Legionella phosphotyrosine phosphatase activity towards MAPKs

Andrew Quaile1, Peter J Stogios1, Olga Egorova1, Elena Evdokimova1, Dylan Vallee1, Boguslaw Nocek2, Purnima S Kompelia2, Sergio Peisajovich3, Alexander F Yakunin1, Alexander W Ensminger4, Alexei Savchenko1,5*

1Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada.
2Structural Biology Center, Advanced Photon Source, Argonne National Laboratory.
3Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada,
4Department of Biochemistry, Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada
5Department of Microbiology, Immunology and Infectious Diseases, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Running title: Legionella phosphotyrosine phosphatase activity towards MAPKs

To whom correspondence should be addressed: Dr. Alexei Savchenko. 5Department of Microbiology, Immunology and Infectious Diseases, Cumming School of Medicine, University of Calgary, 3330 Hospital Drive NW, Calgary, Alberta, Canada; Telephone: (403)-210-7980; E-mail: alexei.savchenko@ucalgary.ca

Keywords: Legionella pneumophila, crystallography, phosphotyrosine, protein phosphatase, peptide array, p38 MAPK, bacterial effector.

ABSTRACT

Host colonization by Gram-negative pathogens often involves delivery of bacterial proteins called “effectors” into the host cell. The pneumonia-causing pathogen Legionella pneumophila delivers more than 330 effectors into the host cell via its type IVB Dot/Icm secretion system. The collective functions of these proteins is the establishment of a replicative niche from which Legionella can recruit cellular materials to grow while evading lysosomal fusion inhibiting its growth. Using a combination of structural, biochemical, and in vivo approaches, we show that one of these translocated effector proteins, Ceg4, is a phosphotyrosine phosphatase harboring a haloacid dehalogenase–hydrolase domain. Ceg4 could dephosphorylate a broad range of phosphotyrosine-containing peptides in vitro and attenuated activation of MAPK-controlled pathways in both yeast and human cells. Our findings indicate that L. pneumophila’s infectious program includes manipulation of phosphorylation cascades in key host pathways. The structural and functional features of the Ceg4 effector unraveled here, provide first insight into its function as a phosphotyrosine phosphatase, paving the way to further studies into L. pneumophila pathogenicity.

Host colonization by Gram-negative pathogens often involves delivery of specific sets of bacterial proteins called “effectors” into the host cell by specialized secretion systems. Acting in concert, effectors secure the pathogen’s survival and replication, via manipulation of host processes and the establishment of subcellular environments that favor the pathogen. Disruption of effector translocation (1-3) results in attenuation of intracellular growth, highlighting the essential role of effectors in pathogen-host interactions. Despite significant progress in identification of effector arsenals in bacterial genomes and in defining the molecular functions of individual effectors, many remain uncharacterized, necessitating further studies into their role in pathogenesis. The composition and number of effectors varies dramatically between...
different pathogens. For example, plague and intestinal disease-causing *Yersinia* species encode less than ten effectors delivered by the type III secretion system (4) while the similar secretion system in *Shigella* spp. is involved in translocation of close to 30 effectors (5). Notably, sequence-related effectors are found in pathogens with very diverse invasion strategies suggesting that these effector families are involved in common host manipulation tactics.

A causative agent of severe pneumonia in humans, the *Legionella pneumophila* genome encodes over 330 effector proteins, that are translocated via the Dot/Icm (defect in organelle trafficking/intracellular multiplication), type IVB secretion system (e.g. (6-11), or see (12) for a recent review). *Legionella*’s effectors account for more than 10% of its proteome (13) and represents the largest effector set known to the bacterial world. In their natural habitat of fresh water reservoirs *Legionella* spp. invade diverse amoebae by preventing formation of the phagolysosome (14) followed by modification of the newly co-opted compartment into an organelle ideal for intracellular replication of the bacteria, the *Legionella* containing vacuole (LCV) (15). The ability to apply the same invasion strategy to invade human alveolar macrophages raises the intriguing possibility that *Legionella*’s effectors target host processes that are conserved between distant eukaryotic phyla. The sheer number of *Legionella* effectors and their apparent functional redundancy makes their functional characterization particularly challenging (16).

Studies into the function of bacterial effectors suggested that these pathogenic factors demonstrate unparalleled abilities to manipulate a wide spectrum of host cell’s processes including facilitating alterations in cytoskeletal rearrangement (17,18), vesicular trafficking (19,20), signal transduction (21,22) and transcription regulation (23,24). To achieve these effects, effectors are often involved in post-translational modification (PTM) of specific host proteins. A key regulation mechanism in both eukaryotic and bacterial cells, PTMs typically involve enzymatic covalent modification of targeted proteins at specific residues, which affect that protein’s activity, localization or interactions thus triggering a change in protein function. Effector proteins can possess not only the PTM activities found in the bacterial world but also ostensibly exclusively eukaryotic ones. Of particular prevalence, bacterial effectors have evolved to mimic the activity of the ubiquitin protein ligases, which control the final step in the eukaryote-specific PTM involving attachment of the ubiquitin polypeptide to targeted proteins, usually resulting in this protein’s degradation by the proteasome (25). Effectors with this PTM activity have been identified in the arsenals of many bacterial pathogens, including *E. coli* (26,27), *Shigella flexneri* (24,28), *Salmonella* spp. (26,29) and *Legionella pneumophila* (30,31).

The most common PTM in eukaryotic cells is phosphorylation, in which a phosphate group from a donor molecule such as ATP onto hydroxyl functional groups on residues (serine, threonine or tyrosine) of the targeted protein (32,33). This PTM is catalyzed by kinases and the human genome encodes several hundred protein kinases divided into tyrosine and serine/threonine specific enzymes (34). This PTM mechanism is involved in most if not all known human cell processes.

One of the best-studied examples of phosphorylation-controlled signaling is mediated by Mitogen-Activated Protein Kinases (MAPKs). MAPKs are highly conserved Ser/Thr protein kinases that have been extensively studied for their central roles in mediating signal transduction of extracellular stimuli to the appropriate biological response (35). MAPKs are activated by dual phosphorylation of threonine and tyrosine residues in a Thr-X-Tyr motif located in their activation loops by upstream kinases (MAPK kinases, or MAPKKs) (36). In turn, MAPKs are responsible for phosphorylating and activating downstream MAPK activated protein kinases (MAPKAPKs or MKs). These downstream targets elicit activation of processes including regulation of stress response, proliferation, differentiation and apoptosis. Their control of processes highly relevant to bacterial infections as well as their careful regulation and conservation among eukaryotes have made MAPKs attractive targets for bacterial effectors;

Icm/Dot substrates LegK1-4 share homology to eukaryotic protein kinases and activation of MAPKs in response to *Legionella* is well documented (37-39). Furthermore, while
phosphorylation of SAP/JNK, ERK1/2 and p38 are seen at early time-points even in translocation deficient mutants (39), sustained activation of SAP/JNK and p38 is reliant on effector translocation (38).

Although protein phosphatases have not been previously detected amongst Legionella's complement of effectors (40) precedent for this mechanism amongst other effectors of other species undoubtedly exists; Yersinia pestis YopH is a phosphotyrosine phosphatase capable of removing the pTyr signal through hydrolysis, thereby muting its activity (41), while Shigella flexneri OspF is a phosphothreonine lyase that irreversibly dephosphorylates the activating threonine of MAPKs (42). A recent analysis of the effector repertoire in the causative agent of Q fever, Coxiella burnetii also identified effectors Cbu1676 and Cbu0885 as phosphatases targeting the MAP kinase pathway in the eukaryotic host surrogate S. cerevisiae (43). S. cerevisiae possesses five main MAP kinase pathways regulating filamentation, cell wall integrity, sporulation, mating and hyperosmosis, with the latter two pathways, controlled by Fus3 and Hog1 respectively, being most broadly conserved among diverse eukaryotic organisms (44). The closest human homologues of Fus3 and Hog1, are ERK2 and p38. The first indication of Cbu1676 and Cbu0885 function came from in silico analysis that pointed to the presence of the haloacid dehalogenase-like (HAD-like) domain in these bacterial proteins (43). Proteins containing HAD-like domains have a broad range of activities including dehalogenase, phosphonatase and phosphomutase activity and although the majority characterized thus far are phosphatases and ATPases (45,46), protein phosphatase activity has only been observed in eukaryotes (47-50).

All HAD-like domains share a common overall fold featuring a core Rossmann fold, consisting of at least two pairs of α-helices that sandwich the core five-stranded parallel β-sheet in the order ‘54123’, with a squiggle and flap motif at the end of β1-strand (45,46). The common molecular architecture between HAD superfamily members includes four (I to IV) highly conserved sequence motifs co-localized to the active site. Motifs I and IV feature conserved aspartate residues that coordinate the Mg2+ ion required for catalysis. In addition to squiggle and flap motifs, HAD-like domains typically contain an insertion to the catalytic core domain called a cap domain, which controls active site access and is involved in substrate binding (51-53). Despite these recognizable sequence motifs, significant variation in substrate specificity and activity of the HAD-like protein family necessitates detailed structural analysis and rigorous substrate specificity studies.

In addition to the Coxiella effectors mentioned above, putative HAD-like domains have been identified in the uncharacterized effectors Lpg0096 (also known as Ceg4 / “co-regulated with the effector encoding genes 4”), Lpg1101 and Lpg2555 from L. pneumophila (43). Here, using x-ray crystallography and biochemical activity screening, we show that Ceg4 is an atypical HAD-like phosphotyrosine phosphatase able to attenuate the activation of MAP kinases in both human and yeast cells. These results indicate that L. pneumophila facilitates its infectious program by manipulation of phosphorylation cascades of key pathways in its host cells.

