A Legionella effector kinase is activated by host inositol hexakisphosphate

The transfer of a phosphate from ATP to a protein substrate, a modification known as protein phosphorylation, is catalyzed by protein kinases. Protein kinases play a crucial role in virtually every cellular activity. Recent studies of atypical protein kinases have highlighted the structural similarity of the kinase superfamily despite notable differences in primary amino acid sequence. Here, using a bioinformatics screen, we searched for putative protein kinases in the intracellular bacterial pathogen Legionella pneumophila and identified the type 4 secretion system effector Lpg2603 as a remote member of the protein kinase superfamily. Employing an array of biochemical and structural approaches, including *in vitro* kinase assays and isothermal titration calorimetry, we show that Lpg2603 is an active protein kinase with several atypical structural features. Importantly, we found that the eukaryote-specific host signaling molecule inositol hexakisphosphate (IP6) is required for Lpg2603 kinase activity. Crystal structures of Lpg2603 in the apo-form and when bound to IP6 revealed an active-site rearrangement that allows for ATP binding and catalysis. Our results on the structure and activity of Lpg2603 reveal a unique mode of regulation of a protein kinase, provide the first example of a bacterial kinase that requires IP6 for its activation, and may aid future work on the function of this effector during *Legionella* pathogenesis.

Protein kinases are a class of enzymes that catalyze phosphorylation, a post-translational modification involving the transfer of the terminal phosphate from ATP to protein substrates (1). The human genome encodes over 500 protein kinases that play a crucial role in cellular function and are implicated in virtually every cellular activity (2). The protein kinase superfamily can be broadly divided into two groups: the eukaryotic protein kinases and the atypical protein kinases. Eukaryotic protein kinases consist of an N-lobe and a C-lobe that harbor conserved amino acid motifs necessary for catalytic function. In contrast, atypical protein kinases lack easily detectable sequence similarity to the eukaryotic protein kinases but still retain catalytic activity. Several atypical members of the kinase superfamily have emerged in recent years that share the structural kinase fold despite significant sequence divergence (3).

We have taken a bioinformatics approach to analyze and identify these atypical and uncharacterized members of the protein kinase superfamily. With this strategy, we identified the Golgi casein kinase, Fam20C, that phosphorylates serine residues within the Ser-X-Glu/pSer consensus motif, found in roughly 75% of human plasma and cerebrospinal fluid phosphoproteins (4, 5). While searching for proteins with sequence similarity to Fam20C, we identified the atypical protein kinase, CotH, present in many bacterial and eukaryotic spore forming organisms. CotH phosphorylates spore coat proteins for effective germination in *Bacillus subtilis* (6). Furthermore, we identified a distant member of the kinase superfamily, the SelO family, that is conserved from bacteria to humans and catalyzes protein AMPylation instead of phosphorylation (7).

A subset of proteins that our bioinformatics approach continues to bring to the forefront are bacterial effector proteins. Bacterial effector proteins serve as a diverse pool of proteins in which many atypical and novel biological mechanisms can be found (8). HopBF1, for example, is a family of bacterial effector proteins from *Pseudomonas syringae* that utilize a novel molecular “mimicry” to phosphorylate host cell Hsp90 to evade the host immune defense during infection (9). Likewise, *Legionella pneumophila*, a Gram-negative intracellular pathogen known to be the causative agent of Legionnaires disease, contains a plethora of effector proteins with exciting biology. *L. pneumophila* translocates over 300 effectors to alter host cellular processes to form a replicative niche and evade degradation (10, 11). One such effector protein, SidJ, also retains a protein
kinase-like fold but catalyzes protein polyglutamylation that depends on the host cofactor calmodulin (12–15).

Interestingly, the \textit{L. pneumophila} genome encodes for five eukaryotic-like protein kinases that manipulate host cell signaling (16, 17). LegK1–4 and LegK7 are serine/threonine kinases that target various host pathways including actin remodeling, protein synthesis, the immune response, the protein folding machinery and the Hippo pathway (18–23). In search of atypical protein kinases in \textit{L. pneumophila}, we identified a kinase domain in the type 4 secretion system effector protein, Lpg2603 (lem28, sdmB), which has also been predicted to have a kinase fold by Burstein \textit{et al.} (24). Previously, a conserved phosphatidylinositol-4-phosphate (PI4P)–binding domain at the C terminus of Lpg2603 was identified and shown to localize the protein to the \textit{Legionella}-containing vacuole during infection (25).

