Legionella pneumophila effector Lem4 is a membrane-associated protein tyrosine phosphatase

Ksenia Beyrakhova1†, Lei Li‡, Caishuang Xu‡, Alla Gagarinova‡, and Miroslaw Cygler1†§

From the 1Department of Biochemistry, University of Saskatchewan, Saskatoon, Saskatchewan S7N 5E5 and the 2Department of Biochemistry, McGill University, Montreal, Quebec H3G 1Y6, Canada

Edited by Joseph M. Jez

Legionella pneumophila is a Gram-negative pathogenic bacterium that causes severe pneumonia in humans. It establishes a replicative niche called Legionella-containing vacuole (LCV) that allows bacteria to survive and replicate inside pulmonary macrophages. To hijack host cell defense systems, L. pneumophila injects over 300 effector proteins into the host cell cytosol. The Lem4 effector (lpg1101) consists of two domains: an N-terminal haloacid dehalogenase (HAD) domain with unknown function and a C-terminal phosphatidylinositol 4-phosphate-binding domain that anchors Lem4 to the membrane of early LCVs. Herein, we demonstrate that the HAD domain (Lem4-N) is structurally similar to mouse MDP-1 phosphatase and displays phosphotyrosine phosphatase activity. Substrate specificity of Lem4 was probed using a tyrosine phosphatase substrate set, which contained a selection of 360 phosphopeptides derived from human phosphorylation sites. This assay allowed us to identify a consensus pTyr-containing motif. Based on the localization of Lem4 to lysosomes and to some extent to plasma membrane when expressed in human cells, we hypothesize that this protein is involved in protein–protein interactions with an LCV or plasma membrane–associated tyrosine-phosphorylated host target.

Legionella pneumophila is the main causative agent of Legionellosis, a severe and potentially fatal pneumonia that develops as a result of bacterial colonization of alveolar macrophages in susceptible patients (1, 2). Legionellae are ubiquitous in natural and artificial aquatic environments where they survive as intraacellular parasites of freshwater amoeba (3) or as a part of biofilm communities (4). Although protozoa are primary hosts of L. pneumophila, humans may also be infected by inhaling contaminated aerosols (3).

Once engulfed by a phagocytic cell, L. pneumophila inhibits phagosome–lysosome fusion (5) and establishes a specific compartment where it can safely reside and multiply, a so-called Legionella-containing vacuole (LCV) (6). The main virulence factor of L. pneumophila that promotes conversion of phagosome to LCV is the Dot/Icm type IV secretion system (T4SS), which delivers more than 300 bacterial effector proteins into the host cell (7, 8). Effector proteins interfere with signaling pathways of the host cell, manipulate membrane trafficking, induce actin cytoskeleton rearrangements, and hamper autophagy to facilitate the intracellular survival of the pathogen (9–12).

Phosphoinositide lipids (PIs) and small Rab GTPases define the membrane-bound compartment identity in the cell; these components are specifically targeted by Legionella effectors during infection onset and LCV formation (13, 14). Within 2 h post-infection, the Legionella-containing phagosome gradually exchanges phosphatidylinositol 3-phosphate, a marker of endosomal pathway, for phosphatidylinositol 4-phosphate (PI(4)P), which is normally present on the Golgi apparatus and secretory vesicles (15). This process depends on a functional Dot/Icm T4SS system, indicating that translocated effectors control the dynamics of LCV PIs (16). For instance, one of these effectors, SidM (DrrA), is a guanosine nucleotide exchange factor for a small GTPase Rab1 that is able to activate and recruit Rab1 to LCVs (17). SidM is anchored to LCV through a specific PI(4)P-binding domain P4M that is unique to the genus Legionella (18). The P4M domain of SidM has a higher affinity to PI(4)P than other known phosphatidylinositol phosphate-binding proteins (19). This domain has been identified in 36 putative effectors across 38 sequenced Legionella genomes (8). It is typically present at the C terminus of a multidomain effector and thus targets the functional activity of N-terminal domain(s) to the surface of LCV. L. pneumophila has three effectors with a P4M domain: SidM, lpg1101(Lem4), and lpg2603 (Lem28) (20). Although the role of SidM in infection has been extensively studied, little is known about the function of the other two effectors.

Lem4 (lpg1101) is a translocated effector (20, 21), unique to L. pneumophila (8). Expression of Lem4 was demonstrated to

This work was supported in part by Canadian Institutes of Health Research Grant MOP 48370 (to M. C.). The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Figs. S1–S2 and Tables S1–S2.

The atomic coordinates and structure factors (codes 6CGK, 6CDW, and 6CGF) have been deposited in the Protein Data Bank (http://wwpdb.org/).

1 Recipient of financial support from the Canadian Institutes of Health Research Training Grant in Health Research Using Synchrotron Techniques (CIHR-THRUST).

2 To whom correspondence should be addressed: Dept. of Biochemistry, University of Saskatchewan, 107 Wiggins Rd., Saskatoon, Saskatchewan S7N 5E5, Canada. Tel.: 306-966-4361; E-mail: miroslaw.cygler@usask.ca.

3 The abbreviations used are: LCV, Legionella-containing vacuole; HAD, haloacid dehalogenase; PI(4)P, phosphatidylinositol 4-phosphate; T4SS, type IV secretion system; PI, phosphoinositide; aa, amino acid; pNPP, p-nitrophenyl phosphate; ER, endoplasmic reticulum; EGFR, epidermal growth factor receptor; PDGF-R, platelet-derived growth factor receptor; TEV, tobacco etch virus; SeMet, selenomethionine; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; AUC, area under curve; PDB, Protein Data Bank; PFA, paraformaldehyde; AP, affinity purification; FBS, fetal bovine serum; DMEM, Dulbecco’s modified Eagle’s medium.

© 2018 Beyrakhova et al. Published under exclusive license by The American Society for Biochemistry and Molecular Biology, Inc.
be regulated by the CpxRA system (22), and its translocation is affected by the IcmS–IcmW chaperon complex (23). Lem4 is a 322-residue protein that consists of two domains: a C-terminal PI(4)P-binding P4M domain that apparently drives ectopically expressed Lem4 to early LCVs in Legionella-infected mammalian cells (20), and an N-terminal haloacid dehalogenase-like (HAD-like) domain of unknown function (Fig. 1).

Here, we use a structural approach to demonstrate that the N-terminal haloacid dehalogenase domain of Lem4 is an active protein tyrosine phosphatase. Our detailed studies of Lem4 localization upon overexpression in mammalian cells show that it is targeted by P4M domain to lysosomes, which differs from the cellular localization of P4M domain of SidM (24). We used co-immunoprecipitation coupled with MS to identify potential cellular interactors was enriched with membrane proteins and proteins involved in vesicular transport. Moreover, within the list we found a significant representation of components of the Wnt signaling pathway, an evolutionarily conserved pathway that is mostly known for regulating embryonic development and tissue homeostasis but also plays a role in inflammation and host immune response to pathogens (25).

Results

Lem4-N is a structural homolog of mouse MDP-1 phosphatase

Full-length Lem4 (aa 2–322) was overexpressed in Escherichia coli strain BL21(DE3) as a fusion protein with a His6 tag and purified by metal affinity and size-exclusion chromatography. During purification, the protein tended to undergo a spontaneous degradation, yielding a fragment that was determined by MS to encompass residues 6–218. This stable fragment containing the predicted HAD-like domain of Lem4 was cloned into the same vector as the full-length protein, and the resulting construct Lem4(6–218), or Lem4-N, was successfully purified and crystallized. Crystals of native Lem4-N diffracted to 1.75 Å. Lem4-N has overall dimensions of 51 \times 38 \times 34 Å. The core domain consists of a 5-stranded parallel β-sheet (β1–β5) surrounded by five α-helices (α3–α7). A small 3_{10} helical turn (Pro-130–Glu-132) is formed at the N terminus of α6 (Fig. 1B).