RESULTS

Legionella effector Ceg4 demonstrates phosphotyrosine specific phosphatase activity in vitro–The Dot/Icm-dependent translocation of L. pneumophila Ceg4 protein encoded by the lpg0096 gene was previously demonstrated using CyaA fusion translocation assays (8). Our sequence analysis using Phobius (54) suggested that in addition to the N-terminal haloacid dehalogenase-like (HAD-like) domain mentioned above, this effector contains two C-terminal transmembrane (TM) helices (residues 266-289 and 295-320). As no TM signatures were detected in Lpg1101 or Lpg2555 we hypothesize that Lpg0096/Ceg4 may be a member of a functionally diversified family that relies of localization to appropriately direct its activity toward the host. Using data from two recent large scale comparative genomics studies of Legionella species (55), we performed phylogenetic analysis of Ceg4 with the 33 other putative HAD domain-containing effectors (Figure 1). Ceg4, Lpg1101 and additional 12 homologues formed a distinct clade corresponding to Legionella orthologue group LOG_02908, while Lpg2555 and five
other effectors were localized in a distinct group. Comparative sequence analysis across the Ceg4 containing clade indicated that sequence conservation among these effectors was highest within the HAD-like domain, with significant sequence variation in their C-terminal domains. Lani_0822 from L. anisa was the sole exception to this observation, possessing an additional 50 residues at both the N and C termini. Lpg1101 is also unique within LOG_02908, having homology to the phosphatidylinositol 4-phosphate binding domain of SidM/DrrA-like (56,57) at its C-terminal. Most notably however, eight out of the 14 Ceg4 homologues possess two TM domains in their C-terminal portion, suggesting that localization to the host membrane is an important and common feature to this subset of HAD-like effectors in Legionella.

To confirm the general activity of the Ceg4-HAD-like domain, we purified an N-terminal fragment spanning residues 1 to 193 of L. pneumophila Ceg4 (see Materials and Methods for details). In line with predictions, the Ceg4[1-193] fragment demonstrated robust phosphatase activity against the generic phosphatase substrate para-nitrophenylphosphate (pNPP) (58). Highest reaction rates were observed between pH 6.5 and 8, with activity dropping markedly above pH 8. Ceg4 also demonstrated a strict requirement for Mg$^{2+}$ ions (Supplementary data, Figure S1B,C,D). Expanding on these results, we tested Ceg4[1-193] for activity against a library of 94 phosphorylated metabolic substrates allowing for querying a broad range of possible specificities (58). In these assays, Ceg4[1-193] demonstrated the highest activity toward phosphotyrosine (Figure 2C, see also Supplemental Figure S1E). To determine if Ceg4[1-193] is active against protein substrates, we screened for its ability to remove the phosphate group from a selection of 53 phosphopeptides, chosen for their importance to signaling in Saccharomyces cerevisiae, which has been successfully used as a model eukaryotic system in characterization of bacterial effector functions including those from Legionella (30,59). This set included pSer-, pThr- and pTyr-containing sequences. Consistent with our previous results, Ceg4[1-193] demonstrated robust phosphatase activity against nine diverse peptides comprising the full pool of pTyr-containing sequences in this substrate array (Figure 1E, Supplemental Figure S1F), and moreover, at levels 4 to 5 times higher than for pSer or pThr peptides. Combined, this data showed that Ceg4 is a phosphotyrosine specific phosphatase active against peptide substrates.

Crystal structure of HAD-like domain provides molecular insight into phosphatase activity of Ceg4—To gain further insight into the molecular function of Ceg4, we determined the crystal structure of the Ceg4[1-208] fragment to 1.88 Å by the single wavelength anomalous dispersion (SAD) method (see Table 1 for x-ray crystallographic statistics). The final structural model spanned Ceg4 residues 1-204 and a portion of the N-terminal fusion tag sequence (GQENLYFQG) corresponding to the TEV protease cleavage site (Figure 3A). In addition to a core Rossmann-like fold, the HAD-like domain of Ceg4 included a cap subdomain consisting of three α-helices and two long loops inserted between the D9-X-D11 “squiggle motif” and the α1 helix (Figures 2A and 2B). The location of the cap subdomain insertion in relation to the conserved motifs I and II of the core Rossmann fold classified it as a C1 cap (46). Such C1 caps were previously identified in cN-III nucleotidase (60), Eya2 protein tyrosine phosphatase (61) and MDP-1 sugar phosphatase (62,63). However, according to our analysis, the Ceg4 cap subdomain did not show any significant structural similarity with these or any other C1 caps of structurally-characterized HAD-like domains (Supplementary Figure S2).

In addition to overall structural similarity to other Rossmann-like folds, the core fold of Ceg4 possessed several conserved features consistent with the canonical motifs of other HAD-like phosphatases, such as D9 and D11 (motif I), T103 (motif II), K135 (motif III), D157 and D158 (motif IV) (Figure 3C). Collectively, these residues formed a small pocket ~350 Å$^3$ in volume, with N20 from the cap subdomain forming a “lid” over the pocket (Figure 3D). Inspection of the active site appears to confirm the correct positioning of D9 and D11 for their putative functions as the nucleophile and general acid/base residues respectively. The active site also contained a Mg$^{2+}$ ion that was coordinated by D9, D158 and three ordered water molecules; the position of the magnesium ion is conserved with those of mono- and divalent ions bound to other...
HAD-like phosphatases (60-62). The active site also contained a Cl\textsuperscript{-} ion associated with the sidechain of K135, the backbone of K104 and two ordered water molecules; its position is conserved with the position of sulfate or phosphate ions trapped in the active site of other crystallized HAD-like phosphatases (60-62). Finally, the active site also contained a highly-coordinated water that interacted with the F8, D9, D11, T103 and the Cl\textsuperscript{-} ion.

To confirm the involvement of these residues with Ceg4 catalytic activity, we performed site-directed mutagenesis and tested the resultant Ceg4[1-193] mutants \textit{in vitro} for activity towards pNPP and pTyr substrates as described above. In accordance with their predicted contributions to catalysis, the D9A, D11A, D11N, D157A, D158A and D162A mutations abrogated phosphatase activity, validating their essentiality for the catalytic function of this protein (Supplementary Figure S3).

Inspection of the molecular packing in the Ceg4[1-208] crystal lattice did not reveal significantly extensive contacts indicative of oligomerization. However, by size exclusion chromatography (Supplementary Figure S4A), we observed a mixture of monomeric and dimeric species. Consistent with this, we observed that the nine residues that we resolved of the N-terminal fusion tag comprising the TEV protease cleavage sequence (G(-8)Q(-7)E(-6)N(-5)L(-4)Y(-3)F(-2)Q(-1)G(0)) contacted the active site of the adjacent molecule in the crystal lattice (Supplementary Figure S4B). This peptide adopted an extended, nearly β-strand-like conformation (Supplementary Figure S4C). Notably, the tyrosine residue in the tag sequence deeply bound in the active site pocket, with its hydroxyl group forming a network of interactions including hydrogen bonds with the sidechain of D11, the bound Cl\textsuperscript{-} ion, and with two ordered water molecules (Supplementary Figure S4C). Other interactions included hydrogen bonds between the sidechain of Q(-1) of the tag and E40, between the backbone amide of F(-2) residue of the tag and N20 of the C1 cap, between the sidechain of K104 and the backbone carboxyl of Q(-6), and R(-7) of the tag formed two interactions (with E165 and the backbone carboxyl of G133). The C1 cap also interacts with the fusion tag peptide via hydrophobic interactions between Y43 and Y(-3) and F24 and Q(-1) and a hydrogen bond between E40 and Q(-1).

Given Ceg4 specificity toward pTyr peptides, we hypothesized that our observed binding of the tyrosine from the N-terminal fusion tag may be representative of Ceg4 interactions with its substrate. To further examine this, we modified the N-terminal fusion TEV cleavage sequence to match the sequence of pTyr carrying peptides – Q(-7)M(-6)T(-5)G(-4)Y(-3)V(-2)S(-1)T(0) identified as a Ceg4 substrate in our peptide array screening and representing the activation loop sequence of the yeast Hog1 MAP kinase (64). Hog1 is involved in a signaling pathway regulating yeast hyperosmotic adaptation (64) and is a close homologue of human p38 MAP kinase, a pathway previously implicated in \textit{Legionella} pathogenesis (38,39,65). As mentioned previously, \textit{Coxiella} HAD-like effectors efficiently modulated the activity of other yeast MAP kinases prompting us to hypothesize that our \textit{in vitro} activity results may also be indicative of Ceg4 activity against MAPK kinases.

The structure of the tag-modified Ceg4[1-208]\textsubscript{HOG1p} fragment was solved to 1.9 Å by Molecular Replacement (Table 1). This crystal structure was almost identical to the original Ceg4[1-208]\textsubscript{TEV site} structure described above, and superimposed with RMSD of less than 0.3 Å across the entire protein backbone. In this crystal structure, we resolved the positions of the six Hog1-derived residues T(-5)G(-4)Y(-3)V(-2)S(-1)T(0) from the modified fusion tag. As with the Ceg4[1-208]\textsubscript{TEV site} structure, the crystal packing of Ceg4[1-208]\textsubscript{HOG1p} showed the N-terminal peptide interacted with the active site of the adjacent Ceg4[1-208]\textsubscript{HOG1p} molecule (Figure 4A), with the Y(-3) residue bound deeply in the pocket (Figure 4B). The HOG1p peptide also adopted an extended β-strand like conformation and its structure was strikingly similar with that of the TEV cleavage site peptide, most especially across residues -5 through 0 (RMSD 0.5 Å over the six matching Cα atoms) (Figure 4C). As we could resolve only a shorter region of this peptide, we were able to identify fewer interactions between this peptide and Ceg4[1-208], and those were similar the ones observed in the Ceg4[1-208]\textsubscript{TEV}
Site structure (i.e. the interactions between Y(-3) and the active site, N20 and the backbone amide of residue (-2) of the tag, and hydrophobic interactions between Y43 and Y(-3) plus F24 and residue -1).