Here, we show that Lpg2603 is an active protein kinase with several unusual structural features not typical of canonical protein kinases. Importantly, we found that Lpg2603 requires binding of the host cofactor inositol hexakisphosphate (IP6) for activation through a unique mode of active site rearrangement. Our results not only provide insights into the mechanism by which Lpg2603 catalyzes phosphorylation but also highlight a unique mechanism of allosteric regulation of kinases by IP6.

**Results**

**Lpg2603 has sequence similarity to protein kinases**

A bioinformatic screen for \textit{Legionella} effectors bearing sequence similarity to protein kinases was performed using the FFAS algorithm (29) and yielded Lpg2603 as a likely protein kinase, with 10–18% sequence identity to known bacterial kinases including OspG, NleH, YopO, PpkA, and several human kinases. The FFAS alignments allowed unequivocal assignments of the typical features of protein kinases: the Gly-rich loop and active site residues (numbering in parentheses corresponds to PKA). The logo for canonical kinases was built using 3998 homologs (Pfam domain PF00069). B, autoradiograph depicting the incorporation of \(^{32}\)P from \(^{32}\)P/ATP into MBP by recombinant Lpg2603 WT or the inactive mutant, D201A, in the presence of different cofactors: IP6, PI4P, ubiquitin (Ub), calmodulin (CaM), globular actin, or 14–3–3 protein. Note that CaM migrates at the same molecular weight as MBP. The reactions were terminated after 10 min by the addition of EDTA, and the products were resolved by SDS-PAGE and visualized by Coomassie Blue staining (lower panel) and autoradiography (upper panel). The results are representative of three independent experiments.

\footnote{The abbreviations used are: PI4P, phosphatidylinositol-4-phosphate; IP6, inositol hexakisphosphate; PKA, protein kinase A; MBP, myelin basic protein; ITC, isothermal titration calorimetry; RMSD, root-mean-square deviation; BTK, Bruton’s tyrosine kinase; AMP-PNP, adenosine 5’-(β,γ-imidotriphosphate); CK2, casein kinase 2; PH, pleckstrin homology; SeMet, selenomethionine; ADP, adenosine diphosphate.}

**Lpg2603 is an active kinase**

To determine whether Lpg2603 is an active kinase, we expressed \textit{L. pneumophila} Lpg2603 in \textit{Escherichia coli} as a...
His₆–Sumo fusion protein and purified the protein by nickel–nitrotriacetic acid affinity chromatography. We also purified recombinant Lpg2603 containing an alanine mutation in the predicted catalytic Asp²⁰¹ (PKA nomenclature Asp¹⁶⁶). Following removal of the His₆–Sumo tag, we performed kinase assays in the presence of [γ-³²P]ATP. Recombinant L. pneumophila Lpg2603, however, did not phosphorylate the generic protein kinase substrate, myelin basic protein (MBP) (Fig. 1B, first and second lanes). Bacterial effectors often utilize eukaryotic specific cofactors to regulate their activity within the host cell while remaining inactive in the bacterial cell (30). Therefore, we tested the activity of Lpg2603 in the presence of some common cofactors. Remarkably, WT Lpg2603, but not the predicted catalytically inactive D201A mutant, phosphorylated MBP in the presence of inositol hexakisphosphate (IP6) but none of the other cofactors that we tested (Fig. 1B).

**Lpg2603 requires IP6 for optimal activity**

Given the structural similarity in inositol phosphates, we tested whether other inositol phosphates could activate Lpg2603. Lpg2603 displayed a strong preference for IP6 compared with myo-inositol, inositol phosphate, inositol diphosphate, inositol triphosphate, inositol tetraphosphate, and inositol pentaphosphate with optimal MBP phosphorylation observed at ~65 μM IP6 (Fig. 2, A and B). To determine the binding affinity of IP6 to Lpg2603, we measured the dissociation constant (Kₐ) using isothermal titration calorimetry (ITC). IP6 bound to Lpg2603 with a Kₐ of ~315 μM, which is in accordance with eukaryotic cellular IP6 concentration (Fig. 2C) (31).