A small cap domain (aa 34–71) is inserted between β1 and β3 and contains two antiparallel α-helices (α1 and α2) flanked by two 3_{10} helical turns (Gly-35–Glu-37 and Asn-68–Ile-70). The cap helices α1 and α2 form extensive contacts with the perpendicularly positioned α4. The Lem4-N fold belongs to the large HAD-like protein superfamily, containing enzymatic activities such as phosphatases, phosphonatases, P-type ATPases, β-phosphoglucomutases, phosphomannomutases, and dehalogenases, which are all involved in a variety of cellular processes ranging from amino acid biosynthesis to detoxification (26).

A search for structural homologs of Lem4-N with the DALI server (27) identified more than 30 proteins with the Z-score above 10. All these homologs belong to the haloacid dehalogenase superfamily. The top hit was the mouse magnesium–dependent phosphatase-1 (MDP-1, PDB code 1U7P, Z-score 12.1, root mean square deviation 3.3 Å for 163 Cα atoms, identity 13%; 1.6 Å for 81 Cα atoms out of 164 in 1U7P, within the core of the protein fold) (Fig. 1C) (28). The overall topology of the

Figure 1. HAD-like domain of Lem4 is homologous to MDP-1. A, schematic representation of domain organization of Lem4. B, overall fold of Lem4-N in rainbow colors from blue at the N terminus to red at the C terminus. The secondary structure elements are labeled. Mg^{2+} ion bound in the active center is shown as a green sphere. C, superposition of Lem4-N (cyan) and MDP-1 (orange). Mg^{2+} ions are colored cyan in Lem4-N and orange in MDP-1. Tungstate ion bound in the active site of MDP-1 is shown as orange sticks. The cap insertions are marked by oval dashed line.

ASBMB
**Structure and function of Legionella Lem4 effector**

**Figure 2. Lem4-N activity against a set of potential small molecule substrates.** The release of phosphate was measured using Malachite Green assay. Lem4(6–218) was added to 40 μl of solution containing 0.1 mM substrate, 50 mM sodium acetate, pH 5.5, 0.1 mM NaCl, 5 mM MgCl₂ to the final concentration of 10 μg/ml. Reactions were carried out for 30 min at 30 °C. The mean values of the two replicates are shown; the bars indicate the range of the data.

core domain of the two enzymes is very similar. Lem4-N has a more prominent loop between β3 and α4 (aa 122–132) with an additional 3_10 helical turn. The main difference between these two proteins is the α-helical cap domain of Lem4-N, which corresponds to an insert in MDP-1 consisting of two antiparallel β-strands (Fig. 1C).

**Lem4-N is an active tyrosine phosphatase**

To experimentally test whether Lem4-N displays phosphatase activity, we used the p-nitrophenyl phosphate (pNPP) as a generic substrate. For initial assessment of Lem4-N activity, we used the conditions optimized for murine MDP-1 (assay buffer with Mg^{2+} ions and pH 5.5 (29, 30)). The kinetic parameters determined for Lem4-N ($K_m = 0.74 \pm 0.037$ mm, $k_{cat} = 14.3$ s⁻¹) were in the same range as those previously reported for MDP-1 ($K_m = 1.7$ mm, rate = 4.0 s⁻¹) (30). We next tested small molecules that were identified as substrates for MDP-1, namely O-phosphotyrosine (pTyr) and some closed-ring sugars, as well as O-phosphoserine (pSer), O-phosphothreonine (pThr), several nucleotides, and phosphoinositol derivatives. The latter substrates were included because the presence of the C-terminal P(4)P-binding domain in Lem4 suggested that they could be a substrate for Lem4-N. Of these, only pNPP and pTyr were good substrates for Lem4-N (Fig. 2). Neither pSer nor pThr were dephosphorylated by Lem4-N to a significant level, suggesting that this protein is a specialized tyrosine phosphatase. Kinetic parameters of pTyr dephosphorylation by Lem4-N ($K_m = 2 \pm 0.17$ mm, $k_{cat} = 17.8$ s⁻¹) were very similar to those quantified for pNPP (see above). The optimal Hfor pTyr dephosphorylation was determined to be 5.5 (Fig. S1) and was used in all the following experiments that assessed Lem4-N activity.

**Peptide substrate profiling of Lem4-N**

To test Lem4-N activity and specificity toward pTyr in a peptide context, we used a general peptide substrate for tyrosine phosphatases, a phosphorylated nona-peptide ENDpYINASL (31). Lem4-N dephosphorylated the peptide at a lower rate than pTyr, but still significantly faster than any of the other small molecule substrates tested (data not shown).

To establish the Lem4-N substrate specificity profile, we tested its activity on a tyrosine phosphatase substrate set (JPT, Berlin, Germany), containing a selection of 360 peptides derived from known human phosphorylation sites. All the synthetic 11-mer peptides in the array contained a single phosphorylated tyrosine in the central position. The final concentration of each peptide in the reaction mixture was 10 μM. The conditions and timing of the dephosphorylation reactions were adjusted to distinguish between faster and slower dephosphorylating peptides. Peptides that showed >60% dephosphorylation (10 μM peptide per well, free phosphate concentration higher than 6 μM as assessed by malachite green assay) were used to determine the preferred sequence motif recognized by Lem4-N (Fig. 3A). Positively charged amino acids (Arg and Lys) dominated in the “+” positions (pTyr, position 0; – sites toward the N terminus; + sites toward the C terminus) and in particular in the +1 to +4 positions, with a strong preference for Lys or Arg at the +2 position. Acidic residues (Asp and Glu) were under-represented in the + positions. On the − side there was a noticeable preference for Asp in the −2 position and Asn in the −4 position (Fig. 3A). This pattern of acidic/polar side chains on one side of the Tyr and basic side chains on the other side suggests a strong electrostatic component in the substrate recognition.

Among the top 10 phosphopeptides that were almost completely dephosphorylated by Lem4-N (free phosphate concentration at the end of the reaction was close to the initial peptide concentration of 10 μM) (Table 1), only one belonged to a protein that had been previously shown to reside on LCV during *L. pneumophila* infection (macrophage colony-stimulating fac-
tor 1 receptor (CSF1R)) (32). Indeed, two CSFR1 peptides derived from the physiologically important sites of Tyr-708 and Tyr-809 were efficiently dephosphorylated by Lem4-N. Autophosphorylation of CSF1R at Tyr-708 and Tyr-809 is important for the receptor interaction with downstream signaling molecules (33).

To validate the results obtained in the peptide array, we synthesized several peptides and repeated the Lem4-N dephosphorylation reaction in a larger volume. The apparent initial rate of the reaction for the two best substrate peptides identified in the array was significantly higher than for the phosphotyrosine or for the generic phosphatase substrate ENDpYINASL (Fig. 3B). Moreover, the peptide KLNTEEpYLRVIGK that best conforms to the identified consensus motif is a better substrate of Lem4-N than IHLEKKpYVRRDSG, which has lysine residues in −1 and −2 positions instead of negatively charged residues of the consensus sequence.
Structure and function of Legionella Lem4 effector

