**Bacterial PhyA protein-tyrosine phosphatase-like myo-inositol phosphatases in complex with the Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃ second messengers**

Received for publication, March 29, 2017, and in revised form, August 24, 2017. Published, Papers in Press, August 27, 2017, DOI 10.1074/jbc.M117.787853

Lisa M. Bruder①, Robert J. Gruninger①, Colyn P. Cleland①, and Steven C. Mosimann①,②

From the ①Department of Chemistry and Biochemistry, University of Lethbridge, Lethbridge AB T1K 3M4, Canada and the ②Lethbridge Research Centre, Agriculture and Agri-Foods Canada, Lethbridge AB T1J 4B1, Canada

Edited by Wolfgang Peti

---

**myo-Inositol phosphates (IPs)** are important bioactive molecules that have multiple activities within eukaryotic cells, including well-known roles as second messengers and cofactors that help regulate diverse biochemical processes such as transcription and hormone receptor activity. Despite the typical absence of IPs in prokaryotes, many of these organisms express IPases (or phytases) that dephosphorylate IPs. Functionally, these enzymes participate in phosphate-scavenging pathways and in plant pathogenesis. Here, we determined the X-ray crystallographic structures of two catalytically inactive mutants of protein-tyrosine phosphatase-like myo-inositol phosphatases (PTPLPs) from the non-pathogenic bacteria *Selenomonas ruminantium* (PhyAsr) and *Mitsukella multacida* (PhyAmm) in complex with the known eukaryotic second messengers Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃. Both enzymes bound these less-phosphorylated IPs in a catalytically competent manner, suggesting that IP hydrolysis has a role in plant pathogenesis. The less-phosphorylated IP binding differed in both the *myo-*inositol ring position and orientation when compared with a previously determined complex structure in the presence of *myo-*inositol-1,2,3,4,5,6-hexakisphosphate (InsP₆ or phytate). Further, we have demonstrated that PhyAsr and PhyAmm have different specificities for Ins(1,2,4,5,6)P₇, have identified structural features that account for this difference, and have shown that the absence of these features results in a broad specificity toward Ins(1,2,4,5,6)P₇. These features are main-chain conformational differences in loops adjacent to the active site that include the extended loop prior to the penultimate helix, the extended Ω-loop, and a β-hairpin turn of the Phy-specific domain.

**myo-Inositol phosphates (IPs)** containing between one and eight phosphoryl groups are ubiquitous in eukaryotic species and have diverse biological activities (1). The most abundant, *myo-*inositol-1,2,3,4,5,6-hexakisphosphate (InsP₆ or phytate), has multiple important roles in eukaryotic cellular processes including the regulation of plant hormone receptors, DNA repair, RNA processing, mRNA export, plant development, apoptosis, and pathogenicity (2–9). Less-abundant, less-phosphorylated IPs have also been implicated in many important biological processes including second messenger activities (1, 10). For example, the role of Ins(1,4,5)P₃ and Ins(1,3,4,5)P₄ in Ca²⁺ mobilization has been well-characterized (1, 11). Ins(1,4,5)P₃ stimulates the release of Ca²⁺ from the endoplasmic reticulum resulting in further cellular responses, and Ins(1,3,4,5)P₄ acts to increase sensitivity and generate a longer-lasting signal. In general, highly phosphorylated IPs (InsP₆ and InsP₇) serve as cofactors, whereas less-phosphorylated IPs are utilized as second messengers in signal transduction pathways (10). In contrast to eukaryotic cells, IPs are typically absent from prokaryotes (12), despite the presence of enzymes specific for IPs (IPases) in many organisms. Known microbial IPases have few characterized roles in prokaryotes (13). Most serve as phosphate-scavenging proteins, and more recently others have been implicated in pathogenesis (4, 14, 15).

Microbial IPases that specifically hydrolyze the C3- and C4-phosphoryl groups of InsP₆ (d-*myo* numbering) are common, and experimental structures of both classes of enzyme are available (16–18). Structural features that determine C3-phosphoryl group specificity of IPases (3-phytases) have been identified for two distinct enzyme families: protein-tyrosine phosphatase-like myo-inositol phosphatase (PTPLPs) and histidine-acid phosphatases (16, 17). In each of these cases, the InsP₆ substrate adopts its lowest energy conformation, and steric interactions restrict the only axial phosphoryl group (C2) to a specific location adjacent to the scissile phosphoryl group (16, 17).

PTPLPs and histidine-acid phosphatases hydrolyze InsP₆ to produce *myo-*inositol-2-monokisphosphate (Ins(2)P) and phosphatase-like *myo-*inositol phosphatase; InsP₆, *myo-*inositol-1,2,3,4,5,6-hexakisphosphate; PhyAsr, *S. ruminantium* PTPLP (PhyAsr); PhyAmm, *M. multacida* PTPLP (PhyAmm); Ins(1,3,4,5)P₃, *myo-*inositol-1,3,4,5-tetraakisphosphate; Ins(1,4,5)P₃, *myo-*inositol-1,4,5-trikisphosphate; Ins(2)P, *myo-*inositol-2-monokisphosphate; LppA, *L. pneumophila* PTPLP (PhyA); GA-loop, general acid loop; P-loop, phosphate-binding loop; LSQ, least squares; HPIC, high-performance ion chromatography; PDB, Protein Data Bank; BME, β-mercaptoethanol; r.m.s.d., root-mean-square deviation.
phosphate after prolonged incubation (18–22). Characterized PTPLPs have been shown to hydrolyze InsP₆ to Ins(2)P via different dephosphorylation pathways (19–22). An example is PhyA from *Selenomonas ruminantium* (*PhyAsr*), which first removes the C3-phosphoryl group (P3) followed (in order) by P1, P6, P5, and P4 (19). PTPLPs such as PhyA from *Mitsuokella multacida* (*PhyAmm*) and *Legionella pneumophila* (*LppA*) have demonstrated activity toward the Ins(1,4,5)P₃ second messenger (23, 24). It is likely that these enzymes can also hydrolyze additional IP second messengers and that this activity is shared by other PTPLPs.