We next assayed the optimal divalent metal for kinase activity. Lpg2603, but not the inactive mutant, phosphorylated MBP in the presence of Mg²⁺ and Mn²⁺ (Fig. 2D). Lpg2603 demonstrates a preference for Mn²⁺ which, interestingly, is inhibitory at high concentrations (Fig. 2F). Similar to canonical kinases, Lpg2603 has a Kₐ for ATP of 18 μM (Fig. 2F). Furthermore, Lpg2603, but not the inactive mutant, phosphorylated MBP in a time-dependent manner (Fig. 2G). Collectively, our results demonstrate that Lpg2603 is a bacterial effector kinase that requires IP6 for catalytic activity.

**Crystal structure of Lpg2603 reveals an IP6-binding pocket and mechanism of activation**

To gain further insight into the mechanism of Lpg2603 activity, we solved the crystal structure of a fragment of Lpg2603, which lacks the P14P-binding region at a resolution of 2.10 Å (Fig. 3A). The apo structure displayed a unique N-terminal extension that is rich in β-strands. There was no observable electron density for residues 76–117 of the N-lobe, indicating a flexible protein region. Electron density is absent for residues that are important for canonical kinase activity including the Gly-rich loop and ion pair Lys from the VAIK motif. The C-lobe consists of five α-helices including the catalytic Asp²⁰¹ and metal-binding Asp²²⁵. The apo structure can be superimposed onto PKA with an RMSD of 4.8 Å over 178 C-α backbone atoms.

Next, we sought to compare the apo structure to the IP6-bound structure to determine the effect of cofactor binding. Surprisingly, the crystal structure of Lpg2603 bound to IP6 revealed a kinase fold with a highly ordered N-lobe with clear electron density for residues 92–117, which form three of the five strands of the canonical N-lobe β-sheet (Fig. 3B). The IP6 is coordinated in a positively charged pocket composed of residues from the N-lobe β-sheet (corresponding to strands β1–β5 of classical kinases) and residues from the N-terminal extension that includes a three-stranded β-sheet. The IP6-binding site is very well-conserved in Lpg2603 homologs, suggesting a common mechanism of activation (Fig. S1). Although the C-lobes of the apo- and IP6-bound structures remain similar, conformational shifts are observed in the IP6-binding cradle, αC-helix, and metal-binding residues that favor ATP binding (Fig. 3, A and B). The apo structure can be superimposed onto the IP6-bound structure with an RMSD of 2.3 Å over 251 C-α backbone atoms. Thus, we hypothesize that binding of IP6 stabilizes the N-lobe of Lpg2603 to allow for ATP binding and kinase activity.

We next solved the crystal structure of Lpg2603 bound to IP6 and adenosine diphosphate (ADP) to identify the residues that facilitate nucleotide coordination upon IP6 binding (Fig. 3C). The overall conformation of the kinase in the presence and absence of nucleotide is very similar with an RMSD of 0.7 Å over 274 C-α backbone atoms (Fig. 3D). One notable difference upon ADP binding is the ordering of the Gly-rich loop that folds over the nucleotide in the active site (Fig. S2). Structural homology searches using DALI identified membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase (Myt1), troponin I–interacting kinase (TNNI3K), and Bruton’s tyrosine kinase (BTK) as the closest structural homologs of Lpg2603.

**Structure guided mutagenesis highlight the residues important for nucleotide and IP6 coordination**

Using the Lpg2603 holoenzyme structure, we identified key residues that form the IP6-binding pocket (Fig. 4A). To identify IP6-binding residues that are important for kinase activity, we purified recombinant Lpg2603 proteins with alanine substitutions at Lys₁₁¹, Lys₁₅⁶, Asn₁₅⁴, Lys₁₀⁷, Arg₅⁰, and Lys₇⁶. As expected, mutations in these residues to an Ala either weakened or completely abolished Lpg2603 activation by IP6 (Fig. 4B). These results support our hypothesis that proper coordination of IP6 is necessary for Lpg2603 activation.