Table 1
List of the best Lem4-N peptide substrates, as determined by the screening using tyrosine phosphatase substrate set (JPT, Berlin, Germany)

| Peptide sequence         | Uniprot ID | Proteins of origin |
|--------------------------|------------|--------------------|
| SENFDDypYMKALGV          | P02689     | Myelin P2 protein (Myp2) |
| IMDNSNypYIVKGN           | P07335     | Macrophage colony-stimulating factor 1 receptor (CSF1R) |
| EAALNeypYLRKTV           | Q75891     | Cytosolic 10-formyltetrahydrofolate dehydrogenase (AllI1) |
| KLNTEnypYLRVIKG          | P50748     | Kinetochore-associated protein1 (KNTC1) |
| TAENAepYLRVAQ1           | P00533     | Epidermal growth factor receptor (Egfr) |
| AFDPDpYWNHSLP            | Q15303     | Receptor tyrosine-protein kinase erbB-4 (ERBB4) |
| RRRGpYvNLREGP            | P12821     | Angiotensin-converting enzyme (ACE) |
| IDLRRypYLRVETI           | Q00763     | Acetyl-CoA carboxylase 2 (ACACB) |
| LEVGApypYLRDVKG          | Q08174     | Protocadherin-1 (PCDH1) |
| IHELEKlpYVRD5G           | P07335     | Macrophage colony-stimulating factor 1 receptor (CSF1R) |
| SHVNDNeypYSOPPRN         | P06127     | T cell surface glycoprotein CDS (CDS) |
| RADLAgpYNPRP            | Q93184     | Sialidase-2 (NEU2) |
| SVDSSlpYNPRSY            | Q13480     | GRB2-associated--binding protein1 (GAB1) |
| RNPLGQpYYQVRRP          | Q04910     | Adhesion G protein-coupled receptor 1 (AGRL1, Adgr11) |
| NPPPeapYPPPRAL           | Q08397     | Lysyl oxidase homolog 1 (LOXL1) |
| GLARlpypKDYYYV           | P35968     | Vascular endothelial growth factor receptor 2 (KDR) |
| EAPTTApYKTTTP           | P14317     | Hematopoietic lineage cell-specific protein (HCLS1) |
| NCTHDlpYIMIMREC         | P22607     | Fibroblast growth factor receptor3 (FGFR3) |

Active site of Lem4-N

The structural similarity to MDP-1 allowed us to propose a putative active site for Lem4-N. The HAD superfamily members that perform phosphoryl transfer have four signature motifs located in the loops, which carry the conserved catalytic residues. Lem4 contains all four of these motifs (Fig. 4A). Side chains of nucleophilic Asp-25 (loop I) and Asp-157 (loop IV), as well as the backbone carbonyl oxygen of Asp-27 (loop I), coordinate the Mg$^{2+}$ ion present in the active center. The Mg$^{2+}$ adopts an octahedral coordination with the three other ligands being water molecules. Residues Lys-134 (loop III) and Ser-96 (loop II), which are known to form H-bonds with the phosphate group, are also conserved (Fig. 4B). All these residues are located at the bottom of a deep tunnel, ~9 Å from the protein surface. The entrance to the tunnel is highly restricted by the presence of a phosphate in the active site was clear (Fig. 5A). This suggested that either the mutant Lem4-N displayed a low residual activity or that a spontaneous hydrolysis of the pTyr occurred under crystallization conditions. Interestingly, three of the four phosphate oxygens’ positions were occupied by water molecules in the native structure (Fig. 5B).

Co-crystallization of Lem4-N with phosphotyrosine and with a peptide substrate

We have shown that a peptide containing the pTyr-708 sequence from CSFIR is a substrate for Lem4-N. We attempted to capture the WT enzyme–substrate complex using either pTyr or the synthetic phosphorylated peptide IHELEKK-pYVRD5G, where underline indicates position 708. Unfortunately, these attempts were unsuccessful. We therefore created several Lem4-N mutants with compromised activity by introducing single point mutations into the main conserved motifs of the active center, namely D25N in loop I, S96A in loop II, and D157N in loop IV. Indeed, two of these mutants showed very weak residual activity (S96A) or no detectable activity (D25N) with pNPP (Fig. S2), pTyr, or pTyr-708—containing peptide (data not shown) as substrates. D157N mutant underwent partial degradation during expression and purification, which might indicate that the mutation affected protein stability.

Lem4-N(D25N) and Lem4-N(S96A) produced diffraction quality crystals in similar conditions to the WT Lem4-N. These crystals were soaked with pTyr or with the pTyr-708–containing peptide, and their diffraction data were collected. Although neither tyrosine nor the peptide were visible in the structure, the presence of a phosphate in the active site was clear (Fig. 5A). This suggested that either the mutant Lem4-N displayed a low residual activity or that a spontaneous hydrolysis of the pTyr occurred under crystallization conditions. Interestingly, three of the four phosphate oxygens’ positions were occupied by water molecules in the native structure (Fig. 5B).

As a control, we collected diffraction data from a crystal of Lem4-N(D25N) alone. We found some electron density in the active site; however, the density was flattened and indicated a planar molecule. We modeled an acetate molecule, which was present in a 200 mm concentration in the crystallization solution, into this density (Fig. 5C).

The phosphate is tightly held in the active site. One of its oxygens coordinates the Mg$^{2+}$ ion, whereas other oxygens form H-bonds with OD2 and NH of Asp-27, OD1 of Asn-25, NH of Asp-97, and NZ of Lys-134. Upon phosphate binding, the active center of Lem4-N undergoes several rearrangements: (i) the Lys-134 side chain shifts toward phosphate and forms an H-bond with its oxygen atom, and (ii) the side chain of the conserved Arg-97, stabilized by a salt bridge with Asp-161 in the absence of a substrate, becomes almost completely disordered in the enzyme–substrate complex, providing the necessary space for the substrate leaving group.

Lem4 co-localizes with lysosomes

To explore the function of Lem4 in mammalian cells, we investigated the subcellular localization of Lem4. Full-length Lem4 was C-terminally tagged with GFP and overexpressed in HEK293 and HeLa cells (Fig. 6A). Cells transfected with the GFP-encoding vector served as a control. The subcellular localization of these constructs was examined using fluorescence microscopy.
GFP control was evenly distributed in the cytosol (Fig. 6A, upper panel), whereas Lem4-GFP predominantly showed a punctate pattern both in HEK293 and HeLa cells (Fig. 6A, lower panel). Lem4 possesses a predicted P4M PI(4)P-binding domain at the C terminus, and the subcellular localization of ectopically expressed Lem4 suggested that it may bind to a specific host organelle. Markers of several organelles (cis-Golgi (GM130), endoplasmic reticulum (ER) (calnexin), and lysosomes (LAMP1, LysoTracker Red)) were used to visualize various compartments in HEK293 cells. In this experiment, Lem4-GFP co-localized with lysosomes but was also present at the plasma membrane. No co-localization with the ER or cis-Golgi was observed (Fig. 6, B and C).

Localization of Lem4 is determined by the C-terminal domain

We next inquired which Lem4 domain is responsible for the observed localization. For this purpose, the individual Lem4 domains were tagged with GFP and expressed in HEK293 cells. The HAD-like domain (aa 1–200) was evenly distributed in the cytosol but had no specific localization (Fig. 7, A and B). However, the P4M domain (aa 200–322) was localized to lysosomes and the plasma membrane, similarly to the localization of a full-length protein (Fig. 7, A and B). Similar localization was observed in the HeLa cells (data not shown) and the macrophage cell line RAW264.7 (Fig. 7C). The C-terminal P4M domain was sufficient to localize Lem4, indicating that the binding to PI(4)P is essential for localization.

Potential interactors suggested by MS

To identify potential interaction partners of Lem4 in the host cells, GFP fusion of Lem4-inactive substrate-trapping mutant (D25N) was overexpressed in HEK293 cells and subjected to co-immunoprecipitation experiments coupled with protein identification by MS. Cells overexpressing GFP alone were used as a control.
Structure and function of Legionella Lem4 effector
Significant positive ((Lem4(D25N)–GFP area under the curve (AUC))/(GFP AUC)) ratios ≥2.4, with corresponding distribution-based \( p \) values of \( \leq 1.5 \times 10^{-4} \), were obtained for 143 proteins, and an additional 117 proteins were recovered only in Lem4(D25N)–GFP samples. Of these 260 proteins, 258 identifications were mapped to Gene Ontology (GO) annotations.

