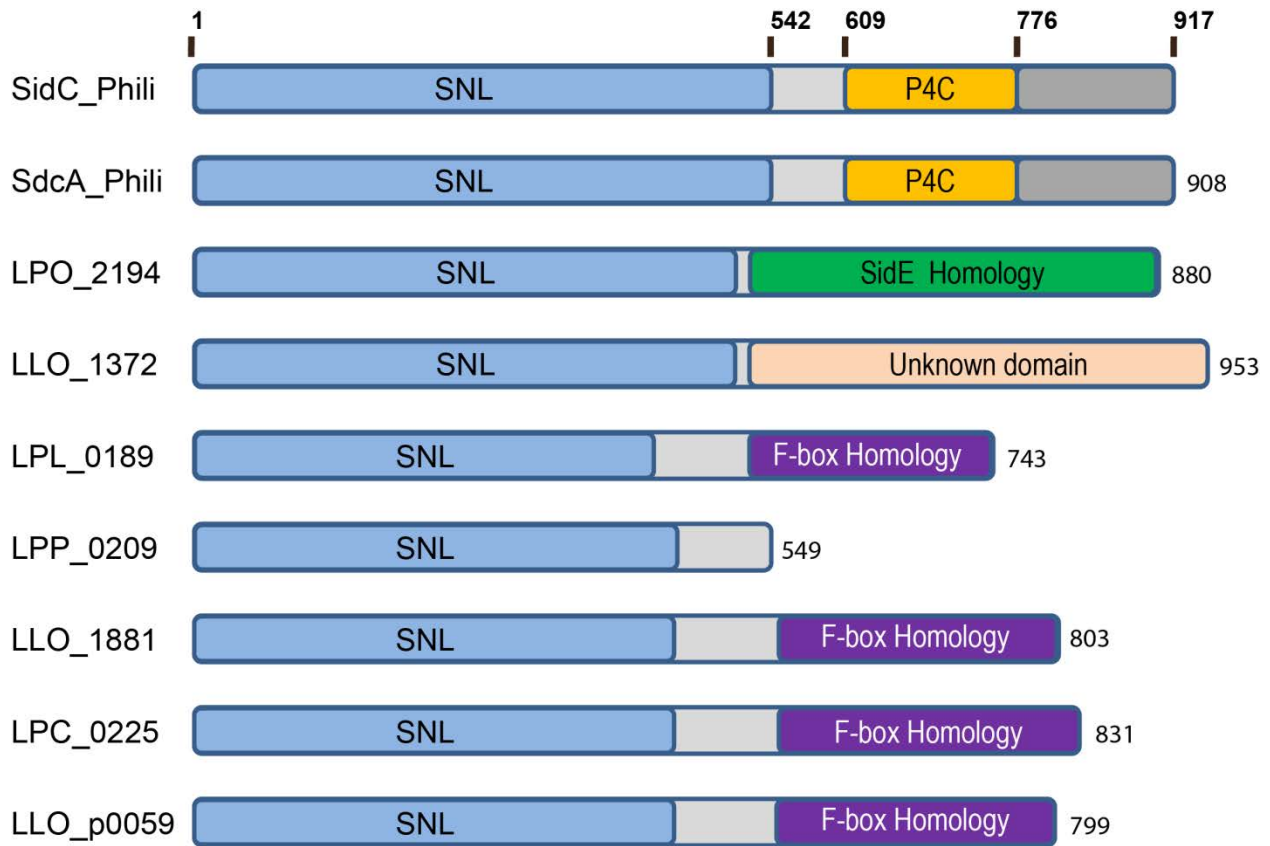
SILAC labeling and immunoprecipitation. HEK293T cells were grown for 2 weeks in complete media containing normal lysine and arginine (“light”) or [¹³C₆,¹⁵N₂] lysine and [¹³C₆,¹⁵N₄] arginine (“heavy”, Sigma) before proceeding to DNA transfection. For HA-immunoprecipitation, cells were collected in lysis buffer containing 20 mM Tris pH 7.5, 150 mM NaCl, 1% Triton and EDTA-free protease inhibitor cocktail (Roche). Soluble fractions were collected by centrifugation at 18,000 rpm for 15 min at 4°C and incubated with anti-HA affinity gel (Sigma) for 2 h at 4°C. Immunoprecipitates were eluted in 1% SDS and 100 mM Tris pH 8.0, boiled for 5 min and then precipitated with 49.9% acetone, 50% ethanol, and 0.1% acetic acid for removal of SDS. Proteins were digested with trypsin (Promega) and desalted in a C18 column. The peptides were dried in a speed-vac. The final sample was dissolved in 0.1% trifluoroacetic acid, and analyzed by LC-MS/MS using a Q-Exactive mass spectrometer (Thermo). Data analysis was carried out using a Sorcerer system (Sage-N) running Sequest for protein identification and Xpress for peptide quantitation. The protocol was adapted from Ohouo et al. (12).