Next, we assayed recombinant Lpg2603 proteins with mutations in the nucleotide-binding pocket (Fig. 4C). The highly conserved catalytic loop in Lpg2603 consists of a noncanonical HPD motif with Asp²⁰¹ likely acting as the catalytic base. Mutation of the catalytic Asp²⁰¹ abolished kinase activity of Lpg2603 (Fig. 4D). A metal binding DXD motif is present in Lpg2603 where the Asp²²⁵ coordinates the Mn²⁺ ions in the active site. Mutation of the predicted metal-binding residues, Asp²²⁵ and Asn²¹³, eliminated activity of the kinase. Lys₁¹⁴ from the β3 strand (PKA equivalent Lys⁷²) coordinates the α and β phosphates of ATP (Fig. 4C). However, the glutamate corresponding to the canonical Glu⁹¹ (PKA) from the αC-helix that forms the conserved Lys⁷²–Glü⁹¹ ion pair is positioned away from the active site, whereas Asp²²⁷ coordinates the ion pair Lys₁¹⁴. Interestingly, mutation of Lys₁¹⁴ completely abolished kinase activity, whereas mutation of Asp²²⁷, which forms an ion pair with Lys¹¹⁴, retained more than 50% of activity. This is reminisc
cent of the hypomorphic E74A ion pair mutant of the bacterial effector kinase HopB1 (9). Collectively, our results highlight the residues important for IP6 binding and phosphotransfer.

IP6 binding facilitates nucleotide binding in Lpg2603

Given the structural reorganization observed with IP6, we investigated whether IP6 binding is required for nucleotide binding. To assay for ATP binding, we performed ITC with the weakly hydrolyzable ATP analog AMP-PNP. In the absence of IP6, Lpg2603 no longer binds to the nucleotide (Fig. 5A). Upon addition of IP6, Lpg2603 binds to AMP-PNP with a $K_d$ of 3.6 μM (Fig. 5B). The IP6-bound Lpg2603 crystal structure and the in vitro kinase assays indicate Lys111 as a key residue for IP6-induced activation of kinase activity. Lpg2603 K111A does not bind to AMP-PNP in the presence or absence of IP6 (Fig. 5C and D). These results provide further evidence that IP6 allosterically activates Lpg2603 kinase activity.
**IP6-dependent kinase activation**

**Figure 3. Crystal structure of Lpg2603 reveals a unique mode of kinase activation by IP6.** A, ribbon representation of apo-Lpg2603. The N-lobe and C-lobe are depicted in magenta and teal, respectively. The αC-helix and the N-terminal extensions are shown in orange and white, respectively. B, ribbon representation of Lpg2603 bound to IP6. The ligand IP6 is shown in ball-and-stick form, colored according to atom. C, ribbon representation of Lpg2603 bound to IP6 and ADP. The nucleotide ADP and ligand IP6 are shown in ball-and-stick form, colored according to atom. The Mn\(^{2+}\) ion is shown as a yellow sphere. D, superposition of the apo-Lpg2603 (shown in pink) with the nucleotide and ligand-bound Lpg2603 (shown in green) depicting the disordered N-lobe in the apo structure.

**Discussion**

Lpg2603 harbors a conserved *Legionella*-effector PI4P-binding region at the C terminus, which is required for anchoring the kinase to the *Legionella* containing vacuole during an infection (25). We have modeled this domain by using the homologous PI4P-binding domain from the *Legionella* effector DrrA and docked it against the kinase domain (Fig. S3). The full-length structure model displays the lipid-binding domain adjacent to the C-lobe of the kinase, suggesting that PI4P-binding and/or membrane localization may affect kinase function of Lpg2603. This also suggests that its substrates may be membrane-anchored or membrane-proximal. Despite our efforts to identify interacting proteins, we were unable to determine the host substrate of Lpg2603. Nevertheless, our results of the IP6-dependent kinase activation will pave the way to identify the *in vivo* substrate during *Legionella* infection.

We have identified a new mechanism of bacterial effector kinase regulation by a host signaling molecule, IP6. IP6 is a eukaryote-specific ligand and is therefore, absent in bacteria (32). Interestingly, it is also the most abundant inositol phosphate in eukaryotic cells where it regulates several processes including growth factor signaling, cell cycle progression, and vesicle trafficking (31). Cleverly, bacterial effectors have evolved to hijack host signaling molecules such as IP6 to spatially regulate activity (30). Less than a handful of other bacterial effectors from different pathogens are activated by IP6 including acetyl transferases (HopZ1, YopJ, and AvrA) and a cysteine protease (VPA1380) (30). Thus, the bacterial enzyme remains inactive within the bacterial cell until it is delivered to its destination within the eukaryotic host cell. Notably, this is the first example of a bacterial effector kinase that requires IP6 for activation. The IP6-binding site of Lpg2603 is spatially distinct from the kinase active site, suggesting allosteric regulation rather than IP6 acting as a cofactor for catalysis (Fig. 3C). Hence, we hypothesize that binding of IP6 at the N-terminal extension triggers conformational changes in multiple amino acids that functionally link the IP6-binding pocket to the kinase active site.