The presence of the PI(4)P-binding domain and its role in determining Lem4 localization to the membranes and particularly to lysosomes is consistent with the over-representation of specific GO cellular component categories, including the “lysosomal membrane” (16 proteins, 3.7-fold enrichment, \( p \) value \( 1.18 \times 10^{-2} \)), “endocytic vesicle” (14 proteins, 4.07-fold enrichment, \( p \) value \( 5.2 \times 10^{-2} \)), “lysosome” (26 proteins, 3.2-fold enrichment, \( p \) value \( 3.5 \times 10^{-4} \)), and “endolysosome membrane” (4 proteins, 23.3-fold enrichment, \( p \) value \( 4.2 \times 10^{-2} \)) (Table S1). Similarly, phosphatase activity of Lem4 is consistent with Lem4’s associations with GO molecular function categories, including “kinase-binding” (28 proteins, \( p \) value \( 1.18 \times 10^{-2} \)) and “purine ribonucleoside triphosphate binding” (106 proteins, \( p \) value \( 1.4 \times 10^{-40} \)). Enrichments for GO Biological Process categories may suggest functions for Lem4 or its downstream effects. These include “regulation of vesicle-mediated transport” (20 proteins, \( p \) value 0.018), “regulation of cellular

*Figure 6. Lem4 co-localizes with lysosomes in transfected mammalian cells.* A, HEK293 cells transfected with a plasmid encoding Lem4(1–200)-GFP or Lem4(200–322)-GFP (green) and stained with anti-LAMP1 antibody (followed by Alexa Fluor-546 (red) secondary antibody) (left panel) or LysoTracker (red) (right panel), and then cells were fixed and imaged by confocal microscopy. Scale bars, 5 \( \mu \)m. B, quantitation of co-localization of truncated Lem4 with lysosome marker by Pearson coefficient. The Pearson correlation coefficient was determined with coloc 2 plugin of the ImageJ software and was calculated with at least 30 cells (mean and S.D. of the mean from each experiment, \( *** \), \( p < 0.001 \)). C, macrophage cell line Raw264.7 was transfected with plasmids encoding GFP alone, Lem4-GFP, Lem4(1–200)-GFP, or Lem4(200–322)-GFP, fixed with 4% (w/v) of PFA 24 h post-transfection, and imaged by confocal microscopy. Scale bars, 5 \( \mu \)m.

*Figure 7. C-terminal PI(4)P-binding domain of Lem4 is responsible for its localization in mammalian cells.* A, HEK293 cells transfected with a plasmid encoding Lem4(1–200)-GFP or Lem4(200–322)-GFP (green) and stained with anti-LAMP1 antibody (followed by Alexa Fluor-546 (red) secondary antibody) (left panel) or LysoTracker (red) (right panel), and then cells were fixed and imaged by confocal microscopy. Scale bars, 5 \( \mu \)m. B, quantitation of co-localization of truncated Lem4 with lysosome marker by Pearson coefficient. The Pearson correlation coefficient was determined with coloc 2 plugin of the ImageJ software and was calculated with at least 30 cells (mean and S.D., of the mean from each experiment, \( *** \), \( p < 0.001 \)). C, macrophage cell line Raw264.7 was transfected with plasmids encoding GFP alone, Lem4-GFP, Lem4(1–200)-GFP, or Lem4(200–322)-GFP, fixed with 4% (w/v) of PFA 24 h post-transfection, and imaged by confocal microscopy. Scale bars, 5 \( \mu \)m.
Structure and function of Legionella Lem4 effector

Table S2

models (48) and plays a role in inflammatory response in infectious pathogen's lifecycle by binding small GTPases Cdc42 and Rac1 (Table S2) (36). Interestingly, these proteins are involved in Wnt signaling that these proteins harbor protein phosphatase domains suggesting a role in Wnt signaling. Inactivation of Wnt/β-catenin signaling has been shown to increase tyrosine phosphorylation events introduced by Lem4. KEGG pathway analysis revealed that three pathways were affected in the Lem4-overexpressing cell line: Jak–Stat-signaling pathway (EGFR, JAK1, JAK2, PDGF-Rβ, SHP-1, SHP-2, STAT1, and TYK2); focal adhesion pathway (CRK, β-catenin, EGFR, PDGF-Rβ, JNK1/2/3, FAK, SHC1, and vinculin); and regulation of actin cytoskeleton (CRK, EGFR, PDGF-R, focal adhesion kinase, caveolin-1, and vinculin).

Discussion

L. pneumophila binding and uptake by the host cell leads to significant changes in phosphorylation state of multiple host proteins (38, 39). Secreted effectors of L. pneumophila were shown to interfere with mitogen-activated protein kinase (MAPK) (40) and NF-κB (nuclear factor κ-light chain enhancer of activated B cells) pathways to promote infection (41).

Until recently, no Legionella effectors with protein phosphatase activity that could directly alter the phosphorylation state of host cell proteins were known. However, a structural approach toward the determination of bacterial effector function proved to be particularly efficient in identifying protein phosphatases. Structures of effectors WipA and WipB revealed that these proteins harbor protein phosphatase domains related to the eukaryotic phosphoprotein phosphatase family (42, 43). WipB displays a Ser/Thr phosphatase activity and is targeted by its nonenzymatic C-terminal domain to lysosomes. There, it was shown to interact with the components of the lysosomal nutrient sensing (LYNUS) apparatus (42). WipA, however, despite being structurally similar to Ser/Thr phosphatases, preferably dephosphorylates phosphorylated phosphotyrosine-containing peptides and proteins in vitro (43).

Recently, the structure of another member of the HAD superfamily present in L. pneumophila, Ceg4, was reported (44). Like Lem4 (reported here), Ceg4 was also identified as a tyrosine phosphatase based on the structural information and was confirmed experimentally. The Ceg4 structure serendipi-

Figure 8. Effect of Lem4 on tyrosine phosphorylation of signal proteins in HEK293T cell line, as detected by tyrosine phosphorylation ProArray (Full Moon BioSystems). Phosphorylation levels were increased by at least 30% at 30 sites and decreased by at least 30% at four sites in the cells overexpressing Lem4-GFP. Red dashed line indicates the level of phosphorylation in the control GFP-expressing cell line.

Figure 8. Effect of Lem4 on tyrosine phosphorylation of signal proteins in HEK293T cell line, as detected by tyrosine phosphorylation ProArray (Full Moon BioSystems). Phosphorylation levels were increased by at least 30% at 30 sites and decreased by at least 30% at four sites in the cells overexpressing Lem4-GFP. Red dashed line indicates the level of phosphorylation in the control GFP-expressing cell line.

component organization” (58 proteins, p value 0.0013), and “immune system process” (65 proteins, p value $6.3 \times 10^{-5}$) (Table S1).

We then filtered the list of potential targets based on the presence of physiologically relevant Tyr phosphorylation sites that matched the substrate recognition pattern of Lem4. Almost one-half of the top 25 hits were proteins involved in membrane dynamics and cytoskeleton rearrangement, including catenin-δ-1 that contributes to regulation of cell–cell adhesion and plays a role in inflammatory response in infectious models (34), as well as the AP2 complex subunit β (responsible for clathrin-dependent endocytosis (35) and Ras GTPase-activating-like protein IQGAP1 that regulates actin cytoskeleton by binding small GTPases Cdc42 and Rac1 (Table S2) (36). Interestingly, these proteins are involved in Wnt signaling, a pathway modulated by many pathogenic bacteria as part of their infectious cycle (37). Reactome pathway analysis of the list of 258 potential Lem4 interaction partners also demonstrated significant enrichment in the following categories: “β-catenin–independent WNT signaling (R-HSA-3858494)” (20 proteins, p value $5.0 \times 10^{-12}$); “signaling by Wnt (R-HSA-195721)” (23 proteins, p value $5.0 \times 10^{-08}$); and “degradation of β-catenin by the destruction complex (R-HSA-195253)” (eight proteins, p value $1.76 \times 10^{-02}$) (Table S2).