Legionella strains and infection. Strains of *L. pneumophila* used include the wild type Lp02 (13), the Dot/Icm deficient Lp03 (13) and the *sidc-sdcA* mutant (3). *Dictyostelium discoideum* strain AX4 stably expressing HDEL-GFP was cultured at 21.5 °C in HL-5 medium supplemented with penicillin and streptomycin (100 U/ml), and 10 µg/ml of G418. *D. discoideum* cells were seeded onto poly-lysine coated coverslips at 2x10⁵ cells per well. After

incubation for 2 hrs at 25 °C, cells were infected with *L. pneumophila* grown to post-exponential phase for 2 hrs at an MOI of 2. To detect translocated SidC, 2×10^7 U937 cells were plated onto 10-cm petri dish 12 hours before infection. Cells were infected with post-exponential *L. pneumophila* strains for 2 hours at an MOI of 5. Cells collected by centrifugation were resuspended and lysed in 50 μ L of PBS containing 0.2% saponin. After 30 min incubation on ice, lysates were cleared by centrifugation at 10,000 g for 10 min at 4 °C. The supernatants were collected analyzed by SDS-PAGE followed by Western blot with appropriate antibodies. U937 cells were cultured and prepared for infection as described in the Materials and Methods section of our manuscript. For ubiquitin recruitment assay, 2×10^5 of U937 cells were seeded into 24-well plates with coverslips 12 hours before infection. Cells were infected with the indicated *L. pneumophila* strains for 2 hours at an MOI of 2.

Antibodies, immunostaining and Western blot. Anti-*L. pneumophila* (14) and anti-SidC (3) were described previously. Anti-tubulin antibody was purchased from (DSHB, University of Iowa). Infected *D. discoideum* samples were fixed and stained as described earlier (15). For ubiquitin immunofluorescent analyses, U937 cells were fixed and stained by standard procedures (16). Anti-*L. pneumophila* antibodies were used at a dilution of 1:20,000. Poly-ubiquitinated proteins were stained by FK1 antibody at a dilution of 1:50 followed by 555-conjugated goat anti-mouse IgM (Invitrogen, Carlsbad, CA). Recruitment of GFP-HDEL and poly-ubiquitinated proteins was examined by Olympus IX-81 fluorescence microscope. Images obtained from an Orca camera were processed with the IPlab software package (Scanalytic, Inc. Fairfax, VA). Western blots were performed following standard protocols (14); anti-SidC and anti-tubulin antibodies were used at 1:10,000; 1: 20,000 and 1:5000, respectively.