In addition to bacterial effectors, two eukaryotic protein kinases, BTK and protein kinase CK2, have also been shown to be activated by IP6 binding (33, 34). BTK is a nonreceptor tyrosine kinase that undergoes IP6-induced dimerization and activation (34). BTK is critical for proper B-cell function, and mutations are implicated in chronic lymphocytic leukemia, X-linked agammaglobulinemia, and several other autoimmune diseases (35). BTK is composed of multiple domains: pleckstrin homology (PH)–Tec homology domain, kinase domain, SH2 and SH3 domain (36). IP6 binds to the PH–Tec homology domain to induce dimerization and trans-autophosphorylation of the kinase domains, leading to activation (34). Notably, the mechanism of IP6 activation of BTK is distinct from Lpg2603 because we did not observe any dimerization of Lpg2603 upon incubation with IP6 (data not shown). In contrast to BTK, IP6 binds to the basic patch in the substrate recognition site of the C-lobe in CK2 (33). CK2 is a serine/threonine kinase that regulates cell proliferation (37); however, the physiological relevance of regulation of CK2 by IP6 is unclear.

Compared with the IP6-binding site of BTK (Protein Data Bank code 4Y94) and CK2 (Protein Data Bank code 3W8L), the IP6-binding site in Lpg2603 is more tightly coordinated by charged residues (Figs. 4A and 6). Moreover, in BTK and CK2, the sites are rather shallow and are built by residues from two different monomers (Fig. 6). The N-lobe of Lpg2603 and the N-terminal extension together form a β-sandwich that binds IP6 reminiscent of PH domains (Fig. S4) (38). However, connectivity of the strands in the two β-sheets is different in Lpg2603 from that of PH domains. This is likely an example of convergent evolution of a structural subdomain. The IP6-binding pocket in Lpg2603 consists of amino acids from the N-terminal extension and the N-lobe of the kinase that are highly conserved within homologs of Lpg2603. Furthermore, the apo structure revealed a fairly disordered N-lobe and active site in the absence of IP6, highlighting the importance of binding this ligand for activity. Hence, our crystallographic and biochemical analysis reveals a previously undocumented mode of allosteric regulation of the kinase superfamily.
Experimental procedures

Reagents

Selenomethionine media was purchased from Molecular Dimensions (MD12-500). Inositol 1,3,4,5-tetraphosphate (Q-1345), inositol 1,4-bisphosphate (Q-0014), inositol 1,4,5 phosphate (Q-0145), and phosphatidylinositol 4-phosphate (P4008a) were purchased from Echelon Biosciences. D-myo-Inositol-1,3,4,5,6 pentaphosphate (10009851) was purchased from Cayman Chemicals. myo-Inositol 1-dihydrogen phosphate (S860360), myo-inositol (I7508), inositol hexakisphosphate (P8810), myelin basic protein (M1891), and AMP-PNP (A2647) were purchased from Sigma. Ubiquitin and calmodulin were expressed as His6 fusion proteins and purified from E. coli as described (12). Globular actin was a generous gift from Michael Rosen (39).

Generation of constructs

*L. pneumophila* Lpg2603 residues 10-C, 21-322, or 10-322 Lpg2603 were cloned into a modified pet28a bacterial expression vector (ppSumo), containing an N-terminal His6 tag followed by the yeast Sumo (smt3). The coding sequence for the yeast homolog of 14-3-3, BMH2, was amplified by PCR using *Saccharomyces cerevisiae* BY4741 gDNA as a template. The amplified open reading frames were cloned into pGEX4T containing an N-terminal GST tag.

Protein expression and purification

*L. pneumophila* Lpg2603 10-C, 21-322, or 10-322 ppSumo were transformed into E. coli Rosetta (DE3) competent cells. The cells were grown in LB medium at 37 °C until the A600 reached ~0.5–0.7. Protein expression was induced by 0.4 mM isopropyl β-D-thiogalactoside overnight at room temperature. The cells were harvested by centrifugation and lysed in 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, and 0.1% β-mercaptoethanol) by sonication. Cell lysates were centrifuged at 25,000 × g for 25 min. The cleared lysate was incubated with nickel–nitritotriacetic acid–agarose for approximately 1 h at 4 °C. The beads were washed over a column and washed with 20 column volumes of 50 mM Tris, pH 8, 300 mM NaCl, 10 mM imidazole. Protein was eluted off the beads with 50 mM Tris, pH 8, 300 mM NaCl, and 300 mM imidazole. The proteins were cut overnight at 4 °C with His6-tagged
ULP Sumo protease and further purified using a Superdex 200 size-exclusion chromatography column attached to an AKTA Pure FPLC system (GE Healthcare).

For purification of the yeast homolog of 14-3-3, BMH2 pGEX4T was transformed into *E. coli* Rosetta (DE3) competent cells. The cells were grown in LB medium at 37 °C until the $A_{600}$ reached 0.5–0.7. Protein expression was induced by 0.4 mM isopropyl-β-D-thiogalactoside overnight at room temperature. The cells were harvested by centrifugation and lysed in 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, and 0.1% β-mercaptoethanol) by sonication. Cell lysates were centrifuged at 25,000 $g$ for 25 min. The cleared lysate was incubated with GSH–Sepharose for approximately 1 h at 4 °C. The beads were washed with 20 column volumes of 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 0.1% β-mercaptoethanol. Protein was eluted off the beads with 50 mM Tris, pH 8, 300 mM NaCl, 10 mM reduced GSH.

**In vitro kinase assays**

*In vitro* kinase assays were performed using untagged Lpg2603 10-C in a reaction mixture containing 50 mM Tris, pH 7.5, 100 μM MnCl$_2$, 100 μM [γ-32P]ATP (SA = 1000 cpm/pmol), 1 mM DTT, 167 μg/ml MBP, 21 μg/ml Lpg2603, and 20 μM IP6. Reactions were incubated at 25 °C for 10 min and terminated by the addition of 0.5 mM EDTA. SDS loading buffer was added to the samples and boiled. The reaction products were separated by SDS-PAGE and visualized by Coomassie Blue staining.

For comparison of cofactors, reactions were performed as above with the following modifications: reactions contained 20 μM IP6, 188 μM PIP4P, 40 μg/ml ubiquitin, 40 μg/ml calmodulin, or 40 μg/ml BMH2. For comparison of inositols, reactions were performed as above with the following modifications: reactions were incubated with 20 μM of the indicated inositols. For comparison of metals, reactions were performed as above with the following modifications: reactions contained 100 μM of the indicated metals. For comparison of Lpg2603 point mutants, reactions were performed as above with the following modifications: reactions contained 100 μM [γ-32P]ATP (SA = 5000 cpm/pmol). For the kinetic analysis, reactions were performed as above with the following modifications: reactions contained 300 μM MnCl$_2$, 300 μM [γ-32P]ATP (SA = 5000 cpm/pmol). Reactions were incubated at 25 °C for 15 min.

**Crystallization and structure determination**

For apo-Lpg2603, recombinant Lpg2603 10-322 in 5 mM Tris-HCl, pH 8, 30 mM NaCl was concentrated to 10 mg/ml. The crystals were grown at 20 °C by the sitting-drop vapor-diffusion method using a 1:1 ratio of protein:reservoir solution containing 18% PEG3350 and 0.2 M LiCl and were flash-frozen.

Figure 5. Nucleotide binding to Lpg2603 depends on IP6 binding. A, representative isothermal titration calorimetry data for Lpg2603 binding to AMP-PNP in the absence of IP6. For all instances, Lpg2603 is at 100 μM in the cell, and AMP-PNP is present at 2 mM in the titration syringe to a final molar ratio of 1:4. B, representative isothermal titration calorimetry data for Lpg2603 binding to AMP-PNP in the presence of IP6. Best fit parameters were as follows: $n = 1.54$; $\Delta H = -5.79$ kCal/mol. C, representative isothermal titration calorimetry data for Lpg2603 K111A binding to AMP-PNP. D, representative isothermal titration calorimetry data for Lpg2603 K111A binding to AMP-PNP in the presence of IP6. The results are representative of two independent experiments.
IP6-dependent kinase activation

in 20% PEG3350, 0.2 m LiCl, 30 mm NaCl, and 30% ethylene glycol. Apo-Lpg2603 crystals exhibited the symmetry of space group P3121 with cell dimensions of a = 62.96 Å, c = 176.41 Å, contained one apo-Lpg2603 per asymmetric unit, and diffracted to a minimum Bragg spacing (d\text{min}) of 2.10 Å when exposed to synchrotron radiation. Selenomethionine labeled-protein was obtained by expressing Lpg2603 10–322 in B834 cells grown in SelenoMet media. SeMet Lpg2603 10–322 in 5 mM Tris-HCl, pH 8, 30 mm NaCl, and 1 mM IP6, and 45% ethylene glycol. IP6-bound Lpg2603 crystals exhibited the symmetry of space group P2\text{1}2\text{1}2\text{1} with cell dimensions of a = 52.54 Å, b = 77.07 Å, c = 72.56 Å contained one IP6- and ADP-bound Lpg2603 per asymmetric unit and diffracted to a minimum Bragg spacing (d\text{min}) of 1.77 Å when exposed to synchrotron radiation.

For IP6- and ADP-bound Lpg2603, SeMet Lpg2603 21–322 in 5 mM Tris-HCl, pH 8, 30 mm NaCl, 1 mM DTT, 0.5 mM IP6, 1 mM MnCl2, and 1 mM AMP-PNP was concentrated to 10 mg/ml. The crystals were grown at 20 °C by the sitting-drop vapor-diffusion method using a 1:1 ratio of protein:reservoir solution containing 0.1 m citric acid, pH 4.0, 6% (+/−)-2-methyl-2,4-pentanediol (MPD). The wells were allowed to equilibrate for ~24 h, and crystal growth was initiated by microseeding. The crystals were flash-frozen in 0.1 m citric acid, pH 3.5, 7% (+/−)-2-methyl-2,4-pentanediol (MPD), 30 mm NaCl, 0.5 mM IP6, 1 mM MnCl2, 1 mM AMP-PNP, and 35% ethylene glycol. IP6- and ADP-bound Lpg2603 crystals exhibited the symmetry of space group P2\text{1}2\text{1}2\text{1} with cell dimensions of a = 52.54 Å, b = 77.07 Å, c = 72.56 Å contained one IP6- and ADP-bound Lpg2603 per asymmetric unit and diffracted to a minimum Bragg spacing (d\text{min}) of 1.77 Å when exposed to CuKα radiation from a home source. Although we added AMP-PNP to the protein, we observed ADP in our crystal structure. This may be due to contaminating ATP or weak hydrolysis of AMP-PNP.

Diffraction data were collected at 100 K at the Advanced Photon Source Beamline 19-ID for all data sets except the IP6- and ADP-bound Lpg2603, which was collected with CuKα radiation from a Rigaku Xtal MM003 source. Anomalous data for the SeMet apo-Lpg2603 were collected near the Se K-edge. The data were indexed, integrated, and scaled using the HKL-3000 program package (40). Data collection statistics are provided in Table S1.

**Phase determination and structure refinement**

Phases for SeMet apo-Lpg2603 were obtained from a single-wavelength anomalous dispersion experiment using a selenomethionyl-derivatized protein crystal with data collected at the selenium K-edge to a d\text{min} of 2.15 Å. Seven selenium sites were located using the program SHELXD (41), and phases were refined with the program SHELXE (42), resulting in an over-all figure-of-merit of 0.66 for data between 46.39 and 2.15 Å. Phases were further improved by density modification in the program dm (43). An initial model containing 74% of all apo-Lpg2603 residues was automatically generated in the program ARP/wARP (44).

Because the selenomethionyl derivatized and native crystals were isomorphous, all further calculations for the native structure were performed versus the native data. Additional residues for apo-Lpg2603 were manually modeled in the program Coot (45). Positional and isotropic ADP, as well as TLS ADP refinement, was performed to a resolution of 2.10 Å using the program Phenix (46) with a random 5% of all data set aside for an R\text{free} calculation. The current model contains one apo-Lpg2603 monomer; included are residues 1–75, residues 118–332, and 196 water molecules. The R\text{work} is 0.176, and the R\text{free} is 0.211. A

---

**Figure 6. Structural comparison reveals unique mode of IP6 binding to Lpg2603.** A, surface representation of BTK bound to IP6 (Protein Data Bank codes 4Y93 and 4Y94). B, surface representation of CK2 bound to IP6 (Protein Data Bank code 3W8L). C, surface representation of Lpg2603 bound to IP6. D, surface representation of the superposition of BTK, CK2, and Lpg2603. IP6 bound to BTK, CK2, and Lpg2603 is shown in green, blue, and yellow, respectively.
**IP6-dependent kinase activation**

Ramachandran plot generated with MolProbity (47) indicated that 97.4% of all protein residues are in the most favored regions, and none are in the disallowed regions.