Here we have presented the first structures of bacterial PTPLPs in complex with eukaryotic IP second messengers, representing crystallographic evidence that PTPLPs bind and hydrolyze these compounds. These IPs bind in a different conformation than InsP₆ while utilizing a subset of the previously determined phosphoryl-binding sites (17). Further, we have identified several variable loops adjacent to the active site that influence their dephosphorylation pathways. Consequently, this work both refines and extends previous studies aimed at understanding the function and specificity of these enzymes and likely applies to other IPase families.

**Results**

**Active sites of PTPLPs are pre-formed**

The overall fold of PTPLPs is composed of a catalytic α-β-α sandwich PTP domain and an antiparallel α-β sandwich Phy domain. The active site includes the general acid loop (GA-loop) and phosphate-binding loop (P-loop) from the catalytic PTP domain as well as residues from both domains that participate in substrate binding (17). The fold and active site of PhyAsr are shown in Fig. 1A. Least squares (LSQ) superpositions of PhyAsr crystal structures in the presence and absence of ligand yielded r.m.s.d. of ~0.2 Å (over 1248 main-chain atoms), clearly demonstrating that there are no large-scale main-chain conformational changes associated with ligand binding (supplemental Table S1) (17, 19). The same is true within the active site, as the main- and side-chain conformations of the residues alone or in complex with InsP₆, Ins(1,3,4,5)P₄ or Ins(1,4,5)P₃ are essentially identical (supplemental Table S2).

The PhyAmm monomer is an example of a tandemly repeated IPase that contains two copies of the catalytic PTP and Phy domains (Fig. 1B). Although the N- and C-terminal repeats of PhyAmm have different IP substrate specificities, they are both active toward the Ins(1,4,5)P₃ second messenger (23). Presented here is the first complex structure of PhyAmm: PhyAmmC252S/C548S in complex with Ins(1,3,4,5)P₄, a related second messenger. The PhyAmmC252S/C548S-Ins(1,3,4,5)P₄ structure was solved in a different space group (P1) than PhyAmm without ligand (P2). Between the two structures, the individual repeats of PhyAmm are nearly identical as judged by LSQ superpositions (supplemental Table S3) (23). Small movements in the linker region between repeats subtly alter the relative orientation of the tandem repeats and account for the bulk of the observed differences between the individual monomers and dimers. As reported previously, the dimer interface of PhyAmm is almost exclusively formed between the N-terminal repeats, and these interactions are an extensive network of hydrogen bonds, salt bridges, and van der Waals contacts (23). As seen with PhyAsr, the C-terminal active site of PhyAmmC252S/C548S is virtually identical in the presence or absence of Ins(1,3,4,5)P₄ (supplemental Table S4). Taken together, the structures of both PhyAsr and PhyAmm support previous suggestions that the PTPLP family of enzymes have pre-formed active sites (17).

**Alternate binding modes of IPs**

The electron density for bound ligand is clearly visible in the initial 2mFᵦ – DFᵦ (1.5 σ) and mFᵦ – DFᵦ (3.5 σ) maps of each PhyAsrC252S structure. Refined electron density maps (2mFᵦ – DFᵦ) for the Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃ ligands at 1
Bacterial PTPLPs in complex with second messengers

Figure 2. Clear electron density for the phosphoryl groups and C2-hydroxyl allows for an unambiguous fit of the ligands and places the C1-phosphoryl group (P1) above the cysteine to serine mutations at positions 252 (PhyAsr) and 548 (PhyAmm). The refined 2mFo − DFc electron density is contoured at 1σ (blue mesh) for PhyAsrC252S in complex with Ins(1,3,4,5)P4 (A) and Ins(1,4,5)P3 (B) and for the PhyAmmC252S/C548S C-terminal repeat in complex with Ins(1,3,4,5)P4 (C). Ligand and protein are shown as sticks; with oxygen shown in red, nitrogen in blue; phosphorus in orange; and carbon in gray.

σ are shown in Fig. 2, A and B. Obvious electron density for the axial C2-hydroxyls and each phosphoryl group clearly identify the conformation of each ligand bound within the active site. We determined two separate PhyAsrC252S Ins(1,3,4,5)P4 structures that differ in how the ligand was soaked: 1 mM Ins(1,3,4,5)P4 for 1 h and 10 mM Ins(1,3,4,5)P4 for 15 min. The resulting structures are virtually identical, apart from an additional Ins(1,3,4,5)P4 conformer (0.25 occupancy) modeled in the 1 mM soak, as it reduces the difference density within the active site. The major conformer of the 1 mM soak is identical to the single conformer modeled at full occupancy in the 10 mM soak. Interestingly, PhyAsrC252S binds Ins(1,4,5)P3 in a position identical to that of Ins(1,3,4,5)P4 less one phosphate (Fig. 3).