Site-specific tyrosine phosphorylation profiling of Lem4-overexpressing mammalian cells

To further investigate the potential cellular dephosphorylation targets of Lem4, we probed the phosphorylation level of cellular proteins when Lem4 was expressed in mammalian cells. For detection of phosphorylation, we have used an array featuring 228 site-specific phosphotyrosine antibodies (Full Moon BioSystems).

We generated HEK293 cells lines stably expressing Lem4-GFP or GFP alone and compared tyrosine phosphorylation in these cell lines using the antibody array (Fig. 8). Although we detected several proteins with decreased phosphorylation levels in Lem4-expressing cells, there were significantly more proteins showing increased tyrosine phosphorylation. This indicated downstream effects of the original dephosphorylation events introduced by Lem4. KEGG pathway analysis revealed that three pathways were affected in the Lem4-overexpressing cell line: Jak–Stat-signaling pathway (EGFR, JAK1, JAK2, PDGF-Rβ, SHP-1, SHP-2, STAT1, and TYK2); focal adhesion pathway (CRK, β-catenin, EGFR, PDGF-Rβ, JNK1/2/3, FAK, SHC1, and vinculin); and regulation of actin cytoskeleton (CRK, EGFR, PDGF-R, focal adhesion kinase, caveolin-1, and vinculin).
tously revealed a tyrosine side chain from an uncleaved tag sequence of a symmetry-related molecule bound to the active site pocket. Ceg4 dephosphorylates MAPK p38 and thus attenuates the MAPK-signaling pathway.

HAD phosphatases are known to be rather promiscuous and to show a broad substrate range (45), with their specificity often being determined by the nature of a so-called cap domain (46). Interestingly, the absence of a cap domain leads to higher substrate specificity (45). Being a close structural homolog of mouse tyrosine phosphatase MDP-1 that lacks a full-sized cap domain, Lem4 also demonstrates strong specificity for phenylphosphate−containing substrates like pNPP and pTyr. Using an array of phosphorylated peptides, we detected a consensus motif that was preferably dephosphorylated by Lem4 and identified the CSF1R protein as one of the potential substrates of Lem4.

CSF1R is a receptor tyrosine kinase expressed on the surface of macrophages that is activated by the colony-stimulating factor−1 (CSF-1) and interleukin-34 (IL-34). It exerts pleiotropic effects in development, innate immunity, inflammation, and tissue repair (47). CSF-1 binding promotes rapid dimerization of CSF1R and its autophosphorylation on several tyrosine residues, including Tyr-708 and Tyr-809, which leads to activation of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase and phosphatidylinositol 3-kinase pathways (48, 49). Interestingly, CSF1R proceeds to signal from endosomes even after its CSF-1-induced internalization and thus enables prolonged activation of ERK1/2 and Akt pathways (50). CSF1R has been identified as a part of the LCV proteome in human macrophages (32) and might be targeted by L. pneumophila effectors, including Lem4, which could hamper CSF1R intracellular signaling through tyrosine dephosphorylation.

Even though CSF1R appears to be a good substrate for Lem4, the detected selectivity for peptide substrates is rather broad, indicating that Lem4 could dephosphorylate other pTyr-containing proteins that come in its proximity. Therefore, its localization and the temporal regulation of its expression is crucial for proper functioning during infection. Localization studies of ectopically expressed Lem4 in mammalian cells infected with L. pneumophila demonstrated strong co-localization with LCVs 30 min post-infection, suggesting that it acts very early during infection (20). This is consistent with the fact that Lem4 is up-regulated by the CpxRA system (22), which was shown to activate effectors involved in establishing the LCV compartment (51).

Localization of Lem4 to the host cell membranes enriched in PI(4)P is fully determined by the binding of its nonenzymatic C-terminal P4M domain. Previously, it has been reported that Lem4 predominantly localizes to the plasma membrane upon overexpression in HEK293 cells (20). Our localization experiments demonstrated that Lem4 is only partially bound to plasma membrane and predominantly associates with vesicular structures inside the cell that were also stained with an anti-LAMP1 antibody or the LysoTracker dye for acidic organelles that are used to label and track lysosomes. The P4M domain from another L. pneumophila effector, SidM, has been reported to predominantly bind the Golgi apparatus, but it was also present at the plasma membrane (24). P4M domains from Lem4 and SidM share only 30% sequence identity and 56% similarity, which could rationalize differences in their preferred localization.

The Ser/Thr phosphatase WipB, which has also been shown to be localized to the lysosomes, shares similar domain organization with Lem4. It has an N-terminal phosphatase domain and a C-terminal domain responsible for localization. The C-terminal domain of WipB (residues 399–508) is homologous to a protein-binding domain of another Legionella effector, RavJ (lpg0944, residues 238–374), with 34% sequence identity, as determined by the HHpred homology detection and structure prediction server (52). Both proteins also share a conserved WXRHH motif, which in the case of RavJ was required for interaction with components of the host septin and elongator complexes (53). WipB is targeted by its C-terminal domain to another large host macromolecular assembly, a lysosomal nutrient-sensing (LYNUS) apparatus on the surface of lysosomes. Although WipB and Lem4 are localized to the same cellular compartments, they are driven toward their host targets by different mechanisms: protein–lipid binding in the case of Lem4 and (probably) protein–protein interactions in the case of WipB.

Specific localization of Lem4, as well as tight temporal regulation of its expression during infection, may compensate for the broader specificity of this protein toward phosphotyrosine-containing macromolecular substrates. Our proteomics analysis supports the localization of Lem4 to plasma membrane and lysosomes. Additional filtering of the proteins co-purifying with Lem4 as identified by proteomics and that adhere to the Lem4 substrate pattern (Fig. 3) resulted in a list that was significantly enriched in proteins involved in cell adhesion, tight junction assembly, and cytoskeleton rearrangements. The list included three components of the Wnt/β-catenin−signaling pathway, indicating the potential role of Lem4 in regulating β-catenin levels during L. pneumophila infection.

These findings were supported by phosphotyrosine profiling of Lem4-overexpressing HEK293 cells, which revealed activation of JAK/STAT and focal adhesion signaling pathways, as well as changes in tyrosine phosphorylation levels of several components of the actin cytoskeleton regulation pathway. Although our co-immunoprecipitation experiments provide information about the mammalian proteins potentially involved in protein–protein interactions with Lem4, additional research is required to identify and validate the direct substrates of Lem4 during Legionella infection.

Materials and methods

Cloning

The gene encoding full-length Lem4 (lpg1101) (residues 2–322) was PCR-amplified from the genomic DNA of L. pneumophila strain Philadelphia 1 (ATCC 33152). Based on bioinformatics analysis, we designed and cloned a construct containing residues 15–200. Later, we identified by limited proteolysis a slightly longer fragment 6–218 that behaved well, and we cloned this fragment as well. For expression in E. coli, the PCR products were cloned into the ligation-independent cloning vector pMCSG7, according to the standard protocol (54). The
Structure and function of Legionella Lem4 effector

pMCSG7 vector encodes an N-terminal His$_6$ tag, separated from the target gene by a cleavage site for TEV protease.

For the expression of Lem4 and its truncated constructs in the mammalian cells, the lpg1101 gene or its fragments were PCR-amplified and cloned into Xhol/BamHI sites of pEGFP-N1 (Clontech) expression vector. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs (Ipswich, MA).

Site-directed mutagenesis for D25N, S96A, and D157N was performed with Q5 site-directed mutagenesis kit (New England Biolabs), according to the manufacturer’s instructions. All obtained constructs were verified by DNA sequencing.

Protein expression and purification

Plasmids carrying designed constructs were transformed into BL21(DE3)pLysS cell line. All culture media were supplemented with 100 μg/ml ampicillin and 25 μg/ml chloramphenicol. Overnight culture was started in 50 ml of LB media, supplemented with 0.4% glucose and cultivated for 16 h at 37 °C. The overnight culture was inoculated into 1 liter of Terrific Broth and grown at 37 °C with shaking at 200 rpm. The culture was induced with 0.5 mM isopropyl-1-thio-D-galactopyranoside when $A_{600}$ reached 1.5. The temperature was lowered to 18 °C; cells were incubated for another 16 h and then pelleted by centrifugation for 20 min at 5000 × g at 4 °C. For selenomethionine labeling, a shorter version of the Lem4 HAD domain, His$_2$–322) did not lead to crystals, but the construct containing residues 15–200 with a cleaved N-terminal His tag could be crystallized. This protein was used to obtain SeMet-containing crystals that were used for phase determination. The crystals of Lem4(15–200) were optimized by whisker seeding with a well solution containing 0.1 M MES, pH 6.5, 0.2 M magnesium acetate tetrahydrate, 0.65 M NaCl, and 18% (w/v) PEG 8,000 and were grown at 20 °C. These crystals diffracted to 2.0 Å resolution.

Screening the 6–218 fragment for crystallization with the Gryphon robot (Art Robbins Instruments) using the ComPAS Suite (Qiagen) identified a new set of crystallization conditions. The optimized conditions contained 0.1 M MES, pH 6.5, 0.2 M magnesium acetate tetrahydrate, and 18% (w/v) polyethylene glycol (PEG) 8,000; the crystals grew at 20 °C. These crystals diffracted to 1.75 Å resolution and were used to obtain the final structure. For data collection, the crystals were soaked in a cryo-protectant (reservoir solution supplemented with 25% glycerol) and flash-cooled in liquid nitrogen.

Crystals of Lem4(6–218)–D25N and Lem4(6–218)–S96A were obtained in 0.1 M BisTris, pH 6.5, 0.2 M magnesium acetate tetrahydrate, 0.6 M NaCl, and 22% (w/v) PEG 3350. To obtain the complex with the substrate (phosphate), the phosphorylated peptide derived from CSF1R (synthesized by GenScript (Piscataway, NJ)) was added to the drop at a final concentration of 1–2 mM.

The X-ray diffraction data collection was performed at 100 K on the 081D-CMCF beamline at the Canadian Light Source (Saskatoon, Saskatchewan, Canada) (55) and on the LRL-31-ID-D beamline at the Advanced Photon Source at Argonne National Laboratory. The native and SeMet datasets were processed and scaled with XDS (56).

Structure solution and refinement

The positions of selenium atoms for the Lem4(15–200) construct were found with SHELXD (57), and the initial model was built with Phenix AutoBuild (58). The model was then transferred to the higher resolution native data from the Lem4(6–218) crystal by molecular replacement. The structure was refined using the PHENIX software package (58) combined with manual rebuilding using Coot (59). The model contains residues Glu15–Lys206, one Mg$^{2+}$, and 175 water molecules. The native structure was a starting model for the Lem4(6–218)–D25N with and without soaked substrate. The Lem4(6–218)–D25N model contains residues Ser11–Asn210, one Mg$^{2+}$ and 127 water molecules. The structures were validated with MolProbity (60). The pertinent details of data collection and refinement statistics are shown in Table 2. The coordinates and structure factors for Lem4-N, Lem4-N(D25N)-substrate and Lem4-N(D25N) were deposited at the Protein Data Bank with codes 6CGJ, 6CGK, and 6CDW, respectively.

Phosphatase activity measurements

The assay was carried out in 96-well assay plates (Costar 9017, Corning, NY). For small molecule substrate screening, phosphatase substrates were diluted with the reaction buffer (50 mM sodium acetate, pH 5.5, 0.1 mM NaCl, 5 mM MgCl$_2$) to 0.1 mM concentration. Lem4(6–218) (10 μg/ml final concentration, or ~400 nM) was added to the wells, and the reaction was carried out for 30 min at 30 °C. The same buffer with 2 mM EDTA instead of magnesium was used in control reactions.
After the reaction ran to completion, plates were equilibrated to room temperature for 5 min, and malachite green reagent BIOMOL® Green (Enzo Life Sciences, Farmingdale, NY) was added to measure free phosphate. Plates were incubated for 20 min at room temperature, and the absorbance at 630 nm was measured. The measured values were adjusted by subtracting the values from the control wells (with EDTA). Free phosphate concentrations were calculated using a phosphate standard curve. All reactions were performed in duplicates.

For assessing substrate selectivity, we utilized the peptide library tyrosine phosphatase substrate set (JPT, Berlin, Germany) in a 384-well plate format. The assay was performed following the manufacturer’s instructions, using the buffer and Lem4 concentration as described above for small molecule substrates. The assay was performed in duplicates.

Kinetic measurements were performed in the same assay buffer in the presence of 30 nM Lem4-N for 2 min at 25 °C with various substrate concentrations (0.50 ± 10 μM for pTyr and 0.15 ± 10 μM for pNPP). Free phosphate concentration was measured by malachite green assay as described above.

Activity measurements of Lem4-N in the presence of pTyr or different peptide substrates were performed at 25 °C in the same assay buffer as described above, using the substrate concentration of 0.2 mM and Lem4-N concentration of 50 nM. De-phosphorylated substrates (free tyrosine or tyrosine-containing peptides) were detected by measuring absorbance at 280 nm in a 1-cm cuvette; reaction volume was 1 ml. Each reaction was repeated three times.

**Mammalian cell culture, transient transfection, and microscopy**

HEK293 or HeLa cells were cultured in Dulbecco’s modified Eagle’s medium (Sigma) supplemented with 10% fetal bovine serum (FBS) (Sigma) at 37 °C with 5% CO₂. DNA constructs were transfected into cells using the X-tremeGENE™ HP DNA transfection reagent (Roche Applied Science, catalog no. 06366236001) according to the manufacturer’s instructions.

Cells were fixed in 4% paraformaldehyde for 30 min and then permeabilized with 0.1% Triton X-100 in phosphate-buffered saline (PBS) for 10 min. Cells were then blocked using 5% FBS and 0.05% Tween 20 in PBS. Endogenous GM130, calnexin, or LAMP-1 were stained using anti-GM130 (BD Biosciences), anti-calnexin (BD Biosciences), or anti-LAMP1 (New England Biolabs) antibodies, respectively, followed by Alexa Fluor-546 secondary antibody (ThermoFisher Scientific).

To visualize lysosomes, the cells were stained with Lyso-Tracker® Red DND-99 (ThermoFisher Scientific) according to the manufacturer’s procedure. Microscopy images were collected on a laser scanner confocal microscope (Zeiss LSM700).

To generate cell lines that stably express GFP or Lem4-GFP, the HEK293 cells were transfected with pEGFP or pEGFP-Lem4 plasmids. The cells were selected and maintained in DMEM containing 200 μg/ml G418 (Geneticin, Sigma).

**Mass spectrometry**

Lem4 D25N-GFP was overexpressed in mammalian cells as described above, and cells were harvested and lysed by sonication in the PBS buffer containing PhosSTOP Phosphatase Inhibitor Mixture (MilliporeSigma, Darmstadt Germany) and 2 mM EDTA. Protein complexes were cross-linked with dithiobis(succinimidyl propionate) and used for affinity purification (AP) followed by mass spectrometry (MS) essentially as described previously (61) except for the following: for AP, Miltenyi Biotec anti-GFP beads were used with binding buffer (25 mM Tris-HCl, 10 mM HEPES-KOH, 500 mM NaCl, 1 mM imidazole, 2.5 mM EDTA, 2.5 mM EGTA, 0.5% Triton X-100, 0.5% Nonidet P-40) containing protease and phosphatase inhibitors. Following binding, cells were washed twice with...
Structure and function of Legionella Lem4 effector

wash buffer 1 (50 mM Tris-HCl, pH 8, 500 mM NaCl; 1 mM imidazole, 5 mM EDTA, 5 mM EGTA, 0.1% SDS, 0.5% sodium deoxycholate, 1% Nonidet P-40), then twice with wash buffer 2 (20 mM HEPES-KOH, pH 8, 300 mM NaCl, 20 mM imidazole, 5 mM EDTA, 5 mM EGTA, 1% Triton X-100), and twice with 50 mM ammonium bicarbonate. Control (GFP) and experimental (Lem4-D25N-GFP) AP-MS were independently performed five times.

The Orbitrap Elite raw files were searched against Homo sapiens proteins (taxonomy ID 9606; 20,983 Uniprot entries) and the Lem4-GFP sequence using the SequestHT algorithm in Proteome Discoverer (version 2.1.1.21; ThermoFisher Scientific). MS and MS/MS tolerances were set at 10 ppm and 0.6 Da, respectively. Search criteria included cysteine carbamidomethylation (+57.0214 Da), methionine oxidation (+15.9949 Da), and serine, threonine, and tyrosine phosphorylation (+79.996 Da). Searches were performed with a full tryptic digestion setting and allowed a maximum of two missed cleavages. Percolator node (Proteome Discoverer Version 2.1.1.21; ThermoFisher Scientific) was used to filter the data with maximum ΔCN of 0.05, false discovery rate of 0.01, and q-value–based validation.

Proteins, present only in the Lem4-D25N-GFP samples, were identified by selecting proteins with two or more peptides in each sample and eliminating proteins present in the control samples. In addition, the AUC for each protein across samples was normalized within each sample. A ratio of experimental over control AUC for each protein ((Lem4-D25N-GFP AUC for the protein)/(GFP AUC for the protein)) was calculated. A statistical over-representation test was performed using the PANTHER program (63) version 12.0 with default settings.

Phospho-specific protein microarray analysis

Tyrosine phosphorylation ProArray was purchased from Full Moon Biosystems (Sunnyvale, CA). HEK293 cell lines stably expressing either GFP (control) or Lem4-GFP were used in the experiment. Following the manufacturer’s instructions, 100 μg of cell lysate was labeled with biotin and conjugated with the array antibodies using the Array Assay Kit (Full Moon Biosystems). The conjugated labeled protein was detected using Alexa Fluor-555 streptavidin (Invitrogen). Fluorescence intensity scanning service was provided by Full Moon Biosystems. For data analysis, the intensity of the array was measured using GenePix Pro 6.0 software (Axon Instruments). The normalized data for each array were computed as follows: normalized data = (average signal intensity of replicate spots)/(median signal of the average signal intensity) for all antibodies on the array. The normalized data were then used to determine the fold change between the control (GFP) and Lem4-GFP samples.

Acknowledgments—We acknowledge the Protein Characterization and Crystallization Facility, College of Medicine, University of Saskatchewan, for access to the crystallization robot and the 08ID-1 beamline at the Canadian Light Source for diffraction data collection. The Canadian Light Source is supported by the Canada Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the Canadian Institutes of Health Research. This research used resources of the Advanced Photon Source, a United States Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the Lilly Research Laboratories Collaborative Access Team (LRL-CAT) beamline at Sector 31 of the Advanced Photon Source was provided by Eli Lilly, which operates the facility.

References

1. Fields, B. S., Benson, R. F., and Besser, R. E. (2002) Legionella and Legionnaires’ disease: 25 years of investigation. Clin. Microbiol. Rev. 15, 506–526 CrossRef Medline
2. Newton, H. J., Ang, D. K., van Driel, I. R., and Hartland, E. L. (2010) Molecular pathogenesis of infections caused by Legionella pneumophila. Clin. Microbiol. Rev. 23, 274–298 CrossRef Medline
3. Fields, B. S. (1996) The molecular ecology of legionellae. Trends Microbiol. 4, 286–290 CrossRef Medline
4. Declerck, P. (2010) Biofilms: the environmental playground of Legionella pneumophila. Environ. Microbiol. 12, 557–566 CrossRef Medline
5. Horwitz, M. A. (1983) The Legionnaires’ disease bacterium (Legionella pneumophila) inhibits phagosome–lysosome fusion in human monocytes. J. Exp. Med. 158, 2108–2126 CrossRef Medline
6. So, E. C., Matthes, C., Tate, E. W., Frankel, G., and Schroeder, G. N. (2015) Creating a customized intracellular niche: subversion of host cell signaling by Legionella type IV secretion system effectors. Can. J. Microbiol. 61, 617–635 CrossRef Medline
7. Gomez-Valero, L., Rusniok, C., Cazalet, C., and Buchrieser, C. (2011) Comparative and functional genomics of Legionella identified eukaryotic-like proteins as key players in host-pathogen interactions. Front. Microbiol. 2, 208 CrossRef Medline
8. Burstein, D., Amaro, F., Zusman, T., Lifshitz, Z., Cohen, O., Gilbert, J. A., Pupko, T., Shuman, H. A., and Segal, G. (2016) Genomic analysis of 38 Legionella species identifies large and diverse effector repertoires. Nat. Genet. 48, 167–175 CrossRef Medline
9. Hubber, A., and Roy, C. R. (2010) Modulation of host cell function by Legionella pneumophila type IV effectors. Annu. Rev. Cell Dev. Biol. 26, 261–283 CrossRef Medline
10. de Felice, K. S., Glover, R. T., Charpentier, X., Anderson, O. R., Reyes, M., Pericone, C. D., and Shuman, H. A. (2008) Legionella eukaryotic-like type IV substrates interfere with organelle trafficking. PLoS Pathog. 4, e1000117 CrossRef Medline
11. Choy, A., Dancourt, J., Mugo, B., O’Connor, T. J., Isberg, R. R., Melia, T. J., and Roy, C. R. (2012) Legionella effector RavZ inhibits host autophagy through irreversible Atg8 deconjugation. Science 338, 1072–1076 CrossRef Medline
12. Liu, Y., Zhu, W., Tan, Y., Nakayasu, E. S., Staiger, C. J., and Luo, Z.-Q. (2017) A Legionella effector disrupts host cytoskeletal structure by cleaving actin. PLoS Pathog. 13, e1006186 CrossRef Medline
13. Hilbi, H., Weber, S., and Finsel, I. (2011) Anchors for effectors: subversion of phosphoinositide lipids by Legionella. Front. Microbiol. 2, 91 Medline
14. Haneburger, I., and Hilbi, H. (2013) Phosphoinositide lipids and the Legionella pathogen vacuole. Curr. Top. Microbiol. Immunol. 376, 155–173 Medline
15. Di Paolo, G., and De Camilli, P. (2006) Phosphoinositides in cell regulation and membrane dynamics. Nature 443, 651–657 CrossRef Medline

Author contributions—K. B. and M. C. conceptualization; K. B. and A. G. data curation; K. B. and L. L. validation; K. B., L. L., C. X., and A. G. investigation; K. B. writing–original draft; L. L., C. X., and M. C. methodology; M. C. supervision; M. C. funding acquisition; M. C. writing–review and editing.
34. Hu, G. (2012) p120-Catenin: a novel regulator of innate immunity and inflammation. Crit. Rev. Immunol. 32, 127–138 CrossRef Medline

35. Lau, A. W., and Chou, M. M. (2008) The adaptor complex AP-2 regulates post-endocytic trafficking through the non-clathrin Arf6-dependent endocytic pathway. J. Cell Sci. 121, 4008–4017 CrossRef Medline

36. Briggs, M. W., and Sacks, D. B. (2003) IOGAP1 as signal integrator: Ca2+, calmodulin, Cdc42 and the cytoktoskeleton. FEBS Lett. 542, 7–11 CrossRef Medline

37. Gagnaire, A., Nadel, B., Raoul, D., Neefjes, J., and Gorvel, J.-P. (2017) Collateral damage: insights into bacterial mechanisms that predispose host cells to cancer. Nat. Rev. Microbiol. 15, 109–128 CrossRef Medline

38. Shin, S., Case, C. L., Archer, K. A., Nogueira, C. V., Kobayashi, K. S., Flavell, R. A., Roy, C. R., and Zamboni, D. S. (2008) Type IV secretion-dependent activation of host MAP kinases induces an increased proinflammatory cytokine response to Legionella pneumophila. PLoS Pathog. 4, e1000220 CrossRef Medline

39. Hempstead, A. D., and Isberg, R. R. (2013) Host signal transduction and protein kinases implicated in Legionella infection. Curr. Top. Microbiol. Immunol. 376, 249–269 Medline

40. Fontana, M. F., Shin, S., and Vance, R. E. (2012) Activation of host mitogen-activated protein kinases by secreted Legionella pneumophila effectors that inhibit host protein translation. Infect. Immun. 80, 3570–3575 CrossRef

41. Losick, V. P., Haenssler, E., Moy, M. Y., and Isberg, R. R. (2010) LmbA is a Legionella pneumophila activator of NF-κB. Cell. Microbiol. 12, 1083–1097 CrossRef Medline

42. Prevost, M. S., Pinotis, N., Dumoux, M., Hayward, R. D., and Waksman, G. (2017) The Legionella effector WipB is a translocaled Ser/Thr phosphatase that targets the host lysosomal nutrient sensing machinery. Sci. Rep. 7, 9450 CrossRef Medline

43. Pinotis, N., and Waksman, G. (2017) Structure of the WipA protein reveals a novel tyrosine protein phosphatase effector from Legionella pneumophila. J. Biol. Chem. 292, 9240–9251 CrossRef Medline

44. Quale, A. T., Stogios, P. I., Egorova, O., Evdokimova, E., Vallee, D., Nocek, B., Kompella, P. S., Peaisovich, S., Yakunin, A. F., Ensminger, A. W., and Savchenko, A. (2018) The Legionella pneumophila effector Ceg1 is a phosphotyrosine phosphatase that attenuates activation of eukaryotic MAPK pathways. J. Biol. Chem. 293, 3307–3320 CrossRef Medline

45. Huang, H., Pandya, C., Liu, C., Al-Obaiidi, N. F., Wang, M., Zheng, L., Toews Keating, S., Aono, M., Love, J. D., Evans, B., Seidel, R. D., Hillier, B. S., Garforth, S. J., Almo, S. C., Mariano, P. S., et al. (2015) Panoramic view of a superfamily of phosphatases through substrate profiling. Proc. Natl. Acad. Sci. U.S.A. 112, E1974–E1983 CrossRef Medline

46. Lu, Z., Wang, L., Dunaway-Mariano, D., and Allen, K. N. (2009) Structure-function analysis of 2-keto-3-deoxy-d-glycero-d-galactononate-9-phosphate phosphatase defines specificity elements in type C0 haloalkanoate dehalogenase family members. J. Biol. Chem. 284, 1224–1233 CrossRef Medline

47. Stanley, E. R., and Chitu, V. (2014) CSF-1 receptor signaling in myeloid cells. Cold Spring Harb. Perspect. Biol. 6, a021857 CrossRef Medline

48. Yu, W., Chen, J., Xiong, Y., Pixley, F. J., Yeung, Y. G., and Stanley, E. R. (2012) Macrophage proliferation is regulated through CSF-1 receptor tyrosines 544, 559, and 807. J. Biol. Chem. 287, 13694–13704 CrossRef Medline

49. Baccarini, M., Li, W., Dello Sbarba, P., and Stanley, E. R. (1991) Increased phosphorylation of the colony stimulating factor-1 receptor following transmembrane signaling. Receptor 1, 243–259 Medline

50. Huhn, J., Kwa, M. Q., Cook, A. D., Hamilton, J. A., and Scholz, G. M. (2012) CSF-1 receptor signalling from endosomes mediates the sustained activation of Erk1/2 and Akt in macrophages. Cell. Signal. 24, 1753–1761 CrossRef Medline

51. Feldheim, Y. S., Zusman, T., Speiser, Y., and Segal, G. (2016) The Machinery at endoplasmic reticulum-plasma membrane contact sites contributes to spatial regulation of multiple Legionella effector proteins. PLoS Pathog. 10, e1004222 CrossRef Medline

52. Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., Kühler, J., Lozaicj, M., Gable, F., Söding, J., Lupas, A. N., and Alva, V. (2018) A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J. Mol. Biol. 430, 2237–2243 Medline
Structure and function of Legionella Lem4 effector

53. Urbanus, M. L., Quaile, A. T., Stogios, P. J., Morar, M., Rao, C., Di Leo, R., Evdokimova, E., Lam, M., Oatway, C., Cuff, M. E., Osipiuk, J., Michalska, K., Nocek, B. P., Taipale, M., Savchenko, A., and Ensminger, A. W. (2016) Diverse mechanisms of metaeffector activity in an intracellular bacterial pathogen, Legionella pneumophila. Mol. Syst. Biol. 12, 893 CrossRef Medline

54. Stols, L., Gu, M., Dieckman, L., Raffen, R., Collart, F. R., and Donnelly, M. I. (2002) A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. Protein Expr. Purif. 25, 8–15 CrossRef Medline

55. Grochulski, P., Fodje, M. N., Gorin, J., Labiuk, S. L., and Berg, R. (2011) Beamline 08ID-1, the prime beamline of the Canadian Macromolecular Crystallography Facility. J. Synchrotron Radiat. 18, 681–684 CrossRef Medline

56. Kabsch, W. (2010) XDS. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132 CrossRef Medline

57. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. 58, 1772–1779 CrossRef Medline

58. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L. W., Jain, S., Kapral, G. J., Grosse Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., et al. (2011) The Phenix software for automated determination of macromolecular structures. Methods 55, 94–106 CrossRef Medline

59. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 CrossRef Medline

60. Chen, V. B., Arendall, W. B., 3rd., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21 CrossRef Medline

61. Li, J., Ma, Z., Shi, M., Malty, R. H., Aoki, H., Minic, Z., Phanse, S., Jin, K., Wall, D. P., Zhang, Z., Urban, A. E., Hallmayer, J., Babu, M., and Snyder, M. (2015) Identification of human neuronal protein complexes reveals biochemical activities and convergent mechanisms of action in autism spectrum disorders. Cell Syst. 1, 361–374 CrossRef Medline

62. Babu, M., Arnold, R., Bundalovic-Torma, C., Gagarinova, A., Wong, K. S., Kumar, A., Stewart, G., Samanfar, B., Aoki, H., Wagih, O., Vlasblom, J., Phanse, S., Lad, K., Yeou Hsiung Yu, A., Graham, C., et al. (2014) Quantitative genome-wide genetic interaction screens reveal global epistatic relationships of protein complexes in Escherichia coli. PLoS Genet. 10, e1004120 CrossRef Medline

63. Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., and Thomas, P. D. (2017) PANTHER version 11: expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. Nucleic Acids Res. 45, D183–D189 CrossRef Medline

64. O’Shea, J. P., Chou, M. F., Quader, S. A., Ryan, J. K., Church, G. M., and Schwartz, D. (2013) pLogo: a probabilistic approach to visualizing sequence motifs. Nat. Methods 10, 1211–1212 CrossRef Medline

65. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. 60, 2256–2268 CrossRef Medline

66. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 42, W320–W324 CrossRef Medline