Next, we compared the position of the tyrosine from the expression tags and the active site configuration of the Ceg4[1-208] active site with other structurally characterized HAD phosphatases. This analysis showed that the fusion tag tyrosine hydroxyl group adopts a position that is 3.4 Å from the bound chlorine atom, which itself occupied the same general position as phosphate or phosphate analogs (i.e. phosphate bound to E. coli YrbI (PDB 3I6B, (66)) or beryllium trifluoride bound to Eya2 (PDB 3HB0, (61)) (Supplementary Figure S2B). The position of the Ceg4[1-208]-bound magnesium ions were also conserved in position with other HAD-like phosphatases (Supplementary Figure S2B). This analysis indicates that the active site configuration observed in the Ceg4[1-208] crystal structures is a good approximation of the position of a phosphotyrosine substrate bound to this enzyme.

Overall, our structural analysis revealed a compact active site in the HAD-like domain of Ceg4[1-208] restricted by a unique cap motif. This active site is able to accommodate the phosphotyrosine residue from a peptide substrate, the position of which can be gleaned from the conformation of the fusion tag in the Ceg4[1-208] crystal lattice.

Ceg4 shows activity against Hog1 and Fus3 MAP kinases in Saccharomyces cerevisiae—According to our phosphopeptide library screening, Ceg4[1-193] demonstrated equally robust activity against peptides QMTGpYVSTR and GMTEpYVATR representing the activation loops of yeast Hog1 and Fus3 MAP kinases, respectively (Fig. 1C). To test if in vitro activity of the Ceg4[1-193] fragment against yeast MAP kinase phosphopeptides is representative of in vivo activity of this effector, we tested the ability of full-length Ceg4 to affect the activation of Hog1- and Fus3-controlled pathways in a yeast model system. For this we used two S. cerevisiae strains engineered to express fluorescent reporter proteins (Stl2-BFP or Fus1-GFP) upon activation of Hog1 or Fus3 controlled pathways, respectively (see Material and Methods for details). We expressed full-length Ceg4[1-397] or the Ceg4[1-208] fragment in these S. cerevisiae strains and measured the overall fluorescence signal from 10,000 cells. S. cerevisiae strains overexpressing Ceg4 showed significant reduction of activation of both Hog1 and Fus3 activated pathways (42% and 56%, respectively), as compared to the control strain carrying an empty vector (Figure 5A). The strain expressing the Ceg4[1-208] fragment showed a reduced ability to suppress MAPK activation in the case of Fus3-controlled activation compared to the strain expressing the full-length effector. In contrast, the expression of the Ceg4[1-208] fragment resulted in further decrease in Hog1-controlled activation compared to the strain expressing the full-length effector. Based on these results, Legionella Ceg4 is implicated in regulation of host MAPK controlled pathways as demonstrated by reduced expression of fluorescent reporters for both Fus3 and Hog1. In addition, our data suggested that Ceg4 activity against the Fus3 pathway is dependent on the C-terminal region of the effector that contains the membrane-spanning elements, pointing to the potential role of this domain for in vivo specificity of Ceg4.

Next, to test the link between Ceg4 phosphatase activity and MAPK regulation we probed the effect of individual Ceg4 active site residue substitutions on the ability of this effector to dampen the activation of Hog1-controlled pathways in yeast. The Ceg4 residues targeted by this analysis were chosen based on their direct involvement in phosphatase catalytic activity (D11, K135, D158, D162 and), and their participation in forming the active site by the cap subdomain (K17, S18, N20, V23, F24, E26 and Y43) and the core HAD-like domain (K104, E108, E131, T161 and N186) (Figure 5B and C). In line with our in vitro activity results, alanine substitution of the conserved D11, K135, D158 and D162 residues directly involved in catalytic activity of the HAD-like domain resulted in abrogation of Ceg4 suppression of the hyperosmotic stress response. This observation confirmed our hypothesis that Ceg4-triggered dampening of MAPK activation is directly linked to the phosphatase activity of its HAD-like domain. Substitution of active site pocket residues V23, F24, Y43 and K104 also had a
negative impact on the ability of Ceg4 to control MAPK activation with V23D substitution resulting in complete loss of activity. In contrast, substitutions of Ceg4 K17, E26, E108, E131 and T161 which are located distally from the active site and N20 which partially covers the entrance to the catalytic pocket did not significantly affect its activity against Hog1 activated pathway. These observations are consistent with previously observed mechanisms of catalysis of HADs and furthermore indicate the involvement of certain cap domain residues in substrate recognition.

Combined, these results clearly linked Ceg4’s in vivo activity as a regulator of MAPK pathways with its HAD-like domain phosphatase active site and identified individual residues involved in catalysis and substrate interactions.

Ceg4 localizes to HeLa endoplasmic reticulum and attenuates MAPK p38 activation in vivo—Having determined that Ceg4 is able to act on conserved eukaryotic MAP kinases, and having also established a possible role of C-terminal region of the Ceg4 effector in this activity, we were interested in identifying the subcellular localization of Ceg4 in human cells and in determining if the MAPK dampening activity extended to human MAP kinases.

To that end, human HeLa cells were transfected with constructs expressing wild type Ceg4, the catalytically inactive variant Ceg4D9A, Ceg4[1-207] and Ceg4[208-397] fragments each fused to GFP. In keeping with the presence of the predicted transmembrane regions in the C-terminal portion of Ceg4, the full length (both wild type and D9A mutant) and Ceg4[208-397] fragments demonstrated specific perinuclear localization, while the construct Ceg4[1-207] which lacks the C-terminal region containing the TM domain showed diffuse localization throughout the cell (Figure 6A). Given the differential importance of the C-terminal domain to dampening S. cerevisiae pheromone and hyperosmolarity responses we posit that this portion of Ceg4 is critically important to its function in the host cell. We therefore sought to more specifically determine the sub-cellular localization of Ceg4. HeLa cells transfected with N-terminally GFP tagged Ceg4D9A, counterstained with ER-tracker dye demonstrated that Ceg4 co-localizes predominantly with endoplasmic reticular structures.

With both structural and biochemical analyses indicating MAPK activation loop sequences as potential targets of Ceg4 phosphatase activity, we also tested its ability to perform this function on the closest human homologue of Hog1, the p38 MAPK. HEK293t cells transfected with either Ceg4 or Ceg4D9A were tested for changes in the phosphorylation state of human p38 MAPK upon stimulation with either TPA or anisomycin. Western blot analysis for both total p38 and phospho-p38 (Figure 6C) showed that while the total levels of p38 are the same in cells expressing the wild type and catalytically-inactive mutant, cells harboring wildtype Ceg4 demonstrated a significant reduction of signal corresponding to phosphorylated p38 compared to the same signal in cells carrying the Ceg4D9A mutant.

Combined, our structural and functional analysis has identified Legionella Ceg4 as a bacterial HAD protein tyrosine phosphatase that is able to attenuate the MAP kinase responses in both human and yeast cells in vivo, and does so via removal of the phosphate moiety from phospho-tyrosine in their activation loops that are critical to their activation.

**DISCUSSION**

Translocation of effector proteins inside the host cell is an important and common strategy adopted by many Gram-negative bacteria including important human pathogens. This necessitates the functional characterization of specific effectors as a necessary step in understanding of host-pathogen interactions and for the development of novel anti-bacterial therapies. Here, we show the conserved Legionella effector Ceg4 can modulate the phosphorylation state of eukaryotic MAP kinases through its HAD-like phosphatase domain and we clarify the molecular structure of this domain, providing key molecular details into this effector.

Ceg4 is one of three predicted HAD-like domain containing effectors in the Legionella pneumophila genome that are known to be translocated by the Dot/Icm system. According to our analysis, HAD-like effectors are also found in other Legionella species suggesting that this functional domain is widely used by these pathogens for manipulation of host signaling pathways. Despite commonality among HAD-
like effectors, Ceg4 represents a sequence-distinct group containing eight *Legionella* effectors that feature the combination of an N-terminal HAD-like domain and a C-terminal region carrying two transmembrane helices. The C-terminal region of the Ceg4 effector plays an important part in interactions of this effector with eukaryotic MAP kinases as deletion of this region had a significant effect on the ability of this effector to dampen the signaling by the yeast Fus3p MAPK. Furthermore, alanine substitution of catalytic residues D11 and D158 in the Ceg4 HAD-like domain active site abrogated phosphatase activity and ability to dampen MAPK activation. Combined with the broad substrate specificity of the Ceg4 HAD-like domain toward phosphotyrosine peptides demonstrated by our *in vitro* assays, this observation prompted us to suggest that the C-terminal region may be responsible for tailoring the general phosphatase activity of Ceg4 toward its specific host target.