The modeled IPs adopt the lowest-energy chair conformation with five equatorial hydroxyl/phosphoryl groups and an axial C2-hydroxyl. The C1-phosphoryl group of the ligands are positioned for hydrolysis in the P-loop phosphoryl-binding site by forming extensive interactions with the P-loop (Fig. 3 and Table 1) (17, 25). The P-loop (residues 252–259), GA-loop (residues 222–225), and Lys-312 are the only interactions that originate from the catalytic PTP domain. The remaining contacts are mediated by residue side chains derived from the Phy domain (residue ranges of 51 to 59 and 136 to 203) and the Phy-specific extension of the penultimate helix (residues 291–307). The less-phosphorylated IPs utilize a subset of the phosphoryl-binding sites identified previously in the PhyAsrC252S InsP6 structure (17).

The myo-inositol rings of the less-phosphorylated IP complexes have different relative orientations within the active site when compared with the PhyAsrC252S InsP6 structure. In particular, the myo-inositol rings are rotated by 180°, resulting in the opposite face contacting the enzyme. For the C1-phosphoryl group to maintain contact with the P-loop, the rotated myo-inositol rings tilt toward the GA-loop (Fig. 4). Overall, the IPs shift by more than 1.2 Å (center of mass to center of mass; supplemental Table S5) and fill the space occupied by ordered solvent in the InsP6 complex structure.

The N- and C-terminal active sites of PhyAmmC250S/C548S are non-equivalent and have different substrate specificities (23). The C-terminal active site of PhyAmm is highly active toward InsPγ, and 12 of the 14 residues contacting the ligand are conserved when compared with PhyAsr (Table 1). Not unexpectedly, Ins(1,3,4,5)P4 binding within the PhyAmmC250S/C548S C-terminal active site is nearly identical to that observed in the PhyAsrC252S Ins(1,3,4,5)P4 structure. In contrast, the N-terminal active site of PhyAmm does not bind ligand in a catalytically competent manner. Instead, an inorganic phosphate is bound by the catalytic P-loop, and a partially occupied (0.6) Ins(1,3,4,5)P4 is bound at a novel site more than 8 Å (phosphorus to phosphorus) from the inorganic phosphate (Fig. 3C). The N-terminal active site has residue substitutions as well as a two-residue insertion. The insertion is in the Phy-specific domain and generates a larger α-turn (residues 182–188) that extends into the active site. The equivalent region of PhyAsr and the C-terminal repeat of PhyAmm have smaller β-hairpin turns. The mutations in the N-terminal repeat of PhyAmm reduce the positive electrostatic surface potential and introduce additional steric limitations.

PTLP specificity differences

The active sites of PhyAsr and the C-terminal repeat of PhyAmm have conserved residues that contact the ligands (Fig. 3 and Table 1) (23). However, they contain significant differences in the main-chain conformation of the three loops that contribute to the active sites of these enzymes: the extended loop prior to the penultimate helix (PhyAsr(287–305) and PhyAmm(583–600)); the extended Ω-loop (PhyAsr(73–102) and PhyAmm(367–398)); and a β-hairpin turn of the Phy-specific domain (PhyAsr(186–189) and PhyAmm(482–485)) (Fig. 5). As the observed differences in these loop conformations may affect substrate access to the active site, we determined the InsPγ hydrolysis pathway for these enzymes. The time courses of the InsPγ hydrolysis products separated by high-performance ion chromatography (HPIC) clearly demonstrate that even though the enzymes are highly specific for the C3-phosphoryl group of InsPγ, they have different specificities for the Ins(1,2,4,5,6)P5 substrate (Fig. 6, A and B). Comparisons with standard chromatograms demonstrate that PhyAmm is specific for the C4-phosphoryl group of Ins(1,2,4,5,6)P5, whereas PhyAsr hydrolyzes the C1-phosphoryl (supplemental Fig. S1).
To further assess the contribution of these variable loops to the substrate specificity of PTPLPs, we examined the structure and hydrolysis pathway of PhyA from *Bdellovibrio bacteriovorus* (PhyAbb, previously Bd1204). PhyAbb is one of the smallest known PTPLPs and has large deletions in the extended loop prior to the penultimate helix, the extended $\gamma/H\_9024$-loop, and the Phy-specific domain (Fig. 5) (26). As a result, the active site of PhyAbb is more open and accessible than in PhyAsr or PhyAmm. Taken together, it is expected that PhyAbb has an altered pathway and would have a broader specificity for IPs than either PhyAsr or PhyAmm, which is confirmed by the hydrolysis pathway (Fig. 6C). Although PhyAbb is also specific for the C3-phosphoryl group of $\text{InsP}_6$, making it a 3-phytase, PhyAbb produces four different $\text{InsP}_4$, which include the PhyAsr ($\text{Ins}(2,4,5,6)\text{P}_4$) and PhyAmm ($\text{Ins}(1,2,5,6)\text{P}_4$) products. The much broader specificity of PhyAbb for IPs indicates that the loops that contribute to the active site influence substrate specificity.