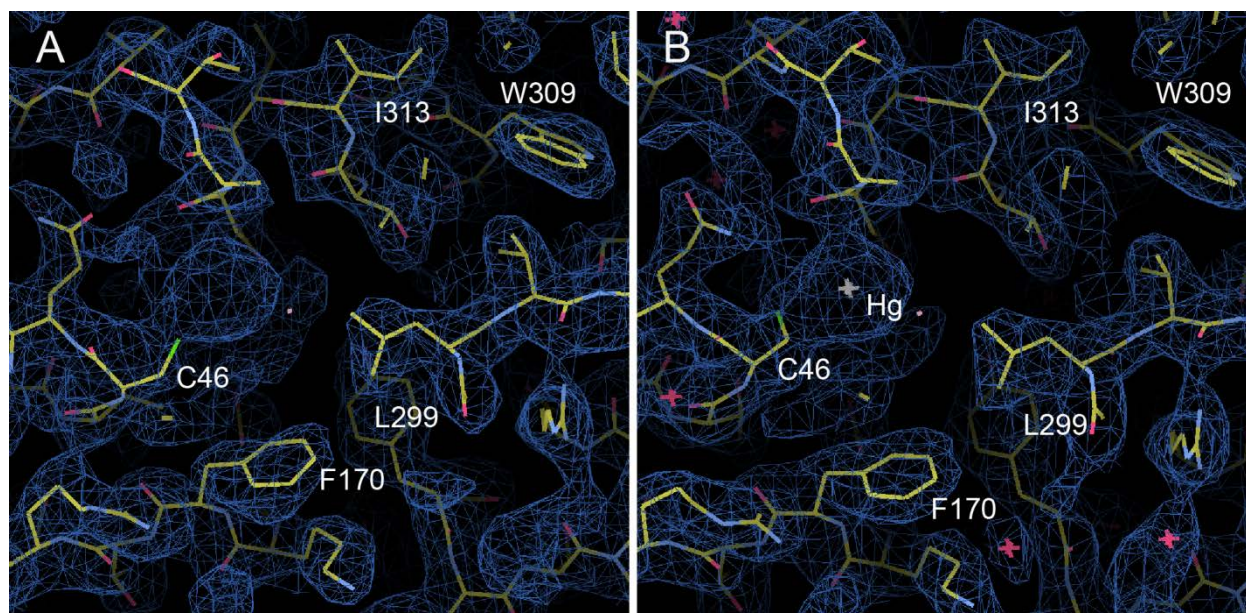
Supplemental Fig. S1



Supplemental Fig. S1. Schematic domain structures of selective members of the SidC family. SidC is a protein comprised of 917 residues. It contains an N-terminal SNL domain (**S**idC **N**-terminal **L**igase domain) and a C-terminal P4C domain (**p**hosphatidylinositol-**4**-phosphate binding of Sid**C**). The SNL domain is located at the N-terminus of all members of the SidC family. Entrez database accession numbers are as follow: SidC_Phili, gi: 52842719; SdcA_Phili, gi: 52842718; LPO_2194, gi: 397664568; LLO_1372, gi: 289164709; LPL_0189, gi: 54293148; LPP_0209, gi: 54296184; LLO_1881, gi: 289165217; LPC_0225, gi: 148358361; and LLO_p0059, gi: 308051561. The C-terminal portion of LPO_2194 contains a domain, which is conserved with the C-terminal part of SidE family proteins. The C-terminal part of LPL_0189, LLO_1881, LPC_0225, and LLO_p0059 is homologous to F-box domain-containing proteins.

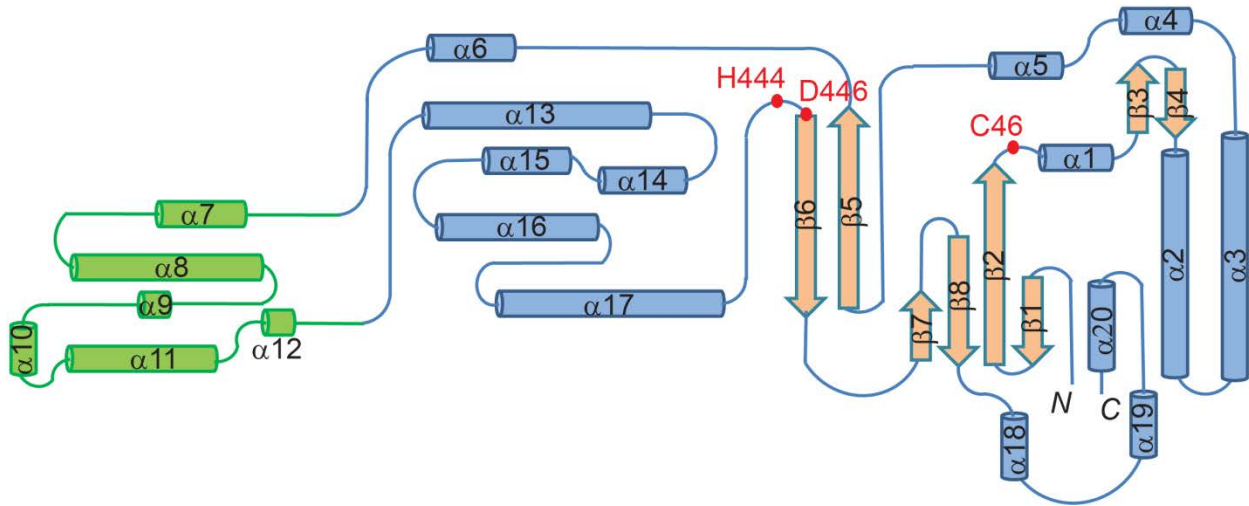
Supplemental Fig. S2. Multiple sequence alignment of the SNL domains of SidC. The sequences corresponding to the SNL domain of SidC (aa. 1-542) from different *Legionella* species were aligned by Clustal Omega (17) and colored by ALSCRIPT (18). Secondary elements are drawn below the alignment. Two conserved sequence clusters are marked with a square. The predicted catalytic triad consisting of C46, H444, and D446 are highlighted with stars. Entrez database accession numbers are listed in the legend of Fig. S1.

Supplemental Fig. S3



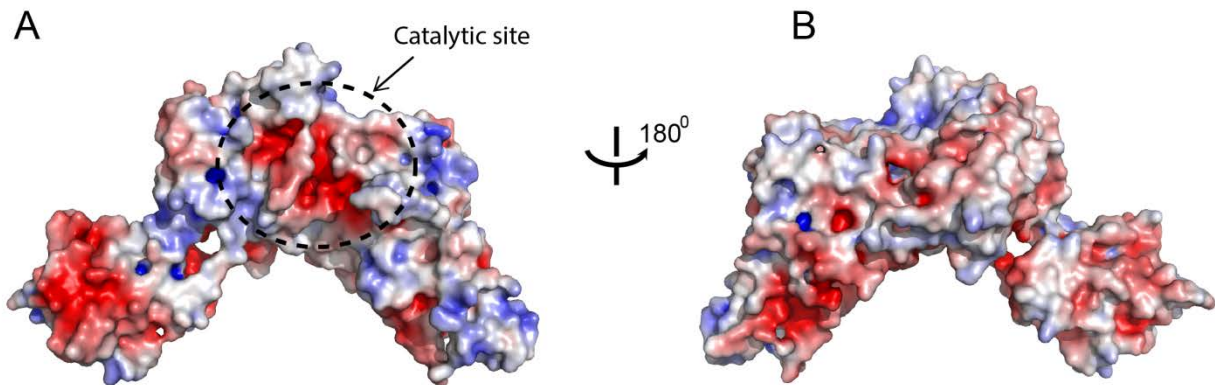
Supplemental Fig. S3. Representative experimental electron density maps. (A) Electron density map contoured at 1σ level after SIRAS phasing with the HKL2MAP program at 2.6 Å resolution. This map covers the area around one of the Hg atoms near C46. A partial polypeptide model automatically built with ARP/wARP is shown. (B) Electron density of the same area after the final cycle of refinement against the same 2.8 Å resolution data.