The crystal structure of the Ceg4[1-208] fragment showed strong similarity with previously structurally-characterized members of the HAD-like protein family including specific features of the active site such as conservation of key catalytic residues and coordination of a Mg$^{2+}$ ion, known to be essential for catalysis (51). Probing the Ceg4 active site cavity with site-directed mutagenesis not only confirmed the role of residues predicted to be directly involved in catalysis but also revealed the subset of residues important for Ceg4 activity against MAP kinase substrates. Substitution of residues such as D11 which acts as a general acid/base, drastically reduced or completely abrogated the activity of this effector against Fus3 MAPK in a yeast model system. Notably, previous prediction of specific residues important for Ceg4 activity based on similarity to *Coxiella* HAD-like effectors (43) was only partially confirmed by our structure and subsequent mutagenesis. Specifically, key residues of motifs II and III were previously predicted to correspond to Ser81 and Lys111. However, our structural analysis pointed instead to Thr103 and Lys135 fulfilling this role. This observation highlights the limits of primary sequence based analysis applied to highly diverse HAD-like proteins, and effectors in general and reiterates the necessity in molecular and structural data to complement functional diversity across large protein families.

The substrate specificity of HAD-like domains is often defined by the ‘cap’ subdomain insertion that controls the access to the catalytic center (67). The Ceg4 HAD-like domain structure features an unusual α-helical cap motif never before described for this domain. This novel cap motif is compatible with the broad specificity of this domain against phosphotyrosine peptides as indicated by our observation that the tyrosine residue from two different N-terminal tag sequences is able to make intimate contacts with the Ceg4 active site of a neighboring molecule in two different crystal lattices. The specific position of the tyrosine residue is the Ceg4 active site is compatible with the position of trapped phosphate/phosphate analog substrates in other structurally-characterized HAD-like phosphatases, suggesting that this crystallization observation may indeed be representative of the interaction between the Ceg4 HAD-like with its phosphoprotein substrate.

Effectors have been demonstrated to target all strata of the MAPK controlled signaling pathways (MAPKKK, MAPKK, MAPK and MKs) using several differing enzymatic activities (68,69). Characterization of Ceg4 phosphatase activity against yeast and human MAP kinases adds a new member to this growing list of MAPK-modulating factors along with the recently-characterised *Coxiella* HAD-like effectors active against the yeast CWI MAPK (43). Human HEK293t cells transfected with Ceg4 demonstrated clear reduction of the amount of phosphorylated p38 MAP kinase, and this was compromised by mutation of the key Ceg4 active site residue D9. Previous work has shown that MAPK phosphorylation in human cells is increased at very early time points of bacterial challenge by *Legionella*, and that this activation is sustained for some time in an effector-dependent manner (39). Therefore, it is tempting to speculate that Ceg4 would serve a functional role only at significantly later points of infection and in keeping with this model, RNA-Seq data taken during infection show that Ceg4 transcription levels are low during the post exponential/infectious stage but increase 10-fold during exponential growth (70). An additional
possibility is that given the C-terminal dependent localization to the endoplasmic reticulum membranes and the loss of ability of Ceg4 to modulate Fus3 activation without this domain it is possible that Ceg4 acts to reduce specific MAPK activation in a localized manner, while allowing general activation throughout the rest of the cell.

Characterization of the structural and functional features of the Ceg4 effector presented in this work provide the first insight into function of this and other HAD-like effectors during Legionella infection, paving the way to further studies into this bacteria’s pathogenic strategy.

**EXPERIMENTAL PROCEDURES**

**Protein purification and size exclusion chromatography**—Based on domain and transmembrane location predictions, genes corresponding to Ceg4 residues 1-208 and 1-193 were cloned into p15Tv-LIC-TEV by ligation independent cloning, and transformed into *E. coli* BL21 CodonPlus (DE3) RIPL competent cells using standard procedures. Several mutants of Ceg4[1-193] in p15Tv-LIC-TEV (D9A, D11N, D11A, E26A, D157A, D158A and D162A) along with a variant of Ceg4[1-208] with the TEV sequence ‘GRQNLYFQG’ mutated to match the activation loop sequence from *S. cerevisiae* Hog1 ‘PQMTGYVST’ were prepared by site-directed mutagenesis. Cultures were grown at 37 °C in M9 media with selenomethionine or LB media, supplemented with kanamycin and at an OD600 of 0.6-0.8 and expression of 6xHis-TEV tagged Ceg4 was induced by the addition of 0.4 mM isopropyl-β-D-1-thiogalactopyranoside and the temperature of the culture was reduced to 16°C overnight. The following day, cultures were harvested by centrifugation and pellets lysed by sonication on ice in 50 mM HEPES (pH 7.5), 150 mM NaCl, and 1 mM PMSF. All further purification was conducted at 4 °C. Cell lysate was clarified by ultracentrifugation at 17,000 × g for 30 min, and 1–5 ml of Ni-NTA resin (Qiagen) was added and incubated with gentle rotation for 30 min. Resin was washed with 50 mM HEPES (pH 7.5), 150 mM NaCl, and 30 mM imidazole, and protein eluted with 50 mM HEPES (pH 7.5), 150 mM NaCl, and 500 mM imidazole. Proteins were further concentrated using a centrifugal concentrator, flash-frozen in liquid N2, and stored at −80 °C. Oligomerization of Ceg4[1-208] was tested by size exclusion chromatography using a Superdex S200 column with running buffer 50 mM HEPES (pH 7.5) and 150 mM NaCl.

**General enzyme activity screening**—General screens for enzyme activity were performed as previously described (58) using Ceg4[1-193]. Briefly, 20 µL of purified protein (at 0.5 µg/µL) was added to wells of a 96-well plate, then 180 µL of protease, phosphatase, phosphodiesterase, dehydrogenase, oxidase, NADH/NADPH oxidase and lipase, or 170 µL of thioesterase mixes (containing buffers, metal cations and substrates) were added, as well as 10 µL of 5,5'-dithiobis(2-nitrobenzoic acid) to the thioesterase mix. Plates were incubated at 37°C for 1 hour. Phosphatase, phosphodiesterase, protease, lipase, and thioesterase results were read at 410 nm, dehydrogenase, oxidase, and NADH/NADPH oxidase results were monitored at 340 nm. All results were obtained using a Spectramax M2 plate reader.

**Natural phosphatase substrate screening**—Screens for activity towards naturally occurring phosphatase substrates were performed as previously described (58) using Ceg4[1-193]. Briefly, two 96-well plates were prepared, one as a blank control and a second for the protein. 10 µL of natural phosphatase substrate at 4 mM (see table S1 for list) was added to each assay well, followed by 150 µL of reaction mixture or reaction mixture containing 2 µg protein to give a final concentration of 50 mM HEPES-K (pH 7.5), 5 mM MgCl₂, 1 mM MnCl₂ and 0.5 mM NiCl₂. Plates were incubated at 37 °C for 30 mins. After incubation, 40 µl malachite green development reagent was added to each well prior to reading the absorbance at 630 nm. (mM phosphate.min⁻¹.mg⁻¹) was calculated based on a KH₂PO₄ standard curve.

**Ceg4 metal and pH dependence assays and kinetics using p-nitrophenylphosphate**—Assays for metal requirements were conducted using 20 mM *p*-nitrophenylphosphate (pNPP) in 50 mM HEPES-K (pH 7.0), 0.1 µg.mL⁻¹ of purified Ceg4 [1-193] and the concentration of metals indicated, in a final volume of 200 µL. For pH optimizations, reactions were conducted using 20 mM pNPP, 15 mM MgCl₂, and 0.1 µg.mL⁻¹ Ceg4[1-193], and one of the following
Legionella phosphotyrosine phosphatase activity towards MAPKs

buffers at 50 mM; MES (pH 6.5)/HEPES (pH 7)/HEPES (pH 7.5)/HEPES (pH 8)/CHES (pH 9)/CHES (pH 9.5)/CHES (pH 10). For kinetics determinations, pNPP was used at the concentrations specified. Reactions were started by the addition of enzyme and monitored using a Spectramax M2 plate reader at 405 nm for 25 minutes at 25 °C. Data analysis and curve fitting was performed using Graphpad Prism 6.

Phosphopeptide phosphatase assay—Phosphatase activity towards a library of 53 phosphopeptide sequences was tested as previously described (46). Briefly, 10 µL peptide solution was mixed with 150 µL of 50 mM HEPES (pH 7.0), 15 mM MgCl₂ and 0.01 µg of SBP-purified Ceg4[1-193]. The reaction was incubated for 10 minutes at 25 °C before addition of 40 µL of malachite green development reagent. Absorbance at 630 nM was recorded. Enzyme velocity (in mM phosphate.min⁻¹.mg⁻¹) was calculated based on a KH₂PO₄ standard curve. Subsequent kinetics determinations with peptide QMTGpYVSTR were performed using the method above, with the peptide concentrations specified in the figure. Data analysis and curve fitting was performed using Graphpad Prism 6.

Crystallization, structure determination and analysis—Crystals of selenomethionine-substituted Ceg4[1-208]TEV site were grown at 23 °C using the hanging-drop vapor diffusion method by mixing 2 µL of 102 mg/mL protein with 2 µL of reservoir solution containing 0.2 M MgCl₂, 0.5 mM MnCl₂, 0.1 M Tris pH 7.3, 30% (w/v) PEG 4K and 2 mM phosphotyrosine. Crystals of native Ceg4[1-208]HOG₁p were grown at 23 °C using hanging-drop vapor diffusion by mixing 2 µL of 50 mg/mL protein with 2 µL of reservoir solution containing 0.2 M NaCl, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350. All crystals were cryo-protected with reservoir solution supplemented with paratone oil.