**Discussion**

***myo-Inositol ring movements compensate for PTPLP active-site rigidity***

There are no changes in the conformation of the active site upon binding of IP ligands to either PhyAsr or PhyAmm (supplemental Tables S2 and S4). The apparent rigidity of these
**Bacterial PTPLPs in complex with second messengers**

**Table 1**

| Residue | Site | Phos/OH Distance (Å) | Phos/OH Distance (Å) | Residue | Site | Phos/OH Distance (Å) |
|---------|------|----------------------|----------------------|---------|------|----------------------|
| Arg-57  | P$_s$ | 2.80/3.17            | 2.80/3.17            | Arg-351 | P$_s$ | 3.16/2.85/2.64       |
| Asp-153 | P$_s$ | 2.60                 | 2.60                 | Asp-449 | P$_s$ | 3.24                 |
| Lys-189 | P$_s$ | 2.96                 | 2.96                 | Lys-485 | P$_s$ | 3.24                 |
| Asp-223 | P$_s$ | 3.16/2.81            | 3.16/2.81            | Asp-519 | P$_s$ | 3.07/3.03            |
| His-224 | P$_s$ | 2.95                 | 2.95                 | His-520 | P$_s$ | 3.02                 |
| Ser-252 | P$_s$ | 3.4                  | 3.4                  |          |      |                      |
| Lys-305 | P$_s$ | 2.51                 | 2.51                 | Ser-548  | P$_s$ | 3.67                 |
| Glu-253 | P$_s$ | 3.06                 | 3.06                 | Glu-549  | P$_s$ | 3.22                 |
| Ala-254 | P$_s$ | 3.17                 | 3.17                 | Ala-550  | P$_s$ | 3.15                 |
| Gly-255 | P$_s$ | 2.91                 | 2.91                 | Gly-551  | P$_s$ | 3.15                 |
| Val-256 | P$_s$ | 2.69                 | 2.69                 | Val-552  | P$_s$ | 3.15                 |
| Gly-257 | P$_s$ | 3.16                 | 3.16                 | Gly-553  | P$_s$ | 3.34                 |
| Arg-258 | P$_s$ | 2.92/3.00/2.99       | 2.92/3.00/2.99       | Arg-554  | P$_s$ | 3.11/2.76            |
| Lys-305 | P$_s$ | 2.52/2.51            | 2.52/2.51            | Lys-600  | P$_s$ | 3.09/3.13/3.15       |
| Tyr-309 | P$_s$ | 2.82                 | 2.82                 | Tyr-604  | P$_s$ | 2.95/2.99            |

This work identifies a significant shift and a 180° rotation of the myo-inositol ring in the less-phosphorylated Ins$_{1,3,4,5}P_4$ and Ins$_{1,4,5}P_3$ complex structures compared with the PhyAsrC252S-Ins$_{6}$ structure (Fig. 4). The ring shift and rotation allows smaller substrates to utilize a different subset of phosphoryl-binding sites. For example, in the PhyAsrC252S-Ins$_{6}$ structure and the less-phosphorylated IP complexes, hydrogen bonds to the C5-phosphoryl group originate from opposite sides of Tyr-309 (Fig. 4). The myo-inositol ring shift also allows phosphoryl groups not adjacent to the scissile phosphoryl group to occupy the P$_s$ site. Importantly, simple modeling studies based on the observed ring shift suggest that, provided the P$_s$ site contains a hydroxyl, the P$_s$-binding site can accommodate equatorial phosphoryl groups adjacent to the scissile phosphate. Alternatively, both the P$_s$ and P$_o$ sites may be able to accommodate equatorial phosphoryl groups if the ring is allowed to adopt ring orientations that are intermediate to those observed in the Ins$_{6}$ and less-phosphorylated IP complex structures. These observations are sufficient to rationalize the formation of minor products of the PhyAsr hydrolysis pathway and suggest that PTPLPs may be able to hydrolyze a wide range of less-phosphorylated IPs.

**Structural determinants of the substrate specificity of PTPLPs**

PhyAsr and PhyAmm hydrolyze Ins$_{1,2,4,5,6}P_5$ to different Ins$_{4}$ products despite having the same specificity for the C3-phosphoryl group of the Ins$_{6}$ and binding Ins$_{1,3,4,5}P_4$ nearly identically (Figs. 3 and 6) (19). The residues that contact the ligands are conserved between PhyAsr and PhyAmm except for residue substitutions in the P-loop (Table 1). In the PhyAmm C-terminal repeat, Glu-549 immediately follows the catalytic Cys-548 (Ser-548 in our structure), whereas the equivalent residue in the PhyAsr is Glu-253. The Glu-549 side chain of PhyAmm is directed toward the active site and may directly or indirectly contact highly phosphorylated IPs providing an additional favorable electrostatic interaction in the P$_s$ site. In contrast, both in the presence and absence of ligand, the side
chain of Glu-253 (PhyAsr) is directed away from the active site. The charge difference and the spatial orientation in the P<sub>a</sub>/ site are capable of influencing the difference in the activity and specificity of these enzymes but do not fully account for the pathway divergence.

The extended loop prior to the penultimate helix, extended Ω-loop, and β-hairpin in the Phy domain are loops around the active site that are different in PhyAsr and PhyAmm and contribute to the divergent pathways (Fig. 5). The main-chain conformational differences in these connecting segments result from residue substitutions, insertions/deletions, and the distinct homodimers formed by each enzyme. The extended loop prior to the penultimate helix of PhyAsr located on the P<sub>a</sub>/P<sub>b</sub> side of the active site contains a single residue insertion, and the C-terminal end of the loop folds into the active site of the enzyme. In contrast, the equivalent loop in PhyAmm is pulled away from the active site and participates in a 2-fold symmetric, β-hairpin turn, resulting in an accessible (RA) active site on the P<sub>a</sub>/P<sub>b</sub> side. Additionally, the position of the extended loop prior to the penultimate helix of PhyAbb is similar in position to the equivalent loop of PhyAmm, and the extended Ω-loop is deleted, leaving the P<sub>a</sub>/P<sub>b</sub> side more accessible (RA). As a result, PhyAbb produces four different Ins<sub>P</sub> products in contrast to PhyAsr and PhyAmm.