Phases for the IP6-bound Lpg2603 were obtained by the molecular replacement method in the program Phaser (48) using the coordinates for the apo Lpg2603 monomer with residues 1–15 removed. Model building and refinement were performed to a resolution of 2.65 Å using a similar protocol to the apo structure. Three IP6-bound Lpg2603 molecules were located in the asymmetric unit, and the electron density for chain C is substantially weaker than for chains A and B. The R_work is 0.253, and the R_free is 0.282; the presence of anisotropy in the data and the weak electron density for chain C is likely the cause of the higher than expected R_work and R_free for this model. A Ramachandran plot generated with MolProbity indicates that 95.4% of all protein residues are in the most favored regions and none in disallowed regions.

Phases for the IP6- and ADP-bound Lpg2603 were obtained by the molecular replacement method in the program Phaser using the coordinates for the IP6-bound Lpg2603 monomer. Model building and refinement were performed to a resolution of 1.77 Å using a similar protocol to the apo structure. One IP6- and ADP-bound Lpg2603 molecule was located in the asymmetric unit. The R_work is 0.213, and the R_free is 0.237. A Ramachandran plot generated with MolProbity indicates that 95.4% of all protein residues are in the most favored regions, and none are in the disallowed regions. Phasing and model refinement statistics for all structures are provided in Table S1.

**Lpg2603 sequence logo**

The similarity of Lpg2603 to kinases was identified by screening the set of *L. pneumophila* subsp. Philadelphia effectors using the FFAS server (29). Homologs of the Lpg2603 kinase domain were collected using BLAST and aligned by Mafft (50). Sequence logos were produced using the WebLogo 3.0 server (51).

**Structural modeling of full-length Lpg2603**

The structural model of the DrrA domain was built using the I-Tasser server (52). The full-length structure model of Lpg2603 was built by docking together the kinase-like domain structure and the DrrA domain structure model using the AIDA server (53).

**Structural comparisons**

Structures were visualized and analyzed in PyMOL. Structure database searches were conducted using the Dali and Fatcat servers (49, 54).

**Isothermal calorimetry**

Recombinant Lpg2603 10-C WT or Lpg2603 10-C K111A was produced as described above under “Protein expression and purification.” For AMP-PNP binding assays, the final buffer consisted of 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM DTT, 1 mM MgCl₂, and 1 mM IP6. For IP6 binding, MgCl₂ and IP6 were excluded from the buffer.

For AMP-PNP binding assays, 2 mM AMP-PNP was titrated into Lpg2603 10-C at 100 μM cell concentration using the Malvern MicroCal iTC200 system at 20 °C. Initial injection was 0.5 μl, and the following 20 injections were 1.9 μl with stirring at 750 rpm. Initial injections were not included in the final analysis. Spacing between injections were 120 s to allow for proper baseline equilibration. Resulting thermograms were integrated using the NITPIC software, and SEDPHAT was used to fit the isotherms assuming a binary interaction model.

For IP6 binding assays, 8 mM IP6 was titrated into Lpg2603 10-C at 400 μM cell concentration as above. Initial injections were 0.5 μl, and the following 20 injections were 1.9 μl. Initial injections were, once again, not included in the final analysis. Spacing between injections were adjusted to 300 s to allow for proper baseline equilibration. The resulting thermograms were analyzed using the Origin software to obtain K_f and confidence intervals.

**Data availability**

All of the data are contained within the manuscript. The Lpg2603 structures have been deposited in the Protein Data Bank with accession codes as follows: apo-Lpg2603 under code 6VVV; Lpg2603 bound to IP6 under code 6VVD; and Lpg2603 bound to IP6, Mn²⁺, and ADP under code 6VVE.

**Acknowledgments**—We thank Joshua Sheetz, Shaeri Mukherjee, Kim Orth, Samantha Yee, and members of the Tagliabracci lab for valuable input. We thank Chad Brautigam and Shih-Chia Tso for help with ITC. The results shown in this report are derived from work performed at the Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source.

**References**

Legionella species identifies large and diverse effector repertoires. Nat. Genet. 48, 167–175 CrossRef Medline


**IP6-dependent kinase activation**