Diffraction data were collected at 100 K at beamline 19-ID at Structural Biology Center, Advanced Photon Source at the wavelength of 0.979 Å (selenium peak). All diffraction data were reduced with HKL-3000 (71). The Ceg4[1-208]TEV site structure was solved first by SAD phasing using PHENIX.solve (72) which identified all five selenomethionine residues in the asymmetric unit, followed by model building by PHENIX.autobuild. The structure of Ceg4[1-208]HOG₁p was determined by Molecular Replacement using the Ceg4[1-208]TEV site as search model using PHENIX.phaser. All structures were refined using PHENIX.refine and Coot (73). The final Ceg4[1-208]TEV site model includes the sequence GRQNLYFQG from the TEV site followed by residues 1-204 of Ceg4; the final Ceg4[1-208]HOG₁p model includes the sequence TGYVST from yeast Hog1 followed by residues 1-204 of Ceg4, with residues 146 and 147 unmodeled due to poor electron density. B-factors were refined as isotropic for all structures. All geometries were verified with PHENIX.refine and the wwPDB Validation server. Structure coordinates were deposited to the Protein Databank under accession codes 6AOK and 6AOJ for the Ceg4[1-208]TEV site and Ceg4[1-208]HOG₁p structures, respectively. Structural orthologs were identified using the Dali lite server (74). Active site volume was calculated by the CastP (75).

Yeast transformation and MAPK pathway activation assay—Ceg4[1-397] or the mutants specified were cloned into pYES2 NT/A, transformed into S. cerevisiae (W303 MATa, bar1::NatR, far1Δ, mfa2::p Fus1-GFP, ura3::Kan-pSTL1-BFP, his3, trp1, leu2) using the LiAc procedure (76) and grown on selective media for two days at 30 °C. For analysis of mating and high osmolarity glycerol (HOG) pathway responses by flow cytometry, transformants were grown either in duplicate or triplicate in selective medium overnight at 30 °C. Empty plasmid was grown as a negative control. Overnight cultures were diluted to OD600 between 0.1-0.2 and grown to early log phase. For analysis of mating pathway response, yeast cells were treated with 1 mM α-factor and incubated at 30 °C for two hours. For analysis of the HOG pathway response, cells were treated with 2 mM KCl and incubated at 30 °C for one hour. For both conditions, cells were then treated with the protein synthesis inhibitor cycloheximide for 30 minutes. For each sample, 10,000 cells were measured with a MACSQuant Vybe (Miltenyi Biotech). Data shown are mean fluorescence (GFP for mating response and BFP for HOG response) and standard deviation of duplicates or triplicates.
**HEK293t and HeLa cell culture and transfection**—For MAPK activation assays, HEK293t cells were maintained in Dulbecco’s Modified Eagle’s Medium (DMEM), supplemented with 10 % FBS at 37 °C and 5 % CO₂, and grown to a confluence of approximately 70 % at the time of transfection. Cells were transfected using Lipofectamine 3000 (Life Technologies) according to the manufacturer’s instructions, with endotoxin free pcDNA3-N-Flag-LIC containing Ceg4[1-397] or Ceg4[1-397]D9A.

For cell localization studies, HeLa cells were grown on poly-L-lysine treated glass coverslips, maintained in DMEM, supplemented with 10 % FBS at 37 °C and 5 % CO₂, and grown to a confluence of approximately 70 % at time of transfection. Cells were transfected with either pEGFP-N1 containing Ceg4[1-397], Ceg4[1-397]D9A, Ceg4[1-207] or Ceg4[208-397] using Lipofectamine 3000 (Life Technologies) according to the manufacturer’s instructions followed by counterstaining with DAPI during fixation. For co-localization studies, pEGFP-N1 Ceg4[1-397]D9A transfected cells were grown in chambered coverslips and incubated with ER-tracker and Lyso-tracker red dyes according to the manufacturer’s instructions prior to fixation and counterstaining with DAPI. Microscopy data were collected using a Nikon TiE inverted microscope and Nikon C2 confocal system, with a 60X oil immersion lens.

**MAPK activation and Immunoblotting**—24 hours post-transfection with Ceg4[1-397] or Ceg4[1-397]D9A, 5 µg/ml anisomycin or 200 nM TPA was added to the culture medium. After 30 minutes, cells were gently washed once with PBS followed by lysis directly into SDS-PAGE loading buffer. SDS-PAGE was performed with the addition of 0.5% 2,2,2-trichloroethanol added to the resolving portion of the gel. After electrophoresis, gels were exposed to UV light for 2 minutes and imaged with a GelDoc (BioRad) to obtain loading controls prior to western blotting. Proteins were transferred to nitrocellulose using a Transblot Turbo (BioRad). Membranes were blocked with blocking buffer (5% w/v BSA in TBS with 0.1% Tween 20) for 1 hour, followed by incubation with either anti-p38 (Cell Signalling Technology, #8690, 1:1000 dilution in blocking buffer) or anti phospho-p38 (Cell Signaling Technology, #4511, 1:1000 dilution in blocking buffer) overnight. After washing, membranes were incubated in 5% non-fat skim milk in TBS with 0.1% Tween 20 containing anti-rabbit HRP (Cell Signalling Technology, #7074, 1:4000 dilution) for 1 hour. Blots were developed with BioRad Clarity Western plus reagent and imaged using a GelDoc (BioRad) and visualized using ImageLab (BioRad).

**Visualization of total cellular tyrosine phosphorylation by immunoblot in S. cerevisiae and HEK293T.** BY4741 S. cerevisiae (MATα his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 (77)) were transformed with the indicated constructs in pYES2NT/A using the LiAc procedure (76) and grown on CM agar-uracil selective media for two days at 30 °C. Colonies were picked and cultured overnight in selective media supplemented with 2% (w/v) raffinose. In the morning, 3ODs of cells were harvested by centrifugation, washed and resuspended in 5 ml selective media supplemented with 2% (w/v) galactose and incubated at 30 °C with shaking for 5 hrs. Cultures were harvested by centrifugation and cells were lysed using an alkaline/SDS lysis procedure (78). HEK293T cells were cultured in and transfected in 6 well plates as per procedures described for MAPK activation assays. Cells were washed with PBS and harvested directly using 100 µl/well of SDS-PAGE loading buffer, followed by brief sonication to reduce viscosity and 5 minutes of heating at 95 °C. 20 µl of each yeast lysate and 15 µl was loaded and separated with 12 % SDS-PAGE gels supplemented with 0.5 % (v/v) 2,2,2-trichloro ethanol, followed visualization of protein loading after exposure of gels to UV light for 2-3 minutes. Protein was transferred to nitrocellulose for immunoblotting using a Transblot Turbo (BioRad). For total phosphotyrosine detection, yeast and HEK293T blots were blocked with 5% w/v BSA, 1X TBS, 0.1% Tween 20 at 4°C for 1hr with gentle shaking followed by overnight incubation with α-P-Tyr-100 antibody (Cell Signalling Technology, #9411, 1:2000 in blocking buffer). For Ceg4 detection in yeast, blocking was performed in 5% non-fat skim milk in TBS with 0.1% Tween 20 followed by incubation with α-Xpress (Thermofisher Scientific, #R910-25 1:4000) overnight. For detection of Ceg4 in HEK293T...
Legionella phosphotyrosine phosphatase activity towards MAPKs

cells, blocking was performed in 3% non-fat skim milk in TBS followed by incubation with α-FLAG M2 (1:2000, Cat# F1804, Sigma-Aldrich) for 45 minutes at room temperature.

After 3, 5 minute washes with TBS with 0.1% Tween 20, all blots were incubated in 5% non-fat skim milk in TBS with 0.1% Tween 20 containing anti-mouse HRP (Cell Signalling Technology, #7076, 1:4000 dilution) for 1 hour. Blots were developed with BioRad Clarity Western plus reagent and imaged using a GelDoc (BioRad) and visualized using ImageLab (BioRad).

Acknowledgements: Special thanks to Greg Brown, for his invaluable assistance with screening and activity assays. We would also like to thank the Claycomb lab for training and use of their microscopy facilities.

Conflict of interest: The authors declare no conflicts of interest.