Figure 5. Variable loops implicated in the substrate specificity of PTPLPs. A, stereo view of the superposition of PhyAsrC252S.Ins(1,3,4,5)P<sub>a</sub> (red), PhyAmmC250S/CS48S.Ins(1,3,4,5)P<sub>a</sub> (blue), and PhyAbb (green, PDB 4NX8) as a ribbon diagram with the ligand as sticks. The variable loops (colored segments) that influence substrate specificity include the extended loop prior to the penultimate helix, the extended Ω-loop, and the β-hairpin turn within the Phy-specific domain. B, the PhyAsr (blue) active site is relatively occluded (RO) on the P<sub>a</sub>/P<sub>b</sub> side and relatively accessible (RA) on the P<sub>a</sub>/P<sub>b</sub> side. C, in the case of PhyAmm (red), the relatively occluded and relatively accessible sides are reversed. D, PhyAbb (green) has a 13-residue deletion, which removes the loop that contains the β-hairpin turn, resulting in an accessible (RA) active site on the P<sub>a</sub>/P<sub>b</sub> side.

Taken together, the residue substitution and main-chain conformational differences indicate that the PhyAsr active site is relatively occluded on the P<sub>a</sub>/P<sub>b</sub> side of the active site and relatively accessible on the P<sub>a</sub>/P<sub>b</sub> side (Fig. 5). In the case of the PhyAmm active site, the converse is true. This is consistent with the specificity of PhyAsr for the C1-phosphoryl group of Ins(1,2,4,5,6)P<sub>α</sub>, as it would place the C3-hydroxyl on the relatively occluded side of its active site. Likewise, the specificity of PhyAmm for the C4-phosphoryl group of Ins(1,2,4,5,6)P<sub>β</sub> would place the C3-hydroxyl on the opposite side of its active site, which corresponds to its occluded side.

The structure and pathway of PhyAbb supports the occlusion theory as an explanation of the divergent pathways of PhyAsr and PhyAmm. PhyAbb lacks the β-hairpin in the Phy-specific domain on the P<sub>a</sub>/P<sub>b</sub> side of the active site, leaving it more accessible than in either PhyAsr or PhyAmm (Fig. 5). Further, the deletions of the extended Ω-loop, in the extended loop prior to the penultimate helix, and of two turns of the penultimate helix result in a more accessible P<sub>a</sub>/P<sub>b</sub> side of the active site. The result is a more open and accessible active site.
that allows Ins(1,2,4,5,6)P₅ to bind in multiple orientations and gives rise to four distinct InsP₄.

\textbf{Ins(1,3,4,5)P₄ is not bound at the catalytic site of the PhyAmm N-terminal repeat}

The PhyAmm N-terminal repeat shares 36 and 34% sequence identity with the C-terminal repeat and PhyAsr, respectively (23). Not surprisingly, the main-chain conformation of the N-terminal active site shares a high degree of similarity with both PhyAsr and the C-terminal repeat of PhyAmm (0.71 and 0.28 Å, respectively, 128 atoms). Despite the closely similar main-chain conformations, in our structure only the C-terminal repeat binds Ins(1,3,4,5)P₄ in a catalytically competent manner. Differences between the N- and C-terminal repeat active sites have been discussed previously (23). At present, there is no clear and unambiguous rationale for the lack of binding of Ins(1,3,4,5)P₄ to the N-terminal repeat P-loop. We note that Ins(1,3,4,5)P₄ is not a natural substrate for this enzyme, and the observed binding sterically prevents another Ins(1,3,4,5)P₄ from binding to the P-loop. Further, the natural InsP₄ substrates, which contain a C₂-phosphoryl, cannot bind in the same manner because of sterical clashes involving this axial phosphoryl group. This suggests that what we observed was a binding site that is preferred by non-native substrates. Alternatively, the structure was produced by soaking the substrate into a pre-formed crystal, which may prevent binding if there is an induced fit associated with IP binding to the N-terminal active site.

\textbf{Biological implications}

PhyAsr, PhyAmm, and PhyAbb all function as 3-phytases and yet have different specificities for Ins(1,2,4,5,6)P₅, thus producing alternate InsP₄. In the case of PhyAsr and PhyAmm, we were unable to predict their different specificities by identification of residues directly interacting with the substrates, as they are essentially identical. This suggests that substrate specificity is influenced by structural features that do not directly interact with the bound ligand. We identified three variable loops unique to PTPLPs that alter substrate access to the PhyAsr and the PhyAmm C-terminal repeat active sites, which may explain their observed specificities. This role of the variable loops is supported by our ability to predict and demonstrate that PhyAbb, an enzyme with large deletions in these loops, has a broad specificity toward IP substrates.

The activity of PTPLPs toward second messengers has been demonstrated, and this work confirms that two second messengers, Ins(1,3,4,5)P₄ and Ins(1,4,5)P₃, are capable of binding to the PTPLP active sites in a catalytically competent manner (23, 24). Further, various IPases have been demonstrated as important for the function and survival of pathogenic bacteria in host systems (4, 14, 15). They function by either providing the phosphate for growth or by derangement of the host phosphatidylinositol signaling pathway (4, 14). This suggest that PTPLP virulence factors disrupt phosphatidylinositol signaling pathway (4, 14). This suggest that PTPLP virulence factors disrupt phosphatidylinositol or inositol phosphate signaling pathways as opposed to phosphotryosine-mediated pathways (27, 28, 15).

\textbf{Experimental procedures}

\textbf{Expression and purification}

The phyA genes of the \textit{S. ruminantium} (phyAsrC252S), \textit{M. multacida} (phyAmmC250S/C548S), and \textit{B. bacteriovorus} HD100 (PhyAbb, previously Bd1204), minus the putative signal peptide, were previously cloned into the NdeI site of the pET28b\textsuperscript{Kan} expression vector (EMD Biosciences) (19, 23, 26). The resulting PhyAsrC252S and PhyAmmC250S/C548S proteins are catalytically inactive as a result of the cysteine-to-serine mutation (an isosteric substitution). All contained an N-terminal His\textsubscript{6} tag and were produced and purified as described previously (17, 19, 23, 26). PhyAsrC252S was dialyzed into 20 mM ammonium bicarbonate (pH 8.0) and lyophilized; PhyAmmC252S/C548S was dialyzed into 100 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1 mM β-mercaptoethanol (BME), and 0.1 mM EDTA (pH 8.0) followed by the addition of glycerol to...
and crystals were grown in 10% w/v PEG 8000, 100 mM Tris-
4.5 mg/ml using a Millipore Ultracel 10-kDa centrifugal filter,
PhyAmmC250S/C548S protein solutions were concentrated to
and crystals were grown as described previously (29).
PhyAsrC252S protein solutions were prepared at 20 mg/ml,
was supplemented with 100 mM glycerol. Following a 24-h equilibration, the reservoir solution
Bioscience) or 10 mM (15 min) Ins(1,4,5)P3 (Sigma-Aldrich).
20% v/v; and PhyAbb was dialyzed into 50 mM sodium acetate
(pH 5.0), 300 mM NaCl, 5 mM BME, and 0.1 mM EDTA (pH 8.0).
The protein was used immediately or flash-frozen and stored at
Values in parentheses are for the highest resolution shell.

| Protein                | PDB code    | Diffraction data (Å) | Wavelength (Å) | Ramachandran distribution |
|------------------------|-------------|----------------------|----------------|---------------------------|
| PhyAsrC252S/Ins(1,3,4,5)P4 | 4WTY        | 4WU2                 | 4WU3           |                          |
| PhyAmmC250S/Ins(1,4,5)P3 |             |                      |                |                          |
| PhyAmmC250S/C548S/Ins(1,3,4,5)P4 |         |                      |                |                          |

20% v/v; and PhyAbb was dialyzed into 50 mM sodium acetate (pH 5.0), 300 mM NaCl, 5 mM BME, and 0.1 mM EDTA (pH 8.0). The protein was used immediately or flash-frozen and stored at 193 K.

**Crystallization**

Crystallization experiments were conducted at room temperature using sitting-drop vapor diffusion with drop ratios of 2 µl of protein solution to 2 µl of reservoir solution. PhyAsrC252S protein solutions were prepared at 20 mg/ml, and crystals were grown as described previously (29). PhyAmmC250S/C548S protein solutions were concentrated to 4.5 mg/ml using a Millipore Ultrace 10-kDa centrifugal filter, and crystals were grown in 10% w/v PEG 8000, 100 mM Tris-HCl (pH 8.0), 1 mM BME, 4% v/v ethylene glycol, and 20% v/v glycerol. Following a 24-h equilibration, the reservoir solution was supplemented with 100 µl of glycerol. After 30 days, PhyAsrC252S grew rod-like crystals with approximate dimensions of 30 × 30 × 100 µm, and after 10 days PhyAmmC250S/C548S grew rod-like crystals with approximate dimensions of 100 × 100 × 500 µm. In each case, the crystals were soaked in mother liquor supplemented with 1 mM (60 min; see supplemental material) or 10 mM (15 min) Ins(1,3,4,5)P4 (Echelon Bioscience) or 10 mM (15 min) Ins(1,4,5)P3 (Sigma-Aldrich).

**Data collection and image processing**

Diffraction data (λ = 0.97934 Å) was collected from frozen crystals (100 K) using a Rayonix MX300 CCD detector at beamline 08ID-1 located at the Canadian Light Source (Saskatoon, Canada). The space group and unit cell parameters of the PhyAsrC252S crystals in complex with ligand are equivalent to those of the PhyAsrC252S/InsP6 structure (PDB 3MMJ), whereas the PhyAmmC250S/C548S/Ins(1,3,4,5)P4 crystals have a novel P1 unit cell. All diffraction image data were processed interactively with MOSFLM prior to scaling and merging within AIMLESS of the CCP4 program suite, version 6.3.0 (30–33). Data collection statistics are shown in Table 2 and supplemental Table S6.

**Structure refinement and model validation**

Phases derived from the PhyAsrC252S/InsP4 structure (PDB 3MMJ) and wild-type PhyAmm (PDB 3F41) were used to solve the structures by isomorphous replacement and molecular replacement (MOLREP), respectively. The refined structures have continuous electron density for main-chain atoms of amino acids 33–346 of PhyAsrC252S and 46–636 of PhyAmmC250S/C548S, with the remaining residues located at the termini assumed to be disordered. This includes the N-terminal histidine tag of both proteins, residues 28–32 of PhyAsrC252S and residues 31–45 of PhyAmmC250S/C548S. Refinement was performed using REFMAC, version 5.7, within the CCP4 program suite, and interactive fitting of the models to the electron density was performed in COOT, version 0.6.2 (31,
Bacterial PTPLPs in complex with second messengers

34). PROCHECK and structure validation tools with COOT were used throughout refinement to assess the stereochemistry of the model (35). Unless indicated otherwise, the figures were prepared with CCP4mg, version 2.10.8 (36). The key data processing and refinement statistics for each PhvAsrC252S and PhvAmmC250S/C548S complex structures are presented in Table 2 and supplemental Table S1.

Structure analysis

These and previously determined structures were compared by LSQ superposition using LSQKAB from the CCP4 program suite (31). The main-chain atoms of residues 35–346 of PhvAsrC252S, and residues 47–342 and 343–636 for the N- and C-terminal repeats of PhvAmmC250S/C548S, respectively, were used in overall fold comparisons. Active-site comparisons were made using residues 56–58, 152–154, 189–190, 221–226, 249–262, 304–309, and 311–313 of PhvAsrC252S and residues 448–450, 484–486, 517–522, 545–558, 584–586, and 599–605 of the PhvAmmC250S/C548S C-terminal repeat. Differences in the relative position of the myo-inositol ring of bound ligands of PhvAsrC252S were calculated using the superposed structure coordinates and GEOMCALC in the CCP4 program suite (31, 34).

Identification of hydrolysis products

Hydrolysis of 5 mM InsP₆ (Sigma-Aldrich) was carried out at room temperature in the presence of 10 mM wild-type PhvAsr or PhvAmm (50 mM sodium acetate (pH 5.0), 200 mM NaCl, 1 mM BME, and 0.1 mM EDTA). In the case of PhvAbb, 100 mM protein and 10 mM InsP₆ were used. Aliquots of 200 μl were taken, heat-denatured at 95°C for 2 min, and subjected to HPIC (Waters 1525 binary HPIC pump, Milford, MA) utilizing a CarboPac PA-100 (4 × 240 mm) analytical column (Dionex, Sunnyvale, CA) (37) at room temperature with a post-column reactor flow rate of 0.2 ml/min. Identification of hydrolysis products utilized a standard hydrolysis chromatogram (supplemental methods).

Author contributions—L. M. B. crystallized and solved the 4WTY, 4WU2, and 4WT3 structures, determined the pathway of PhvAmm, and wrote the paper. R. J. G. crystallized and solved the 3O3L structure. C. P. C. determined the pathway of PhvAbb. S. C. M. supervised the work, and all authors analyzed the results and approved the final version of the manuscript.

Acknowledgment—We thank L. B. Selinger (Dept. of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada) for providing the PhvAsrC252S and PhvAmmC250S/C548S clones.