Supplemental Fig. S4



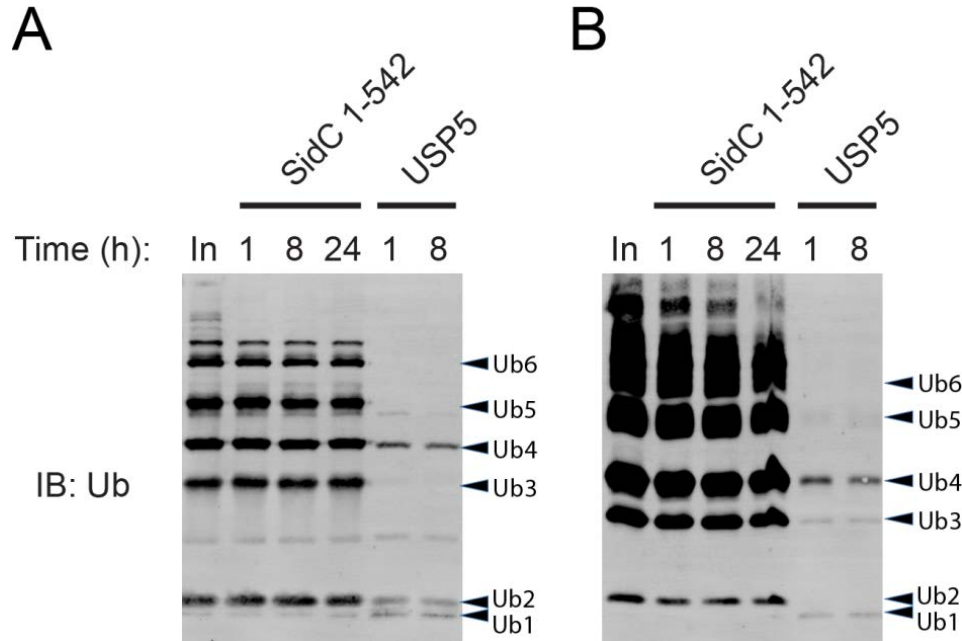
Supplemental Fig. S4. Schematic diagram of the secondary structure topologies of the SNL domain of SidC. Alpha helices are represented by cylinders and beta strands are represented by arrows. The predicted catalytic triad are labeled and marked by red circles.

Supplemental Fig. S5



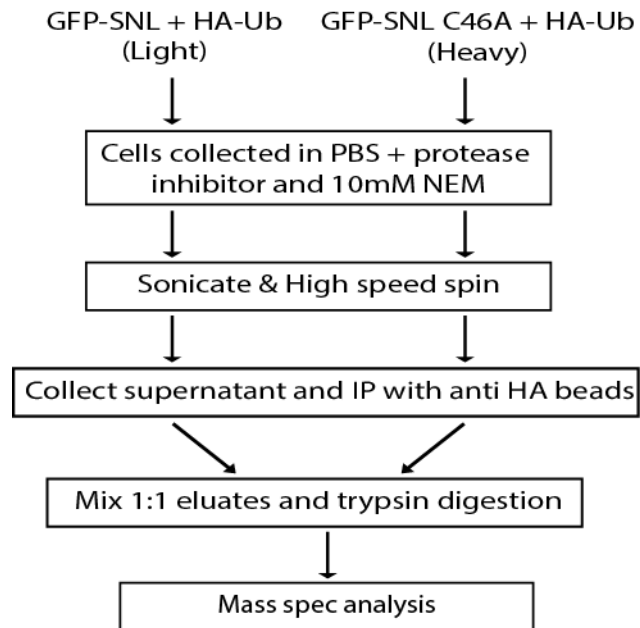
Supplemental Fig. S5. Molecular surface of the SNL domain of SidC. (A) Front view of the SNL domain. The orientation of this view is the same as that in Fig. 1B and D. The dash-lined circle indicates the area containing the catalytic C46, H444, and D446 catalytic triad. (B) Back view of the domain. The surface is colored based on electrostatic potential with positively charged region in blue (+5 kcal/electron) and negatively charged surface in red (-5 kcal/electron).

Supplemental Fig. S6



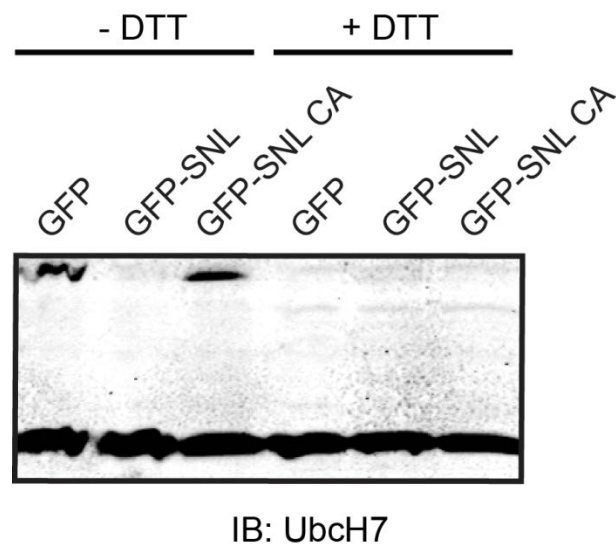
Supplemental Fig. S6. In vitro deubiquitination assay of SidC. (A) Western blot of reaction products of SidC with mixture of K48 Ub(1-6) poly-ubiquitin chains. (B) Similar reactions with K63 poly-ubiquitin chains. Incubation with wild type SidC from 1 to 24 hours. USP5 (isopeptidase T, from Boston Biochem) was used as a positive control. In: Input.

Supplemental Fig. S7



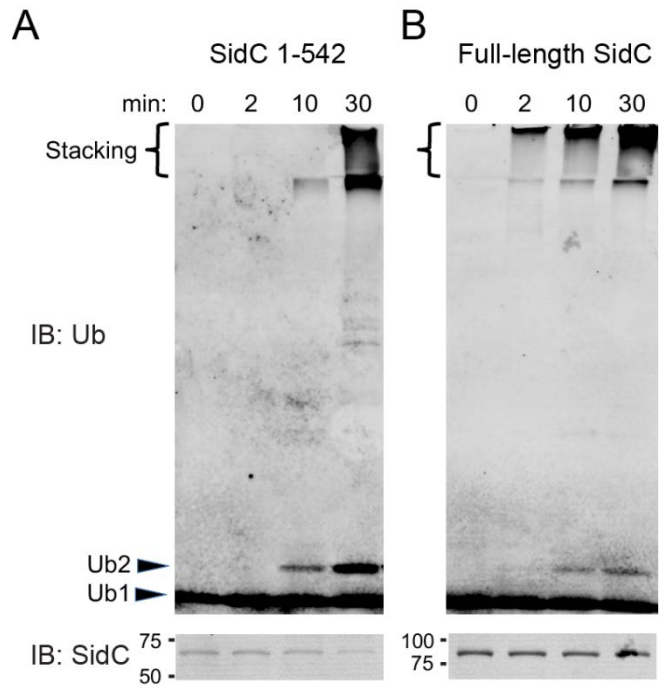
Supplemental Fig. S7. Experimental flow chart of SILAC sample preparation.

Supplemental Fig. S8



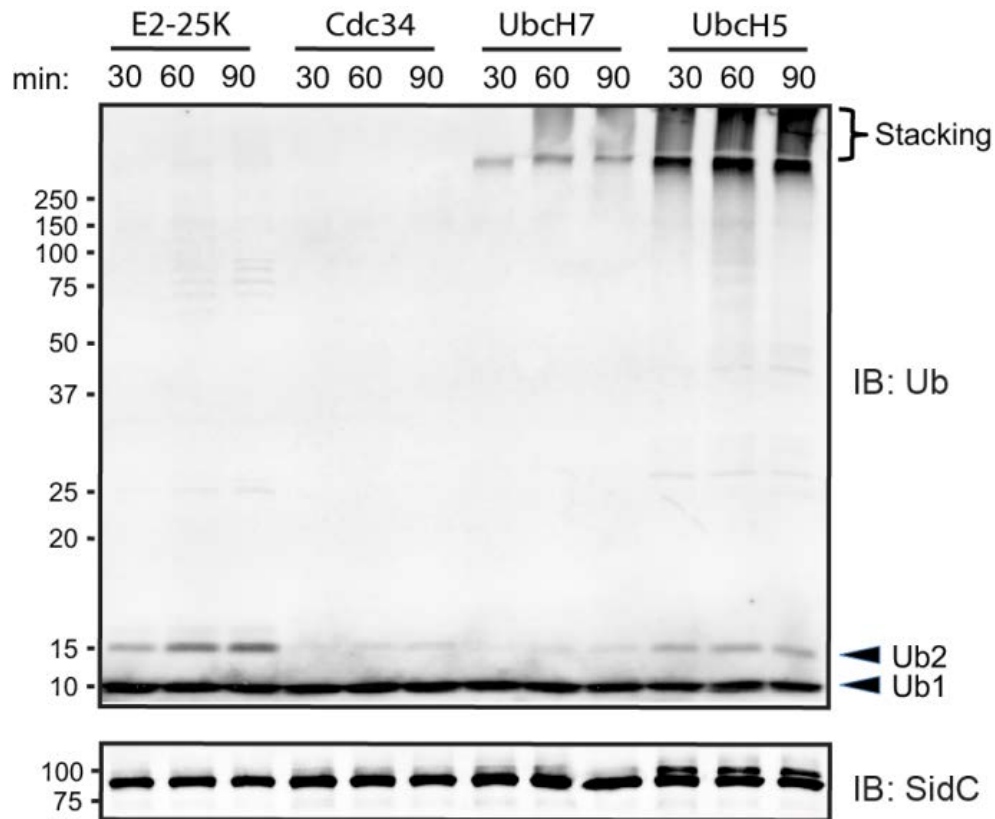
Supplemental Fig. S8. The protein levels of the ubiquitin conjugating E2 enzyme, UbcH7, are not changed in the presence of the SNL domain of SidC. 293T cells were co-transfected with HA-ubiquitin and other indicated plasmids. Cells lysates were prepared in the presence (three left lanes) or absence of DTT (three lanes on the right) and analyzed by specific UbcH7 antibody.

Supplemental Fig. S11



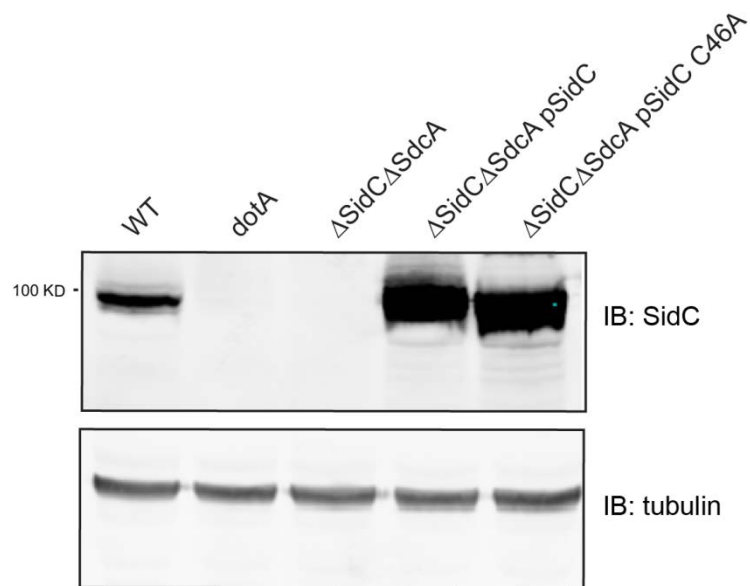
Supplemental Fig. S11. Ubiquitin ligase activity assay of the SNL domain and the full length SidC. (A) Time-dependent in vitro ubiquitination of the SNL domain of SidC (1-542). (B) Time-dependent in vitro ubiquitination of the SidC full length (1-917). (Note that the full length SidC has a comparable activity as that of its N-terminal SNL domain).

Supplemental Fig. S12



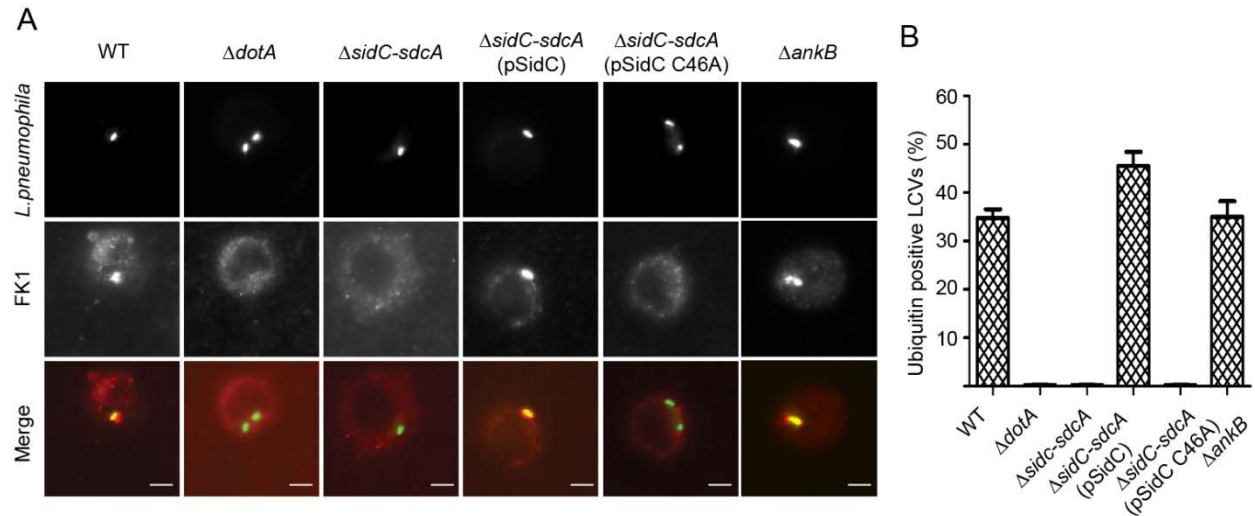
Supplemental Fig. S12. In vitro ubiquitination assays with full length SdcA (1-908) and four representative E2s: E2-25K, Cdc34, UbcH7 and UbcH5. SdcA was detected using the polyclonal antibody against full length SidC, which cross-reacts with SdcA. (Note that unlike SidC, SdcA prefers UbcH5 for efficient poly-ubiquitin chain assembly).

Supplemental Fig. S13



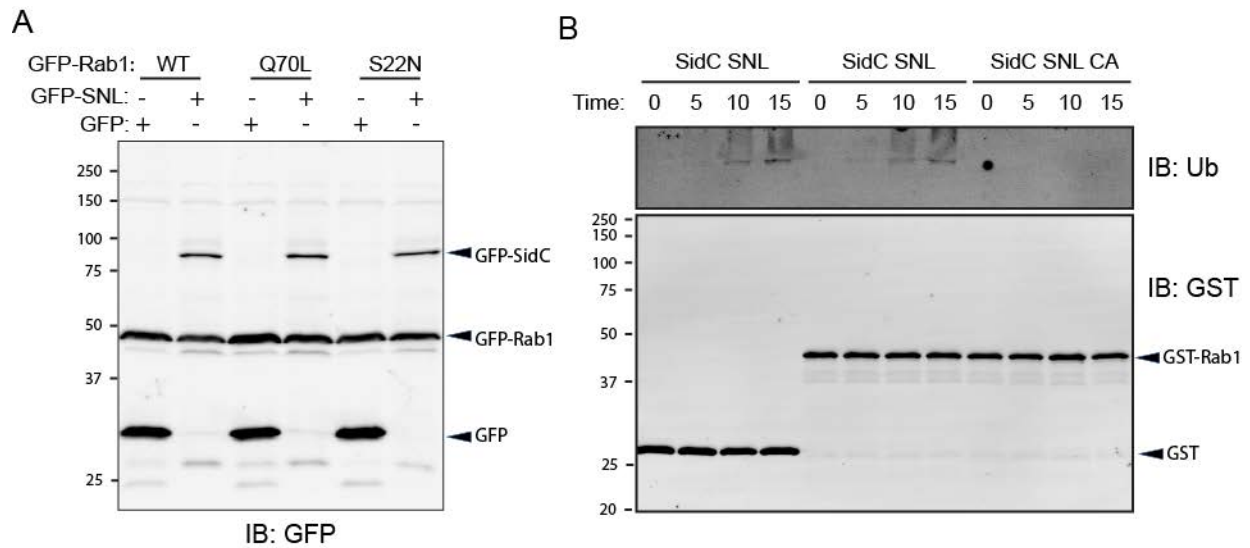
Supplemental Fig. S13. SidC translocation after *Legionella* infection. 2×10^7 U937 cells were infected with designated *Legionella* strains at moi = 5. After 2 hr infection, cells were collected and lysed by 50 μ l 0.2% saponin for 1h at on ice. After 10 min centrifugation at 12,000g, the supernatants were collected and mixed with SDS loading buffer. These samples were blotted by SidC antibody to detect the levels of SidC protein translocated into the host cells (top panel). Tubulin was used as a loading control (bottom panel).

Supplemental Fig. S14



Supplemental Fig. S14. The E3 ubiquitin ligase activity is required for the recruitment of ubiquitin conjugates to the bacterial phagosome. (A) Images show the recruitment of ubiquitinated species (red) to the LCVs in U937 cells infected with the indicated *Legionella* strains (green). Scale bars, 3 μ m. WT: *L. pneumophila* Philadelphia-1 wild type strain Lp02; *dotA*: the type IV secretion system defective strain Lp03; $\Delta sidC-sdcA$: the SidC and SdcA double deletion mutant of the Lp02 strain; $\Delta sidC-sdcA$ (pSidC) and $\Delta sidC-sdcA$ (pSidC C46A): $\Delta sidC-sdcA$ strain complemented with a plasmid expressing wild type or C46A mutant SidC; $\Delta sAnkB$: the AnkB deletion mutant of the Lp02 strain. (B) Percent of cells containing ubiquitin positive LCVs counted from three independent assays under the conditions infected with the indicated *Legionella* strains.

Supplemental Fig. S15



Supplemental Fig. S15. Rab1 is not directly ubiquitinated by SidC. (A) GFP-tagged wild type Rab1, its constitutive active form (Q70L), or dominant negative form (S22N) were co-transfected with the SNL domain of SidC GFP control in 293T cells. Cell lysate were prepared and analyzed by western blot using anti-GFP antibody. (B) Recombinant GST-tagged Rab1 was assayed for in vitro ubiquitination in the presence of Ub, E1, UbcH7, and the SNL domain of SidC. Although self-ubiquitination signals were detected (Top panel), no ubiquitination signals for either GST or GST-tagged Rab1 were detected (lower panel).

Supplemental Movie S1. Mapping of identical residues on the surface of the SNL domain.

Supplemental Table S1. Data collection and structural refinement statistics

A. Data collection statistics		
Space group	P2 ₁ 2 ₁ 2 ₁	
Cell dimensions	a = 61.53 Å, b = 133.44 Å, c = 170.25 Å, α = 90°, β = 90°, γ = 90°	
	Native	Hg
Synchrotron beam lines	MCCHESS A1	NSLS X4C
Wavelength (Å)	0.9789	0.9789
Maximum resolution (Å)	2.03	2.6
Observed reflections	400,367	651,503
Unique reflections	90,100	83,467
Completeness (%) ^a	95.2(93.0)	99.9(99.8)
$\langle I \rangle / \langle \sigma \rangle$ ^a	25.0(1.74)	31.1(10.9)
R _{sym} ^{a,b} (%)	10.7(72.4)	5.9(17.9)
Number of Hg sites		4
B. Refinement statistics		
	Native	Hg
Resolution (Å) ^a	35.88-2.03(2.055-2.03)	36-2.59(2.66-2.59)
R _{crys} / R _{free} (%) ^{a,c}	21.2/24.8(28.8/31.0)	18.9/23.1(25.4/30.3)
Rms bond length (Å)	0.018	0.016
Rms bond angles (°)	1.767	1.637
Ramachandran plot		
Most favored/Additional (%)	97.4/2.6	97.1/2.9
Generous/Disallowed (%)	0/0	0/0
^a Values in parenthesis are for the highest resolution shell.		
^b R _{sym} = $\sum_h \sum_i I_i(h) - \langle I(h) \rangle / \sum_h \sum_i I_i(h)$.		
^c R _{crys} = $\Sigma(F_{obs} - k F_{cal}) / \Sigma F_{obs} $. R _{free} was calculated for 5% of reflections randomly excluded from the refinement.		

Supplemental Table S2. A list of SILAC hits of the proteins with altered ubiquitination.

protein	experiment 1		experiment 2	
	peptide count	gmean of log2 ratio (Wt/mut)	peptide count	gmean of log 2 ratio (mut/Wt)
UBE2NL	9	-3.60566	4	3.08539
UBE2L3/UbcH7	11	-3.46593	12	6.35531
UBE2K/Ubc1	10	-2.64236	7	5.3087
UBE2S	9	-2.56502	7	3.80895
UBA1/Ube1	97	-2.53918	97	2.70132
BRAP	4	-2.37871	2	4.29822
UBE2T/HSPC150	9	-2.02673	3	2.64006
USP5/ISOT	35	-2.02175	34	3.8251
AUP1	8	-1.34807	2	2.34722

Supplemental Table S3. A list of SILAC hits of the proteins with increased ubiquitination.

protein	experiment 1		experiment 2	
	Wt = light; mut = heavy		Wt = heavy; mut = light	
	peptide count	gmean of log2 ratio (Wt/mut)	peptide count	gmean of log 2 ratio (mut/Wt)
PGAM1	8	3.95143875	27	-3.34935
GAPDH	14	3.30782	24	-3.09904
LDHA	7	1.48515571	7	-2.05435
NACA	5	1.98635	4	-1.18603
TARDBP	15	1.54404	2	-1.12998
LDHB	17	0.8431	10	-2.1497
PRDX1	10	1.260868	17	-1.5146565
PRDX2	5	1.12094	5	-1.762

Supplemental References:

1. Zuo X, *et al.* (2005) Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. *J Struct Funct Genomics* 6(2-3):103-111.
2. VanRheenen SM, Luo ZQ, O'Connor T, & Isberg RR (2006) Members of a *Legionella pneumophila* family of proteins with ExoU (phospholipase A) active sites are translocated to target cells. *Infect Immun* 74(6):3597-3606.
3. Luo ZQ & Isberg RR (2004) Multiple substrates of the *Legionella pneumophila* Dot/Icm system identified by interbacterial protein transfer. *Proc Natl Acad Sci U S A* 101(3):841-846.
4. Raasi S & Pickart CM (2005) Ubiquitin chain synthesis. *Methods Mol Biol* 301:47-55.
5. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276:307-326.
6. Matthews BW (1968) Solvent content of protein crystals. *J Mol Biol* 33(2):491-497.
7. Pape T & Schneider TR (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Cryst.* 37:843-844.
8. Langer G, Cohen SX, Lamzin VS, & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3(7):1171-1179.
9. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
10. Murshudov GN, Vagin AA, & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3):240-255.
11. Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. *Acta Cryst. D*(50):760-763.
12. Ohouo PY, Bastos de Oliveira FM, Almeida BS, & Smolka MB (2010) DNA damage signaling recruits the Rtt107-Slx4 scaffolds via Dpb11 to mediate replication stress response. *Mol Cell* 39(2):300-306.
13. Berger KH & Isberg RR (1993) Two distinct defects in intracellular growth complemented by a single genetic locus in *Legionella pneumophila*. *Mol Microbiol* 7(1):7-19.
14. Xu L, *et al.* (2010) Inhibition of host vacuolar H⁺-ATPase activity by a *Legionella pneumophila* effector. *PLoS Pathog* 6(3):e1000822.
15. Liu Y & Luo ZQ (2007) The *Legionella pneumophila* effector SidJ is required for efficient recruitment of endoplasmic reticulum proteins to the bacterial phagosome. *Infect Immun* 75(2):592-603.
16. McLean IW & Nakane PK (1974) Periodate-lysine-paraformaldehyde fixative. A new fixation for immunoelectron microscopy. *J Histochem Cytochem* 22(12):1077-1083.
17. Sievers F, *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539.
18. Barton GJ (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng* 6(1):37-40.