Author contributions: AQ performed human in vivo work with the help of DV. EE and OE purified and crystalized Ceg4. OE and AQ performed enzyme activity screening. PK and SP performed yeast MAPK assays. BN collected crystal data. PS solved the structures and wrote the manuscript. AQ and AS wrote the manuscript with input from AWE and AS. All work was performed under the supervision of AWE, AFY and AS.
REFERENCES
1. Viboud, G. I., and Bliska, J. B. (2005) Yersinia outer proteins: role in modulation of host cell signaling responses and pathogenesis. Annu Rev Microbiol 59, 69-89
2. Watarai, M., Tobe, T., Yoshikawa, M., and Sasakawa, C. (1995) Contact of Shigella with host cells triggers release of Ipa invasin and is an essential function of invasiveness. The EMBO journal 14, 2461-2470
3. Zierler, M. K., and Galan, J. E. (1995) Contact with cultured epithelial cells stimulates secretion of Salmonella typhimurium invasion protein InvJ. Infect Immun 63, 4024-4028
4. Zhang, L., Mei, M., Yu, C., Shen, W., Ma, L., He, J., and Yi, L. (2016) The Functions of Effector Proteins in Yersinia Virulence. Pol J Microbiol 65, 5-12
5. Parsot, C. (2009) Shigella type III secretion effectors: how, where, when, for what purposes? Curr Opin Microbiol 12, 110-116
6. Huang, L., Boyd, D., Amyot, W. M., Hempstead, A. D., Luo, Z. Q., O'Connor, T. J., Chen, C., Machner, M., Montminy, T., and Isberg, R. R. (2011) The E Block motif is associated with Legionella pneumophila translocated substrates. Cell Microbiol 13, 227-245
7. Zhu, W., Banga, S., Tan, Y., Zheng, C., Stephenson, R., Gatley, J., and Luo, Z. Q. (2011) Comprehensive identification of protein substrates of the Dot/Icm type IV transporter of Legionella pneumophila. PLoS One 6, e17638
8. Burststein, D., Zusman, T., Degtyar, E., Viner, R., Segal, G., and Pupko, T. (2009) Genome-scale identification of Legionella pneumophila effectors using a machine learning approach. PLoS Pathog 5, e1000508
9. de Felipe, K. S., Pampou, S., Jovanovic, O. S., Pericone, C. D., Ye, S. F., Kalachikov, S., and Shuman, H. A. (2005) Evidence for acquisition of Legionella type IV secretion substrates via interdomain horizontal gene transfer. J Bacteriol 187, 7716-7726
10. Kubori, T., Hyakutake, A., and Nagai, H. (2008) Legionella translocates an E3 ubiquitin ligase that has multiple U-boxes with distinct functions. Mol Microbiol 67, 1307-1319
11. Luo, Z. Q., and Isberg, R. R. (2004) Multiple substrates of the Legionella pneumophila Dot/Icm system identified by interbacterial protein transfer. Proceedings of the National Academy of Sciences of the United States of America 101, 841-846
12. Ensminger, A. W. (2016) Legionella pneumophila, armed to the hilt: justifying the largest arsenal of effectors in the bacterial world. Curr Opin Microbiol 29, 74-80
13. Chien, M., Morozova, I., Shi, S., Sheng, H., Chen, J., Gomez, S. M., Asamani, G., Hill, K., Nuara, J., Feder, M., Rineer, J., J. G. J., Sheshenko, V., Park, S. H., and Geringer-Sameth, A. (2004) The genomic sequence of the accidental pathogen Legionella pneumophila. Science 305, 1966-1968
14. Horwitz, M. A. (1983) The Legionnaires' disease bacterium (Legionella pneumophila) inhibits phagosome-lysosome fusion in human monocytes. J Exp Med 158, 2108-2126
15. Horwitz, M. A. (1983) Formation of a novel phagosome by the Legionnaires' disease bacterium (Legionella pneumophila) in human monocytes. J Exp Med 158, 1319-1331
16. O'Connor, T. J., Adepoju, Y., Boyd, D., and Isberg, R. R. (2011) Minimization of the Legionella pneumophila genome reveals chromosomal regions involved in host range expansion. Proc Natl Acad Sci U S A 108, 14733-14740
17. Gruenheid, S., DeVinney, R., Bladt, F., Goosney, D., Gelkop, S., Gish, G. D., Pawson, T., and Finlay, B. B. (2001) Enteropathogenic E. coli Tir binds Nck to initiate actin pedestal formation in host cells. Nat Cell Biol 3, 856-859
18. Michard, C., Sperandio, D., Bailo, N., Pizarro-Cerda, J., LeClaire, L., Chadeau-Argaud, E., Pombo-Gregoire, I., Hervet, E., Vianney, A., Gilbert, C., Faure, M., Cossart, P., and
Legionella phosphotyrosine phosphatase activity towards MAPKs

Doublet, P. (2015) The Legionella Kinase LegK2 Targets the ARP2/3 Complex To Inhibit Actin Nucleation on Phagosomes and Allow Bacterial Evasion of the Late Endocytic Pathway. mBio 6, e00354-00315

19. Nagai, H., Kagan, J. C., Zhu, X., Kahn, R. A., and Roy, C. R. (2002) A bacterial guanine nucleotide exchange factor activates ARF on Legionella phagosomes. Science 295, 679-682

20. Hernandez, L. D., Hueffer, K., Wenk, M. R., and Galan, J. E. (2004) Salmonella modulates vesicular traffic by altering phosphoinositide metabolism. Science 304, 1805-1807

21. Mittal, R., Peak-Chew, S. Y., and McMahon, H. T. (2006) Acetylation of MEK2 and IkappaB kinase (IKK) activation loop residues by YopJ inhibits signaling. Proc Natl Acad Sci U S A 103, 18574-18579

22. Selyunin, A. S., Sutton, S. E., Weigele, B. A., Reddick, L. E., Orchard, R. C., Bresson, S. M., Tomchick, D. R., and Alto, N. M. (2011) The assembly of a GTPase-kinase signalling complex by a bacterial catalytic scaffold. Nature 469, 107-111

23. Cui, J., Yao, Q., Li, S., Ding, X., Lu, Q., Mao, H., Liu, L., Zheng, N., Chen, S., and Shao, F. (2010) Glutamine deamidation and dysfunction of ubiquitin/NEDD8 induced by a bacterial effector family. Science 329, 1215-1218

24. Ashida, H., Kim, M., Schmidt-Supprian, M., Ma, A., Ogawa, M., and Sasakawa, C. (2010) A bacterial E3 ubiquitin ligase IpaH9.8 targets NEMO/IKKgamma to dampen the host NF-kappaB-mediated inflammatory response. Nat Cell Biol 12, 66-73; sup pp 61-69

25. Ashida, H., and Sasakawa, C. (2017) Bacterial E3 ligase effectors exploit host ubiquitin systems. Curr Opin Microbiol 35, 16-22

26. Lin, D. Y., Diao, J., and Chen, J. (2012) Crystal structures of two bacterial HECT-like E3 ligases in complex with a human E2 reveal atomic details of pathogen-host interactions. Proc Natl Acad Sci U S A 109, 1925-1930

27. Wu, B., Skarina, T., Yee, A., Jobin, M. C., Dileo, R., Semesi, A., Fares, C., Lemak, A., Coombes, B. K., Arrowsmith, C. H., Singer, A. U., and Savchenko, A. (2010) NleG Type 3 effectors from enterohaemorrhagic Escherichia coli are U-Box E3 ubiquitin ligases. PLoS Pathog 6, e1000960

28. Singer, A. U., Rohde, J. R., Lam, R., Skarina, T., Kagan, O., Dileo, R., Chirgadze, N. Y., Cuff, M. E., Joachimiak, A., Tyers, M., Sansonetti, P. J., Parsot, C., and Savchenko, A. (2008) Structure of the Shigella T3SS effector IpaH defines a new class of E3 ubiquitin ligases. Nature structural & molecular biology 15, 1293-1301

29. Quezada, C. M., Hicks, S. W., Galan, J. E., and Stebbins, C. E. (2009) A family of Salmonella virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases. Proc Natl Acad Sci U S A 106, 4864-4869

30. Quaile, A. T., Urbanus, M. L., Stogios, P. J., Nocek, B., Skarina, T., Ensminger, A. W., and Savchenko, A. (2015) Molecular Characterization of LubX: Functional Divergence of the U-Box Fold by Legionella pneumophila. Structure 23, 1459-1469

31. Ensminger, A. W., and Isberg, R. R. (2010) E3 ubiquitin ligase activity and targeting of BAT3 by multiple Legionella pneumophila translocated substrates. Infect Immun 78, 3905-3919

32. Khoury, G. A., Baliban, R. C., and Floudas, C. A. (2011) Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. Sci Rep 1

33. Cohen, P. (2002) The origins of protein phosphorylation. Nat Cell Biol 4, E127-130
Legionella phosphotyrosine phosphatase activity towards MAPKs

34. Manning, G., Plowman, G. D., Hunter, T., and Sudarsanam, S. (2002) Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* **27**, 514-520
35. Seger, R., and Krebs, E. G. (1995) The MAPK signaling cascade. *FASEB J* **9**, 726-735
36. Cargnello, M., and Roux, P. P. (2011) Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiol Mol Biol Rev* **75**, 50-83
37. N'Guessan, P. D., Etouem, M. O., Schmeck, B., Hocke, A. C., Scharf, S., Vardarova, K., Opitz, B., Flieger, A., Sutterp, N., and Hippenstiel, S. (2007) Legionella pneumophila-induced PKalpha-, MAPK-, and NF-kappaB-dependent COX-2 expression in human lung epithelium. *Am J Physiol Lung Cell Mol Physiol* **292**, L267-277
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
39. Welsh, C. T., Summersgill, J. T., and Miller, R. D. (2004) Increases in c-Jun N-terminal kinase/stress-activated protein kinase and p38 activity in monocyte-derived macrophages following the uptake of Legionella pneumophila. *Infect Immun* **72**, 1512-1518
40. Haenssler, E., and Isberg, R. R. (2011) Control of host cell phosphorylation by legionella pneumophila. *Front Microbiol* **2**, 64
41. Zhang, Z. Y., Clemens, J. C., Schubert, H. L., Stuckey, J. A., Fischer, M. W., Hume, D. M., Saper, M. A., and Dixon, J. E. (1992) Expression, purification, and physicochemical characterization of a recombinant Yersinia protein tyrosine phosphatase. *The Journal of biological chemistry* **267**, 23759-23766
42. Li, H., Xu, H., Zhou, Y., Zhang, J., Long, C., Li, S., Chen, S., Zhou, J. M., and Shao, F. (2007) The phosphothreonine lyase activity of a bacterial type III effector family. *Science* **315**, 1000-1003
43. Lifshitz, Z., Bursttein, D., Schwartz, K., Shuman, H. A., Pupko, T., and Segal, G. (2014) Identification of novel Coxiella burnetii Icm/Dot effectors and genetic analysis of their involvement in modulating a mitogen-activated protein kinase pathway. *Infect Immun* **82**, 3740-3752
44. Chen, R. E., and Thorner, J. (2007) Function and regulation in MAPK signaling pathways: lessons learned from the yeast Saccharomyces cerevisiae. *Biochim Biophys Acta* **1773**, 1311-1340
45. Koonin, E. V., and Tatusov, R. L. (1994) Computer analysis of bacterial haloacid dehalogenases defines a large superfamily of hydrolases with diverse specificity. Application of an iterative approach to database search. *Journal of molecular biology* **244**, 125-132
46. Kuznetsova, E., Nociek, B., Brown, G., Makarova, K. S., Flick, R., Wolf, Y. I., Khusnutdinova, A., Evdokimova, E., Jin, K., Tan, K., Hanson, A. D., Hasnain, G., Zallot, R., de Crey-Lagarde, V., Babu, M., Savchenko, A., Joachimiak, A., Edwards, A. M., Koonin, E. V., and Yakunin, A. F. (2015) Functional Diversity of Haloacid Dehalogenase Superfamily Phosphatases from Saccharomyces cerevisiae: BIOCHEMICAL, STRUCTURAL, AND EVOLUTIONARY INSIGHTS. *The Journal of biological chemistry* **290**, 18678-18698
47. Gohla, A., Birkenfeld, J., and Bokoch, G. M. (2005) Chronophin, a novel HAD-type serine protein phosphatase, regulates cofilin-dependent actin dynamics. *Nat Cell Biol* **7**, 21-29
Legionella phosphotyrosine phosphatase activity towards MAPKs