References

1. Irvine, R. F., and Schell, M. J. (2001) Back in the water: The return of the inositol phosphates. Nat. Rev. Mol. Cell Biol. 2, 327–338
2. York, J. D., Odom, A. R., Murphy, R., Ives, E. B., and Wente, S. R. (1999) A phospholipase C-dependent inositol phospholipase kinase pathway required for efficient messenger RNA export. Science 285, 96–100
3. Hanakahi, L. A., Bartlet-Jones, M., Chappell, C., Pappin, D., and West, S. C. (2000) Binding of inositol phosphate to DNA-PK and stimulation of double-strand break repair. Cell 102, 721–729
4. Chatterjee, S., Sankaranarayanan, R., and Sonti, R. V. (2003) PhvA, a secreted protein of Xanthomonas oryzae pv. oryzae, is required for optimum virulence and growth on phytic acid as a sole phosphate source. Mol. Plant Microbe Interact. 16, 973–982
5. Macbeth, M. R., Schubert, H. L., Vandemark, A. P., Lingam, A. T., Hill, C. P., and Bass, B. L. (2005) Inositol hexakisphosphate is bound in the ADA2 core and required for RNA editing. Science 309, 1534–1539
6. Tan, X., Calderon-Villalobos, I., Sharon, M., Zheng, C., Robinson, C. V., Estelle, M., and Zheng, N. (2007) Mechanism of auxin perception by the TIR1 ubiquitin ligase. Nature 446, 640–645
7. Lupardus, P. J., Shen, A., Bogoy, M., and Garcia, K. C. (2008) Small molecule-induced allosteric activation of the Vibrio cholerae RTX cysteine protease domain. Science 322, 265–268
8. Majerus, P. W., Zou, J., Marjanovic, J., Kisseleva, M. V., and Wilson, M. P. (2008) The role of inositol signaling in the control of apoptosis. Adv. Enzyme Regul. 48, 10–17
9. Gillaspy, G. E. (2011) The cellular language of myo-inositol signaling. New Phytol. 182, 823–839
10. Wilson, M. S., Livermore, T. M., and Saiardi, A. (2013) Inositol pyrophosphates: Between signalling and metabolism. Biochem. J. 452, 369–379
11. Gawler, D. J., Potter, B. V., and Nahorski, S. R. (1990) Inositol 1, 3, 4, 5-tetraakisphosphate-induced release of intracellular Ca²⁺ in SH-SY5Y neuroblastoma cells. Biochem. J. 272, 519–524
12. Chen, L., and Roberts, M. F. (1999) Characterization of a tetrameric inositol monophosphatase from the hyperthermophilic bacterium Thermotoga maritima. Appl. Environ. Microbiol. 65, 4559–4567
13. Mullaney, E. N., Daly, C. B., and Ullah, A. H. (2000) Advances in phytaglyce in phytase research. Adv. Appl. Microbiol. 47, 157–199
14. Norris, F. A., Wilson, M. P., Wallis, T. S., Galyov, E. E., and Majerus, P. W. (1998) SopB, a protein required for virulence of Salmonella dublin, is an inositol phosphatase phosphatase. Proc. Natl. Acad. Sci. 95, 14057–14059
15. Michell, R. H. (2008) Inositol phosphatases: Evolution and functions. Nat. Rev. Mol. Cell Biol. 9, 151–161
16. Lee, D. C., Cottrill, M. A., Forsberg, C. W., and Jia, Z. (2003) Functional insights revealed by the crystal structures of Escherichia coli glucose-1-phosphatase. J. Biol. Chem. 278, 31412–31418
17. Gruninger, R. J., Dobing, S., Smith, A. D., Bruder, L. M., Selinger, L. B., Wieden, H. J., and Mosimann, S. C. (2012) Substrate binding in protein-tyrosine phosphatase-like inositol phosphatases. J. Biol. Chem. 287, 9722–9730
18. Ariza, A., Moroz, O. V., Blagova, E. V., Turkenburg, J. P., Waterman, J., Roberts, S. M., Vind, J., Sjøholm, C., Lassen, S. F., De Maria, L., Glitsoe, V., Skov, L. K., and Wilson, K. S. (2013) Degradation of phytate by the 6-phytase from Hafnia alvei: A combined structural and solution study. PLoS One 8, e65062
19. Puhl, A. A., Gruninger, R. J., Greiner, R., Janzen, T. W., Mosimann, S. C., and Selinger, L. B. (2007) Kinetic and structural analysis of a bacterial protein tyrosine phosphatase-like myo-inositol phosphatase. Protein Sci. 16, 1368–1378
20. Puhl, A. A., Greiner, R., and Selinger, L. B. (2008) A protein tyrosine phosphatase-like inositol phosphatase from Selenomonas ruminantium subsp. lactilytica has specificity for the 5-phosphate of myo-inositol hexakisphosphate. Int. J. Biochem. Cell Biol. 40, 2053–2064
21. Puhl, A. A., Greiner, R., and Selinger, L. B. (2008) Kinetics, substrate specificity, and stereospecificity of two new protein tyrosine phosphatase-like inositol phosphatases from Selenomonas lactilytica. Biochem. Cell Biol. 86, 322–330
22. Puhl, A. A., Greiner, R., and Selinger, L. B. (2009) Stereo specificity of myo-inositol hexakisphosphate hydrolysis by a protein tyrosine phosphatase-like inositol phosphatase from Megasphaera elsdenii. Appl. Microbiol. Biotechnol. 82, 95–103
23. Gruninger, R. J., Selinger, L. B., and Mosimann, S. C. (2009) Structural analysis of a multifunctional, tandemly repeated inositol polyphosphatase. J. Mol. Biol. 392, 75–86
24. Weber, S., Stirmann, C. U., Wieser, M., Frey, D., Meier, R., Engelhardt, S., Li, X., Capitani, G., Kammerer, R. A., and Hilbi, H. (2014) A type IV translocated Legionella cysteine phytase counters intracellular growth restriction by phytate. J. Biol. Chem. 289, 34175–34188

17310 J. Biol. Chem. (2017) 292(42) 17302–17311
25. Chu, H. M., Guo, R. T., Lin, T. W., Chou, C. C., Shr, H. L., Lai, H. L., Tang, T. Y., Cheng, K. J., Selinger, B. L., and Wang, A. H. (2004) Structures of *Selenomonas ruminantium* phytase in complex with persulfated phytate: DSP phytase fold and mechanism for sequential substrate hydrolysis. *Structure* **12**, 2015–2024

26. Gruninger, R. J., Thibault, J., Capeness, M. J., Till, R., Mosimann, S. C., Sockett, R. E., Selinger, B. L., and Lovering, A. L. (2014) Structural and biochemical analysis of a unique phosphatase from *Bdellovibrio bacteriovorus* reveals its structural and functional relationship with the protein tyrosine phosphatase class of phytase. *PLoS One* **9**, e94403

27. Espinosa, A., Guo, M., Tam, V. C., Fu, Z. Q., and Alfano, J. R. (2003) The *Pseudomonas syringae* type III-secreted protein HopPtoD2 possesses protein tyrosine phosphatase activity and suppresses programmed cell death in plants. *Mol. Microbiol.* **49**, 377–387

28. Bretz, J. R., Mock, N. M., Charity, J. C., Zeyad, S., Baker, C. J., and Hutcheson, S. W. (2003) A translocated protein tyrosine phosphatase of *Pseudomonas syringae* pv. tomato DC3000 modulates plant defense response to infection. *Mol. Microbiol.* **49**, 389–400

29. Gruninger, R. J., Selinger, L. B., and Mosimann, S. C. (2008) Effect of ionic strength and oxidation on the P-loop conformation of the protein tyrosine phosphatase-like phytase, PhyA. *FEBS J.* **275**, 3783–3792

30. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell H. R., and Leslie, A. G. W. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281

31. Winn, M. D, Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242

32. Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82

33. Evans, P. R. (2011) An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292

34. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of COOT. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501

35. Vaguine, A. A., Richelle, J., and Wodak, S. (1999) Sfcheck: A unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 191–205

36. McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. (2011) Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 386–394

37. Blaabjerg, K., Hansen-Møller, J., and Poulsen, H. D. (2010) High-performance ion chromatography method for separation and quantification of inositol phosphates in diets and digesta. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **878**, 347–354

Bacterial PTPLPs in complex with second messengers