48. Rayapureddi, J. P., Kattamuri, C., Steinmetz, B. D., Frankfort, B. J., Ostrin, E. J., Mardon, G., and Hegde, R. S. (2003) Eyes absent represents a class of protein tyrosine phosphatases. *Nature* **426**, 295-298

49. Seifried, A., Knobloch, G., Duraphe, P. S., Segerer, G., Manhard, J., Schindelin, H., Schultz, J., and Gohla, A. (2014) Evolutionary and structural analyses of mammalian haloacid dehalogenase-type phosphatases AUM and chronophin provide insight into the basis of their different substrate specificities. *The Journal of biological chemistry* **289**, 3416-3431

50. Kim, Y., Gentry, M. S., Harris, T. E., Wiley, S. E., Lawrence, J. C., Jr., and Dixon, J. E. (2007) A conserved phosphatase cascade that regulates nuclear membrane biogenesis. *Proc Natl Acad Sci U S A* **104**, 6596-6601

51. Seifried, A., Schultz, J., and Gohla, A. (2013) Human HAD phosphatases: structure, mechanism, and roles in health and disease. *The FEBS journal* **280**, 549-571

52. Rinaldo-Matthis, A., Rampazzo, C., Reichard, P., Bianchi, V., and Nordlund, P. (2002) Crystal structure of a human mitochondrial deoxyribonucleotidase. *Nature structural biology* **9**, 779-787

53. Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., and Allen, K. N. (2006) Diversification of function in the haloacid dehalogenase enzyme superfamily: The role of the cap domain in hydrolytic phosphoruscarbon bond cleavage. *Bioorg Chem* **34**, 394-409

54. Kall, L., Krogh, A., and Sonnhammer, E. L. (2004) A combined transmembrane topology and signal peptide prediction method. *Journal of molecular biology* **338**, 1027-1036

55. Burstein, D., Amaro, F., Zusman, T., Lifshitz, Z., Cohen, O., Gilbert, J. A., Popko, 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

56. Finn, R. D., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and Bateman, A. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic acids research* **44**, D279-285

57. Zhu, Y., Hu, L., Zhou, Y., Yao, Q., Liu, L., and Shao, F. (2010) Structural mechanism of host Rab1 activation by the bifunctional Legionella type IV effector SidM/DrrA. *Proc Natl Acad Sci U S A* **107**, 4699-4704

58. Proudfoot, M., Kuznetsova, E., Sanders, S. A., Gonzalez, C. F., Brown, G., Edwards, A. M., Arrowsmith, C. H., and Yakunin, A. F. (2008) High throughput screening of purified proteins for enzymatic activity. *Methods Mol Biol* **426**, 331-341

59. 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

60. Wallden, K., Stenmark, P., Nyman, T., Flodin, S., Graslund, S., Loppnau, P., Bianchi, V., and Nordlund, P. (2007) Crystal structure of human cytosolic 5'-nucleotidase II: insights into allosteric regulation and substrate recognition. *The Journal of biological chemistry* **282**, 17828-17836

61. Jung, S. K., Jeong, D. G., Chung, S. J., Kim, J. H., Park, B. C., Tonks, N. K., Ryu, S. E., and Kim, S. J. (2010) Crystal structure of ED-Eya2: insight into dual roles as a protein tyrosine phosphatase and a transcription factor. *FASEB J* **24**, 560-569
62. Peisach, E., Selengut, J. D., Dunaway-Mariano, D., and Allen, K. N. (2004) X-ray crystal structure of the hypothetical phosphotyrosine phosphatase MDP-1 of the haloacid dehalogenase superfamily. *Biochemistry* **43**, 12770-12779

63. Fortpied, J., Maliekal, P., Vertommen, D., and Van Schaftingen, E. (2006) Magnesium-dependent phosphatase-1 is a protein-fructosamine-6-phosphatase potentially involved in glycation repair. *The Journal of biological chemistry* **281**, 18378-18385

64. Brewster, J. L., de Valoir, T., Dwyer, N. D., Winter, E., and Gustin, M. C. (1993) An osmosensing signal transduction pathway in yeast. *Science* **259**, 1760-1763

65. 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

66. Biswas, T., Yi, L., Aggarwal, P., Wu, J., Rubin, J. R., Stuecky, J. A., Woodard, R. W., and Tsodikov, O. V. (2009) The tail of KdsC: conformational changes control the activity of a haloacid dehalogenase superfamily phosphatase. *The Journal of biological chemistry* **284**, 30594-30603

67. Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D., and Aravind, L. (2006) Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *Journal of molecular biology* **361**, 1003-1034

68. Tegtmeier, N., Neddermann, M., Asche, C. I., and Backert, S. (2017) Subversion of host kinases: a key network in cellular signaling hijacked by Helicobacter pylori CagA. *Mol Microbiol* **105**, 358-372

69. Krachler, A. M., Woolery, A. R., and Orth, K. (2011) Manipulation of kinase signaling by bacterial pathogens. *J Cell Biol* **195**, 1083-1092

70. Weissenmayer, B. A., Prendergast, J. G., Lohan, A. J., and Loftus, B. J. (2011) Sequencing illustrates the transcriptional response of Legionella pneumophila during infection and identifies seventy novel small non-coding RNAs. *PLoS One* **6**, e17570

71. Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta crystallographica. Section D, Biological crystallography* **62**, 859-866

72. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography* **66**, 213-221

73. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126-2132

74. Hasegawa, H., and Holm, L. (2009) Advances and pitfalls of protein structural alignment. *Curr Opin Struct Biol* **19**, 341-348

75. Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y., and Liang, J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic acids research* **34**, W116-118

76. Gietz, R. D., and Schiestl, R. H. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature protocols* **2**, 31-34
77. Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., and Boeke, J. D. (1998) Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast 14, 115-132

78. von der Haar, T. (2007) Optimized protein extraction for quantitative proteomics of yeasts. PLoS One 2, e1078
FOOTNOTES
Results shown in this report are supported in part by a National Institutes of Health grant (GM094585) to AS through the Midwest Center for Structural Genomics and by the U. S. Department of Energy, Office of Biological and Environmental Research (contract DE-AC02-06CH11357). The work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357
### TABLES

Table 1 – X-ray diffraction data collection and refinement statistics

| Structure       | Ceg4[1-208]TEV<sub>sie</sub> | Ceg4[1-208]HOG1p<sub>native</sub> |
|-----------------|-----------------------------|-----------------------------------|
| PDB Code        | 6AOK                        | 6AOJ                              |

| Data collection |                                       |                                   |
|-----------------|----------------------------------------|-----------------------------------|
| Space group     | P21212                                | C2                                |
| Cell dimensions | a, b, c (Å)                            | 40.04, 48.03, 100.10              |
|                 | β (°)                                 | 90.0                              |
| Resolution, Å   | 40.0 – 1.88                           | 50.0 – 1.90                       |
| R<sub>merge</sub> | 0.071 (0.525)<sup>a</sup>           | 0.052 (0.259)                     |
| R<sub>pim</sub>  | 0.035 (0.266)                         | 0.031 (0.149)                     |
| I / σ(I)        | 21.4 (3.9)                            | 39.3 (4.5)                        |
| Completeness, % | 100 (99.9)                            | 97.2 (92.5)                       |
| Redundancy      | 5.1 (4.9)                             | 3.9 (3.8)                         |

| Refinement      |                                       |                                   |
|-----------------|----------------------------------------|-----------------------------------|
| Resolution, Å   | 37.17 – 1.88                           | 36.53 – 1.90                      |
| No. of unique reflections: | 16264, 1461 | 20876, 1044 |
| R-factor/free R-factor<sup>c</sup> | 15.6/19.8 (19.9/27.1) | 16.2/19.3 (22.1/28.3) |
| No. of refined atoms |  |  |
| Protein         | 1811                                   | 1738                              |
| Magnesium       | 1                                      | 1                                 |
| Chloride        | 2                                      | 1                                 |
| Water           | 269                                    | 252                               |

| B-factors       |                                       |                                   |
|-----------------|----------------------------------------|-----------------------------------|
| Protein         | 23.1                                   | 51.4                              |
| Magnesium       | 12.7                                   | 32.7                              |
| Chloride        | 23.5                                   | 38.3                              |
| Water           | 37.8                                   | 59.3                              |
| r.m.s.d.        |                                        |                                   |
| Bond lengths, Å | 0.016                                  | 0.009                             |
| Bond angles, °  | 1.267                                  | 0.954                             |

<sup>a</sup>R<sub>sym</sub> = Σ<sub>h</sub> Σ<sub>i</sub> |I<sub>i(h)</sub>| - ⟨|I(h)|⟩/Σ<sub>h</sub> Σ<sub>i</sub> |I<sub>i(h)</sub>|, where I<sub>i(h)</sub> and ⟨|I(h)|⟩ are the <i>i</i>th and mean measurement of the intensity of reflection <i>h</i>.

<sup>b</sup>Figures in parentheses indicate the values for the outer shells of the data.

<sup>c</sup>R = Σ |F<sub>p</sub>|<sup>obs</sup> - |F<sub>p</sub>|<sup>calc</sup>/Σ |F<sub>p</sub>|<sup>obs</sup>, where |F<sub>p</sub>|<sup>obs</sup> and |F<sub>p</sub>|<sup>calc</sup> are the observed and calculated structure factor amplitudes, respectively.

* = molecules in the active site cleft.
**FIGURE LEGENDS**

**FIGURE 1.** Domain organization and species Ceg4 orthologues. Phylogenetic analysis of the HAD domains of *Legionella* Dot/ICM effectors shows that Ceg4 (Lpg0096) belongs to a group of effectors in which the majority of members possess a two transmembrane (TM) region. Numerals indicate node posterior probabilities, scale bar illustrates substitutions per site.

**FIGURE 2.** Ceg4 is a phosphotyrosine phosphatase. *A,* General screening for a range of activities indicated that Ceg4 is a phosphatase. *B,* Kinetics were obtained for the generic phosphatase substrate pNPP. Error bars denote SEM values. *C,* Incubation of Ceg4 with a variety of naturally occurring phosphatase substrates, and subsequent detection of released phosphates, reveals that Ceg4 preferentially removes the phosphate from phosphotyrosine *in vitro.* See also supplemental data, Figure S1E for the full list of tested substrates. Error bars show 95% confidence intervals from triplicate data. *D,* Kinetics were obtained for the generic phosphatase substrate pNPP. Error bars denote SEM values *E,* Ceg4 dephosphorylates all tested of phosphotyrosine-containing peptides at levels 4-5 times higher than phosphoserine or phosphothreonine-containing peptides *in vitro.* Error bars denote 95% confidence intervals from triplicate data. See also Supplemental Figure S1E for all of tested peptides.

**FIGURE 3.** Crystal structure of Ceg4’s HAD-like domain provides molecular insight into phosphatase activity. *A,* Schematic of the final structural model of Ceg4. Amino acids highlighted in blue belong to the core Rossmann-like fold, gray to the cap domain and orange to the expression tag used for purification. *B,* Overall structure of the Ceg4 HAD-like domain, highlighting positions and topology of the C1-cap subdomain and the “squiggle motif”. *C,* Close-up of the Ceg4 HAD-like domain active site, indicating positions of catalytic site residues (small red spheres denote positions of water). *D,* Catalytic site residues of Ceg4 form a ~350 Å pocket, occupied by a magnesium and a chloride ion and partially covered by N20 of the cap subdomain.

**FIGURE 4.** Ceg4[1-208] crystallizes with the N-terminal fusion tag’s tyrosine bound deeply in the catalytic pocket of the adjacent molecule. *A,* Crystal packing of Ceg4[1-208]HOG1p shows the N-terminal peptide packed against the active site of the adjacent Ceg4[1-208]HOG1p molecule. *B,* Close-up examination of the HOG1p tag-Ceg4 interaction shows the tyrosine bound deeply in the catalytic pocket. *C,* Hog1 and TEV fusion tag peptides bind the active site of Ceg4[1-208] with strikingly similar conformations.

**FIGURE 5.** Ceg4 suppresses MAPK responses in *S. cerevisiae.* *A,* Ceg4[1-397] is able to reduce the Hog1-mediated high osmolarity response to high salt concentrations and the Fus3 mediated pheromone/mating response to α-factor. Removal of the C-terminal TM domain reduced Ceg4’s ability to lessen the pheromone response, but slightly increased its ability to reduce activation of the hyperosmolarity response, suggesting off-target effects. *B,* Ceg4[1-397] mutants K17A, N20A, E26A, E108A, E131A and T161A all retained wildtype levels of activity, whereas mutants D11A, V23D, F24A, Y43A, K104A, K135, and D158 all exhibited reduced ability to supress the pheromone response. See also Supplementary Figures S5 for expression testing. *C,* Location of tested mutants illustrated on the structure of Ceg4. Mutation of residues highlighted in red reduced ability to supress the pheromone response.

**FIGURE 6.** Ceg4 localization to the endoplasmic reticulum requires C-terminal TM domain and dampens activation of human MAPK p38 *in vivo.* *A,* Ceg4 localization is dependent on C-terminal TM domains. N-terminal GFP fusions of Ceg4[1-397], Ceg4[1-397]D9A and Ceg4[208-397] expressed in HeLa cells show distinct subcellular localisation that is lost in Ceg4[1-207] lacking transmembrane regions. *B,* GFP-Ceg4[1-397]D9A colocalizes with ER-tracker red dye. See also Supplemental figure S6. *C,* Activation and phosphorylation of p38 MAPK was achieved by incubating HEK293 cells expressing Ceg4[1-397] or Ceg4[1-397]D9A with TPA or anisomycin. While total levels of p38 remained unchanged, cells expressing Ceg4[1-397] showed reduced levels of phospho-p38 compared to D9A mutant constructs.
Legionella phosphotyrosine phosphatase activity towards MAPKs
Figure 1
**Legionella** phosphytyrosine phosphatase activity towards MAPKs

**Figure 2**

| Substrate                | ΔOD  |
|--------------------------|------|
| Phosphatase              | 2.03 |
| Protease                 | 0.10 |
| Phosphodiesterase        | 0.00 |
| Thioesterase             | 0.00 |
| β-maltosidase            | 0.01 |
| α-maltosidase            | 0.00 |
| Cellobiosidase           | 0.01 |
| Esterase                 | 0.00 |
| Glucopyranosidase        | 0.00 |

**A**

**B**

\[ V_{\text{max}} = 180 \]
\[ K_m = 0.0899 \]

**C**

| Substrate     | mM phosphate min⁻¹mg⁻¹ |
|---------------|------------------------|
| β-glucosyl-1P | 100                    |
| α-glucosidase | 100                    |
| PAP (3'-5' ADP)| 100                   |
| 2,5-ADP       | 100                    |
| GTP           | 100                    |
| GMP-Morpholidate| 100                   |
| Thiamine-PP   | 100                    |
| Phosphorylamine| 100                   |
| p-Tyrosine    | 100                    |
| α-keto-glutarate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
| 2-phosphoglycerate| 100                   |
| 3-phosphoglycerate| 100                   |
Figure 3
Figure 4

A

Ceg4[1-208]^{HOG1p}

B

Ceg4[1-208]^{HOG1p}

C

Ceg4[1-208]^{HOG1p}

Ceg4[1-208]^{TEV site}
Legionella phosphotyrosine phosphatase activity towards MAPKs

Figure 5

A

+KCl (hyperosmolarity), Mean BFP Fluorescence (Stl2-BFP)

|                     | Mean Fluorescence |
|---------------------|-------------------|
| Empty Plasmid       | 4.025             |
| Ceg4[1-397]         | 2.325             |
| Ceg4[1-208]         | 1.805             |

+a-factor (Pheromone), Mean GFP Fluorescence (Fus1-GFP)

|                     | Mean Fluorescence |
|---------------------|-------------------|
| Empty Plasmid       | 1.2335            |
| Ceg4[1-397]         | 0.592             |
| Ceg4[1-208]         | 0.9525            |

B

+a-factor, Mean Fluorescence

|        | Mean Fluorescence |
|--------|-------------------|
| Empty Plasmid | 1.2335             |
| D11A    | 1.044             |
| K17A    | 0.548             |
| S18A    | 0.7755            |
| N20A    | 0.5305            |
| V23D    | 1.2985            |
| F24A    | 0.9885            |
| E26A    | 0.5575            |
| Y43A    | 1.091             |
| K104A   | 1.159             |
| E108A   | 0.6985            |
| E131A   | 0.667             |
| K135A   | 1.0175            |
| D158A   | 1.0595            |
| T161A   | 0.618             |
| D162A   | 0.8145            |
Figure 6

A

Ceg4  Ceg4_{D9A}  Ceg4[1-207]  Ceg4[208-397]

B

DAPI  Ceg4_{D9A}

ER-tracker  Merge

C

|              | Ceg4 | Ceg4_{D9A} | Anisomycin | TPA |
|--------------|------|------------|------------|-----|
| Ceg4        | +    | -          | +          | -   |
| Ceg4_{D9A}  | -    | +          | -          | +   |
| Anisomycin  | +    | +          | -          | -   |
| TPA         | -    | -          | +          | +   |

phospho-p38

p38
The Legionella pneumophila effector Ceg4 is a phosphotyrosine phosphatase that attenuates activation of eukaryotic MAPK pathways
Andrew T. Quaile, Peter J. Stogios, Olga Egorova, Elena Evdokimova, Dylan Valleau, Boguslaw Nocek, Purnima S. Kompella, Sergio Peisajovich, Alexander F. Yakunin, Alexander W. Ensminger and Alexei Savchenko

J. Biol. Chem. published online January 4, 2018

Access the most updated version of this article at doi: 10.1074/jbc.M117.812727

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts