manuscript rr5133 for review

The structure of a Pseudomonas sp calcium-dependent phosphoinositol-specific phospholipase C, the first from a Gram-negative bacterium

Olga Moroz, Elena Blagova, Andrey Lebedev, Allan Nørgaard, Dorotea Segura, Thomas Blicher and Keith Wilson*

Category: *research papers*

Co-editor:

Professor R. Read

Department of Haematology, University of Cambridge, Cambridge Institute for Medical Research, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 0XY, UK

Telephone: 01223 336500

Fax: 01223 336827

Email: rjr27@cam.ac.uk

Contact author:

Keith Wilson

YSBL, Chemistry, University of York, Heslington, York, YO10 5DD, United Kingdom

Telephone: 01904 328262

Fax: 01904328266

Email: keith.wilson@york.ac.uk

The structure of a calcium-dependent phosphoinositide-specific phospholipase C from *Pseudomonas sp-62186*, the first from a gram-negative bacterium

Olga V. Moroz[1], Elena Blagova[1], Andrey A. Lebedev[2], Allan Nørgaard[3], Dorotea R. Segura[3], Thomas H. Blicher[3], Jesper Brask[3], and Keith S. Wilson[1]

[1] Department of Chemistry, York Structural Biology Laboratory, University of York, York YO10 5DD, U.K.

[2] CCP4, STFC Rutherford Appleton Laboratory, Harwell Oxford, Didcot OX11 0QX, UK

[3] Novozymes A/S, Krogshoejvej 36, Bagsvaerd, 2880, Denmark.

Corresponding author emails: olga.moroz@york.ac.uk and keith.wilson@york.ac.uk

Keywords: Phospholipase C, calcium-dependent, phosphoinositide-specific, X-ray structure

Running title: The structure of a *bacterial* calcium-dependent phosphoinositide-specific phospholipase C

Abbreviations: PI-PLC: phosphoinositide-specific phospholipase C, *Bc*PI-PLC: *Bacillus cereus* PI-PLC, *Bt*PI-PLC: *B. thuringiensis* PI-PLC, *Lm*PI-PLC: *Listeria monocytogenes* PI-PLC, *Sau*PI-PLC: *Staphylococcus aureus* PI-PLC, *San*PI-PLC: *Streptomyces antibioticus* PI-PLC, *Ps*PI-PLC: *Pseudomonas sp-62186* PI-PLC

ABSTRACT

Bacterial phosphoinositide-specific phospholipases C (PI-PLCs) are the smallest members of the PI-PLC family that includes much bigger mammalian enzymes responsible for signal transduction as well as enzymes from protozoan parasites, yeast and plants. Eukaryotic PI-PLCs have calcium in the active site, but this is absent in the known structures of gram-positive bacteria where its role is instead played by arginine. In addition to their use in a number of industrial applications, the bacterial enzymes attract special interest because they can serve as convenient models of the eukaryotic enzymes' catalytic domains for *in vitro* activity studies. We report here the structure of a PI-PLC from *Pseudomonas sp-62186,* the first from a gram-negative bacterium and the first of a native bacterial PI-PLC with calcium present in the active site. Solution of the structure posed particular problems owing to the low sequence identity of available homologous structures. Dependence on calcium for catalysis makes this enzyme a better model for studies on the mammalian PI-PLCs than previously used calcium-independent bacterial PI-PLCs.

INTRODUCTION

Phospholipases of type C (PLC) cleave phospholipids on the diacylglycerol side of the phosphodiester bond. Members of the phosphoinositide-specific phospholipase C (PI-PLC) family catalyse the cleavage of the membrane lipid phosphatidylinositol (PI), or its phosphorylated derivatives, releasing diacylglycerol and (phosphorylated) myo-inositol. PI-PLCs have been isolated from plants, mammals, yeast, protozoan parasites and bacteria, reviewed in (Griffith & Ryan, 1999, Katan, 1998). In plants, they are involved in lipid signalling, regulation of growth and development, and response to nutrient deficiency as well as biotic and abiotic stress (Singh *et al.*, 2015). In eukaryotes, they cleave more highly phosphorylated forms of PI forming myo-inositol 1,4,5-trisphosphate (Ins 1,4,5-P3) and diacylglycerol, which act as second messengers in a number of receptor-mediated signalling pathways (Suh *et al.*, 1988, Rhee *et al.*, 1989, Majerus *et al.*, 1986, Cocco *et al.*, 2015). In gram-positive bacteria, they have been proposed to contribute to virulence (Heinz *et al.*, 1995), although direct evidence of their role in pathogenicity is unclear (White *et al.*, 2014). PI-PLCs are also found in non-pathogenic gram-positive bacteria (Griffith & Ryan, 1999).

All PI-PLCs belong to the PI-PLCc_GDPD_SF (catalytic domain of phosphoinositide-specific phospholipase C-like phosphodiesterases superfamily, accession cl14615) superfamily in the NCBI Conserved domains database, CDD (Marchler-Bauer *et al.*, 2015), together with several other enzyme families e.g. glycerophosphodiester phosphodiesterases (GP-GDE), which all share a similar mechanism of general base and acid catalysis with conserved histidine residues and hydrolyze the 3'-5' phosphodiester bonds in different substrates. The active site features relevant to PI-PLCs and the reaction mechanism will be discussed in more detail below. Briefly, the catalytic domain fold is a $(\beta\alpha)_8$, TIM barrel, with conserved active site residues belonging to one half of the barrel (termed the X-box for the mammalian PI-PLCs (Rhee *et al.*, 1989)), and the second half responsible for the specificity towards different ligands.

There are 13 isoforms of mammalian PI-PLCs identified to date, belonging to 6 families: PLCβ, γ, δ, ε, ζ and η, with specific isoforms being linked to specific signalling pathways. The best studied isoform for which structural information is

available, is PI-PLCδ from rat (Essen *et al.*, 1996), which will be used in structure comparisons here. All isoforms have four domains in common (with the additional, different domains present for some, depending on their specific functions): a pleckstrin homology domain (PH) involved in membrane binding, an EF-hand domain (with EF-hand calcium binding motifs, but most probably not linked to calcium regulation in PLCs), a catalytic domain ($(\beta\alpha)_8$, TIM barrel), and a β-sandwich C2 domain (named after the second conserved domain of protein kinase C, with a possible role in orientation of the catalytic domain relative to the membrane) (Bunney & Katan, 2011, Fukami *et al.*, 2010, Williams, 1999). It has been speculated that the bacterial enzymes evolved from eukaryotic PI-PLCs, losing the additional domains and optimizing the catalytic domain for PI cleavage (Heinz *et al.*, 1998). Indeed, bacterial PI-PLCs exhibit significant structural similarity to the catalytic domain of mammalian PI-PLCs (Heinz *et al.*, 1998, Essen *et al.*, 1996), and are viewed as convenient model systems for these more complicated eukaryotic signalling molecules (Griffith & Ryan, 1999). Until recently, the most significant and consistent difference was the presence of an essential, catalytic calcium in the eukaryotic enzymes and its replacement by arginine in all known bacterial structures. As bacterial cell membranes do not normally contain PI (Bishop *et al.*, 1967, Ames, 1968), it was proposed that bacteria may use PI-PLCs for interaction with (higher) eukaryotic organisms (Smith *et al.*, 1995, Goldstein *et al.*, 2012, Marques *et al.*, 1989, Zhao *et al.*, 2013, White *et al.*, 2014). While eukaryotic enzymes need calcium for fine-tuning signal transduction, the bacterial enzymes are thought to have evolved to use arginine instead of calcium to allow them to act in the intracellular space after uptake by the host cell where calcium levels are normally low.

Crystallisation of a calcium-dependent PI-PLC from *Streptomyces antibioticus* (*San*PI-PLC) (Jackson & Selby, 2012) was reported earlier, but while the structure was deposited in the PDB in 2010 (codes 3H4X and 3H4W), there is no associated publication. The purification of *San*PI-PLC was reported and its calcium dependence identified (Iwasaki *et al.*, 1994). The calcium dependence was subsequently confirmed and sequence alignments showed that the enzyme had calcium-binding residues corresponding to those of the eukaryotic enzymes as well as two catalytic histidines (Iwasaki *et al.*, 1998, Zhao *et al.*, 2003). Inspection of the structure reveals that while there is no calcium ion in the deposited PDB file, this enzyme potentially

has calcium-binding residues instead of an arginine in the active site. However, recent research has shown that the reaction mechanism of *San*PI-PLC proceeds in an unusual way with the enzyme capable of cleaving a number of PI analogues, so the exact natural substrates of *San*PI-PLC are still to be confirmed (Bai *et al.*, 2010).

All bacterial PI-PLC structures available from the PDB to date have been from gram-positive bacteria, and none of them are calcium-dependant. The only other bacterial PI-PLC for which a structure with an active site calcium has been reported is that from *Bacillus thuringiensis* (PDB code 1T6M), which has an engineered calcium binding site resulting from an R69D mutation (Apiyo *et al.*, 2005, Kravchuk *et al.*, 2003). We report here the first structure of a wild type PI-PLC from the gram-negative bacterium *Pseudomonas sp-62186*, with calcium clearly identified in the active site in both the apo form and in complex with myoinositol. This makes an interesting update to an established PI-PLC classification.

EXPERIMENTAL

Cloning, expression, purification and activity assay

The *Ps*PI-PLC gene was inserted into a Bacillus expression plasmid. The DNA encoding the mature polypeptide predicted by Signal P (Bendtsen *et al.*, 2004) was cloned in frame to the *Bacillus clausii* secretion signal, BcSP, having the following amino acid sequence MKKPLGKIVASTALLISVAFSSSIASA replacing the native secretion signal sequence with an extra alanine at the C-terminus. This resulted in a recombinant mature polypeptide with an alanine in front of the N-terminus of the mature wild type sequence. The BcSP fused to the PI-PLC gene was expressed under control of a promoter system described in (Widner *et al.*, 2000). Furthermore, the plasmid contained a terminator sequence and a gene coding for chloramphenicol acetyltransferase which was used as a selection maker as described in (Diderichsen *et al.*, 1993) for *Bacillus subtilis*. The β-lactamase gene giving ampicillin resistance and the kanamycin resistance gene were used as cloning selection marker genes for *Escherichia coli* growth. The plasmid also contained an *E. coli* origin of replication.

E. coli TOP10 cells were transformed with the plasmid and one clone containing the correct PI-PLC gene sequence was selected. Competent *B. subtilis* cells were transformed with the plasmid isolated from the selected *E. coli* clone containing the PI-PLC gene. The final expression plasmid integrated into the *B. subtilis* chromosome by homologous recombination into the pectate lyase gene locus. Chloramphenicol-resistant transformants were analysed by PCR to verify the correct size of the amplified fragment. A recombinant *B. subtilis* clone containing the integrated expression construct was selected and cultivated on a rotary shaking table in 500 ml baffled Erlenmeyer flasks each containing 100 ml LB media supplemented with 34 mg/l chloramphenicol. The clone was cultivated for 5 days at 30°C. The enzyme-containing supernatants were separated from the cells by centrifuging the culture broth for 30 minutes at 15.000 x g and the enzyme was purified as described below.

The supernatant from 1 l of culture broth (shake flasks) was filtered through a 0.22 μm PES membrane and buffer-exchanged into 50 mM MES pH 6.5 using a packed bed of G-25 (Sephadex™ G-25 Medium, GE Healthcare). Collected fractions were

analysed by SDS-PAGE (NuPAGE® 4-12% Bis-Tris gel, reducing conditions) and pooled based on the presence of a band with the expected molecular weight (approx. 33 kDa). The pool was loaded on a packed bed of Source™ 15S (GE Healthcare) and bound material eluted using a linear gradient. *Ps*PI-PLC eluted at 50 mM MES, 0.15 M NaCl pH 6.5 (in a gradient going from 0-0.5M NaCl (buffer A: 50 mM MES pH 6.5; buffer B: Buffer A + 0.5 M NaCl) in 10 CV's. Collected fractions were analysed by SDS-PAGE and pooled based on the presence of a band with the expected molecular weight. The pool was concentrated 2-fold by use of centrifugal spinfilters (Vivaspin® 20, 10,000 MWCO PES, cat. No. VS2002, Sartorius AG) to a final conc. of approx. 3.5 mg/mL (based on A_{280}, ε_{280} =1.7 ml*mg^{-1}*cm^{-1}).

For the activity assay, the enzyme was diluted to 0.9 and 0.09 mg/mL in 100 mM citrate buffer, pH 4.0, 5.5 and 7.0. Crude soybean oil (0.25 ml) was pipetted into 2 ml Eppendorf tubes to which the diluted enzyme (25 µl) was added. This resulted in reactions with 10 and 100 mg enzyme per kg oil, 10% water, at pH 4.0, 5.5 and 7.0. The mixtures were incubated in a thermoshaker at 50°C for 2 h. Then adding 0.500 mL internal standard solution (IS, 2 mg/mL triphenyl phosphate in methanol), 0.5 ml chloroform-d ($CDCl_3$) and 0.5 ml Cs-EDTA buffer (0.2 M EDTA adjusted to pH 7.5 with CsOH). Phase separation was obtained after 30 s shaking followed by centrifugation (3 min, 13,400 rpm). For each reaction as well as a blank (no enzyme), the lower phase was transferred to a NMR tube with a pipette and the ^{31}P NMR spectrum was recorded using a 400 MHz Bruker Avance-III HD instrument, operating at 300 K, with 128 scans, 5 s relaxation delay. All signals were integrated. Chemical shifts were approx. 1.7 ppm (PA), -0.1 ppm (PE), -0.5 ppm (PI), -0.8 ppm (PC), -17.8 ppm (IS). The concentration of each phospholipid class was calculated as "ppm P", i.e. mg elemental Phosphorus per kg oil sample. Hence, ppm P = I/I(IS) * n(IS) * M(P) / m(oil).

Crystallization

Native protein

Prior to crystallization, the protein was further concentrated to 24 mg/ml by ultrafiltration in an Amicon centrifugation filter unit (Millipore), aliquoted to 50 µl and aliquots that were not immediately set up for crystallization were flash frozen in

liquid nitrogen and stored at -80°C to use later in optimizations/complex formation. The buffer was not changed from the ion exchange elution conditions 50 mM MES, 0.15 M NaCl pH 6.5. Crystallization screening was carried out using sitting-drop vapour-diffusion with drops set up using a *Mosquito Crystal* liquid handling robot (TTP LabTech, UK) with 150 nl protein solution plus 150 nl reservoir solution in 96-well format plates (MRC 2-well crystallization microplate, Swissci, Switzerland) equilibrated against 54 µl reservoir solution. Experiments were carried out at room temperature with a number of commercial screens. Initial crystals were obtained in PACT premier™ HT-96 (Molecular Dimensions) condition H4 (0.2 M KSCN, 20% PEG 3350, 0.1 M BTP pH 8.5). These were optimised in 24-well Linbro trays (hanging drops) using an Oryx-8 robot (Douglas instruments, (Shah *et al.*, 2005)), by microseeding with a range of seed dilutions and a fine pH grid (between pH 9-10). Seeding stock was prepared, and seeding carried out according to the published protocols (Shaw Stewart *et al.*, 2011). Briefly, crystals from the initial successful drop were transferred on a glass slide, crushed, and collected in a Seed Bead^{TM} (HR2-320, Hampton research) with 50 µl well solution added, vortexed for one minute, and used as an initial seeding stock. New seeding stock was prepared from the first optimisations. Different protein/well/seeds ratios (1/0.9/0.1, 1/0.99/0.01, 1/1/0; these are drop sizes in µl) were used for three hanging drops on one coverslip, in an automated set up using the Oryx8 robot. Unused seeding stocks were stored at -20°C for later experiments. The best crystals grew close to pH 10, in the PEG concentration range 24-27%, with 1:200 ($1_{protein}$:0.99_{well} :0.01_{seed}) seed dilution.

Myoinositol complex

Myoinositol (Merck Millipore, catalogue number 104507, 99+%) was dissolved in water to make a 500 mM stock. This was mixed with the protein, 40 µl protein + 10 µl myoinositol, to give a final concentration of 100 mM. Crystals were obtained by seeding with native protein crystallization stock, in hanging drops with varying protein-to-well solution ratios, using the Oryx-8 robot, under conditions similar to those giving native crystals. No back-soaking was carried out, but the excess myoinositol was possibly removed during cryoprotection.

Data collection, structure solution and refinement

Computations were carried out using programs from the CCP4 suite (Winn *et al.*, 2011), unless otherwise stated. The final models were validated with *MOLPROBITY* (Chen *et al.*, 2010). Data processing and refinement statistics for all three structures are provided in Table 1. In the text below and in Figures 2 and S1, the structure with the smaller unit cell ($a = b \approx 96$ Å) is referred to as Form 1, and that with larger unit cell ($a = b \approx 135$ Å), both apo and with bound ligand, is referred to as Form 2.

Native structure

Data were collected for two crystal forms at the Diamond Light Source and processed with Xia2 (Winter *et al.*, 2013), Table 1. Data for Form 1 were processed to a resolution limit of 1.46 Å in point group 422 and *P*-lattice with unit cell dimensions of a = b = 96.1 Å and c = 113.3 Å, which suggested either one molecule (solvent content of 69%) or two (38.5%) in the asymmetric unit. The second possibility was confirmed since a strong non-origin peak in the Patterson synthesis at (0.50, 0.50, 0.48) indicated two-fold pseudotranslation. The low solvent content implied potential problems for molecular replacement in addition to those associated with the presence of pseudotranslation. Data to 1.7 Å resolution were used in molecular replacement and to 1.5 Å in the final refinement.

The first search model was generated with Sculptor (Bunkoczi & Read, 2011) from PDB entry 3h4w (sequence identity 27%). Sculptor option 1 was selected to provide maximum truncation of the input model, but nevertheless Phaser (Mccoy *et al.*, 2007) rejected all the putative solutions because of clashes. Moreover, there were no solutions which stood out before the rejection, with the Log Likelihood Gain (LLG, difference between the log-likelihoods for a given model and unrelated models) varying from 27 to 29 for several top solutions, well below the value of 120 that is considered as an almost certain indication of a correct solution.

There are several pipelines currently available which try various protein fragments or highly truncated homologous structures as search models and then try to extend such partial solutions through density modification and model building. Before undertaking such time consuming calculations, an attempt was made (and turned out to be

successful) to apply this approach at a much simpler level, with a single truncated model, which was manually prepared using *COOT* (Emsley *et al.*, 2010) from the 3h4w search model. The truncated model contained helices 90-105 and 133-144, strands 45-51, 114-122 and 183-190, and several adjacent residues (helices 2 and 3, and strands II, III and IV of the X-box; the total of 72 residues out of 298 in the target sequence) which seemed to form the most compact fragment of the molecule. *PHASER* now positioned two tNCS-related copies of this model with a LLG of 52 compared to 44 and 42 for the second and third solutions. The best solution confirmed the space group to be $P4_322$. A straightforward attempt to rebuild the MR solution using Buccaneer failed: the resulting R/R-free = 0.49/0.53 for 493 out of 596 residues built indicated incorrect main-chain tracing or sequence assignment. In order to generate a better starting model, two copies of a complete molecule from 3h4w were superposed using COOT on the two truncated copies positioned by PHASER. The complete molecules packed almost perfectly, with only one loop being too close to a neighbouring molecule thus providing strong evidence that the MR solution was correct. The model consisting of complete homologous molecules was rebuilt with a procedure involving refinement with *REFMAC5* (Murshudov *et al.*, 2011) and removal of incorrect fragments (total 300 cycles of refinement, final R/R-free = 0.50/0.53 against all data), density modification with *SHELXE* (8 cycles, final CC of 48%), refinement with *REFMAC5* (24 cycles, R/R-free = 0.42/0.44), model extension and sequence docking with *BUCCANEER* (R/R-free = 0.25/0.28, 592 residues) and final refinement and model correction (R/R-free = 0.13/0.17).

For crystal Form 2, diffraction images were merged using *Xia2* to a resolution of 1.17 Å. Molecular replacement with the model from Form 1 using *MOLREP* identified four monomers per asymmetric unit, all related by pseudotranslation. Manual rebuilding with COOT and refinement with *REFMAC5* resulted in a model with R/R-free of 0.12/0.14 and 92.5 % residues built.

Myoinositol complex

Data were collected at DLS, and processed using XDS (Kabsch, 2010) in space group $P4_32_12$, to a resolution of 1.45 Å. The structure was essentially isomorphous to the native Form 2 and was solved by MR with *MOLREP* using that structure as search

model. Refinement was performed using *REFMAC5* and manual model building/correction in *COOT*.

Sequence alignments and structure superpositions

Sequence alignments and structure superpositions were carried out using ClustalX (Larkin *et al.*, 2007) and SSM (Krissinel & Henrick, 2004) as implemented in CCP4mg (McNicholas *et al.*, 2011), respectively. The results are summarised in Table 2.

RESULTS AND DISCUSSION

Cloning, expression, purification and assay

The gene encoding the *Ps*PI-PLC was amplified by PCR from genomic DNA of the bacterium *Pseudomonas sp-62186*, isolated from a seaweed sample collected in Denmark. Sequencing of 16S rRNA from the *Pseudomonas* species revealed 99.7% identity to *Pseudomonas protegens* CHAO (EMBL:CP003190). The *Ps*PI-PLC gene was cloned, expressed in *Bacillus subtilis*, and purified as described under Experimental. The sequence of the gene coding for the PI-PLC from *Pseudomonas sp-62186* was deposited in GenBank with the accession number GenBank:KY078744.

To assay activity and substrate specificity, the purified enzyme in citrate buffer was incubated with crude (non-refined) soybean oil having a natural content of phospholipids. Assay conditions hence resemble those used in enzymatic plant oil degumming (De Maria et al., 2007). Following an aqueous extraction, the content of non-hydrolyzed phospholipids was characterized and quantified by ^{31}P NMR spectroscopy. This sample preparation was inspired by reports documenting that extraction with aqueous EDTA leads to sharper phospholipid NMR signals (Yao & Jung, 2010). The result (Fig. 1) provided information about the enzyme's pH activity profile and substrate specificity. As expected, the enzyme is highly specific towards PI and prefers neutral over acidic conditions.

Pseudo-symmetry in the two crystal forms

Figure 2 (*a*) illustrates the relation between the symmetries of the two crystal forms, and demonstrates that the diagonal screw-twofold axes in Form 1 ($a = b \approx 96$ Å) become a coordinate screw two-fold axes in Form 2 ($a = b \approx 135$ Å). Accordingly, the HM-symbols for the two space groups are different, $P4_322$ and $P4_32_12$, respectively. The situation is further complicated by the presence of a pseudo-translation in Form 1, which acts as a second pseudo-translation in Form 2, and results in additional rotational pseudo-symmetry elements in both forms. The latter elements coincide in the two forms and are not shown in Figure 2 (*a*). As was shown using the *Zanuda* program (Lebedev & Isupov, 2014), the pseudo-symmetry space group (the group that includes all crystallographic symmetry elements and all pseudo-symmetry elements)

for both structures is $I4_122$ with unit cell dimensions as in Form 1. As outlined in the next paragraph, the space groups realized in each of the two forms presents one of four possible ways of breaking $I4_122$ symmetry to obtain a tetragonal space group with a *P*-cell (Form 1) and a tetragonal space group with twice larger *P*-cell (Form 2). In particular, the actually existing and presented in Fig. 2 (*a*) subgroup-supergroup relation between forms 1 and 2 is not *a priori* imposed by the approximate $I4_122$ pseudo-symmetry or by the point-group symmetry of the diffraction data and the systematic absences.

Figures 2 (*b*, *c*) show the Patterson peaks and thus illustrate the pseudo-translations in the two crystal forms. There are four possible crystallographic space groups ($P4_122$, $P4_12_12$, $P4_322$ and $P4_32_12$) for pseudo-symmetry space group $I4_122$ provided that the crystallographic translation base vectors are **a**, **b** and **c**, and (**a** + **b** + **c**)/ 2 is a pseudo-translation (as in Form 1). There is intensity modulation for reflections $k = l = 0$ for Form 1 (Fig. S1 (*a)*), which, however, are not true systematic absences because the intensities of the most weak reflections with $h = 2n + 1$ nevertheless significantly exceed their experimental uncertainties (Fig. S1 (*c)*). This observation reduces the number of possible space groups to two ($P4_122$, $P4_322$). Similarly, there are four possible space groups (HM symbols $P4_122$, $P4_12_12$, $P4_322$ and $P4_32_12$) for the pseudo-symmetry space group $I4_122$ with crystallographic translation base vectors **a'** = **a** + **b**, **b'** = – **a** + **b** and **c'** = **c** (as in the large-cell Form 2). Two of these (HM symbols $P4_12_12$ and $P4_32_12$) are consistent with the clear systematic absences $k = l = 0$, $h = 2n + 1$ in the Form 2 data (Fig. S1*b*, *d*). During the structure determination, the remaining ambiguity (enantiomorphic space groups) was resolved at the molecular replacement step (Form 1: highest Phaser LLGs of 52.5 for the correct space group $P4_322$ and 36.5 for the incorrect $P4_122$; Form 2: highest Molrep's CCs of 0.269 for the correct and 0.259 for the incorrect) and this space group assignment was confirmed by low *R*-factors (below 0.2) and clear electron density maps for the refined structures.

Overall fold and comparison to other PI-PLCs

Like the structures of other bacterial PI-PLCs reported to date, *Ps*PI-PLC forms an imperfect $(\beta\alpha)_8$ TIM-barrel (where TIM stands for triose phosphate isomerase

(Banner *et al.*, 1975)); bacterial PI-PLCs are family 51699 in the SCOP database (Andreeva *et al.*, 2008). As for other bacterial PI-PLCs, some of the β-strands are longer, and some shorter than in a "canonical" $(\beta\alpha)_8$ barrel (Moser *et al.*, 1997, Heinz *et al.*, 1995), which gives the barrel an irregular shape. In addition, the α-helices are distributed around the barrel in an irregular way, with most clustered on one side. However, in contrast to other bacterial PI-PLC structures, there are connecting α-helices between strands IV and V, and V and VI, and the β-barrel itself is more closed, while the C-terminal α8 helix is absent in *Ps*PI-PLC (Fig. 3, 5). Furthermore, *Ps*PI-PLC has two additional short antiparallel β strands IIb, between II and III, and VIIIb, close to the C-terminus.

X-ray structures of five bacterial PI-PLCs are available from the PDB from *Bacillus cereus* (*Bc*PI-PLC), *Bacillus thuringiensis* (*Bt*PI-PLC), *Listeria monocytogenes* (*Lm*PI-PLC), *Staphylococcus aureus* (*Sau*PI-PLC) and *Streptomyces antibioticus* (*San*PI-PLC); PDB codes for representative structures are given in Table 2. The sequence identity of *Ps*PI-PLC to the closest homologue, *San*PI-PLC, is only 27.1%, and this made molecular replacement a challenge. Comparison of the overall fold for *Ps*PI-PLC and *San*PI-PLC is shown in Fig. 4 (a) together with the outline of the final model. The sequence identity is even lower, between 15.7-18.4% (Table 2), for the other bacterial homologues. Nevertheless, the active sites superimpose relatively well, with the ligand in the structures of complexes occupying essentially the same location (Fig. 4. (a), (b), (c)). In addition, the bacterial sequences have a degree of similarity to the X-box regions of the catalytic domains of eukaryotic PI-PLCs; the name X and Y boxes was given to the N- and C-terminal parts of the split $(\alpha\beta)_8$ barrel (Rhee *et al.*, 1989). While the sequence identity is low (between 13.9 and 19.7% for the X-box and bacterial PI-PLCs from the above list), the catalytic histidines are conserved. Moreover, the calcium-binding residues Asn312, Glu341 and Gln343 from the X-box are present in both calcium-binding PI-PLCs (Fig. 5). This catalytic calcium-binding (or calcium-free, with arginine) motif has been described for the PI-PLC family in an overview of β-barrel proteins (Nagano *et al.*, 2002). The structural similarity, as for the bacterial PLCs, is significant - the X-box of the rat enzyme (296-440; 1djx) superimposes well on the bacterial PI-PLC N-terminal half of the β-barrel. Similar to what has been described for non-calcium binding bacterial PI-PLCs (Heinz *et al.*, 1998), sequence conservation for the second half of the barrel, corresponding to the

Y-box of the mammalian enzyme, is very low, although a number of structural elements superimpose reasonably well (Figs 4, 5).

Ligand binding site

As for all TIM barrel proteins, the ligand binding site is formed by residues from the C-terminal regions of the β-strands plus the connecting loop regions, the so-called "activity face" at the C-terminal end of the $(\alpha\beta)_8$ barrel (Hocker *et al.*, 2001, Nagano *et al.*, 2002). As reported for other PI-PLCs (Essen *et al.*, 1997), there are no major conformational changes upon ligand binding, with the average rmsd between the subunits of the complex and those of the native structures lying within the range of rmsd between the individual subunits of the native structures (5fyo: 0.142 Å, 5fyp: 0.276 Å, 5fyr: 0.245 Å, 5fyo/5fyp: 0.233 Å; 5fyr/5fyo: 0.202 Å, 5fyr/5fyp: 0.220 Å). Myoinositol is coordinated by Asn27, Asn240, Asn189, Glu48, Arg260, three water molecules and most interestingly a calcium ion. Trp262 contributes to the ligand-binding pocket. Asn27, Glu48, Asp50, Asp119, the O2 of myoinositol and a water molecule (Fig. 6 (a)), coordinates the calcium ion. His26 is 3.2Å from the O2 of inositol, equivalent to the catalytic His32 first described in *Bc*PI-PLC (Heinz *et al.*, 1995). In *Ps*PI-PLC Tyr283 has replaced Asp274, which was proposed as a second member of the catalytic triad in *Bc*PI-PLC. While Asp274 in *Bc*PI-PLC forms a hydrogen bond with the ND1 of His32 (equivalent to His26 in *Ps*PI-PLC), stabilising its position for catalysis, Tyr283 possibly plays an equivalent role in making a stacking interaction. The inositol OH2 moiety was proposed to complete a catalytic triad. Another histidine, His70, occupies a similar, although not identical position to His82 in *Bc*PI-PLC, with CE1 ~5.5Å from OH1 of inositol, and has been suggested to be essential for catalysis, protonating the leaving group. Histidines corresponding to His26 and His70 are conserved in all (including mammalian) PI-PLCs, and as mentioned in the Introduction, are present in all the PI-PLCc_GDPD_SF superfamily members. The reaction mechanism is described below.

Active site comparisons

***Streptomyces antibioticus San*PI-PLC**

As expected, the ligand-binding site superimposes closely on that of *San*PI-PLC used as the molecular replacement model. However, a number of features are less similar than for the *Bacillus* or, more interestingly, rat enzymes. This is in agreement with reports that *San*PI-PLC is an unusual bacterial PI-PLC (Iwasaki *et al.*, 1994) being later identified as one of two independent PI-PLCs in *S. antibioticus* cloned by a shotgun approach (Iwasaki *et al.*, 1998). This enzyme was described as "eukaryotic-like" due to a number of characteristics, one of which was its dependence on calcium, in addition to its lack of activity on GPI-anchored proteins. Two structures of the native enzyme were deposited in the PDB in 2009 (3h4w, 3h4x) but no calcium ion was modelled in the potential calcium-binding site. Subsequently, an unusual mechanism for PI catalysis was proposed, with *trans*-1,6-cyclic *myo*-inositol phosphate (1,6-IcP) being formed as an intermediate step (Bai *et al.*, 2009). In addition, *San*PI-PLC was shown to be capable of cleaving a number of structural analogues of PI, with its role and natural substrates still under investigation (Bai *et al.*, 2010).

The calcium binding residues Asn27, Glu48 and Asp50 are all present in *San*PI-PLC (Fig. 5 – sequence alignment, and Fig. 7 (a)). A numbering discrepancy should be mentioned here for clarity: the corresponding residues in *San*PI-PLC would have been Asn17, Glu39 and Asp41 if numbered as for the mature protein excluding the signal peptide from the entry for the proenzyme in Uniprot, B3A043 (Zhao *et al.*, 2003). In the deposited structure (3h4w, 3h4x) the numbering starts from the beginning of the expression tag, so adds 21 to the original numbering making these residues Asn38, Glu60 and Asp62. We use the PDB numbering here to facilitate structure comparisons for the reader. The fourth potential calcium-binding residue Glu129, corresponding to Asp119 in *Ps*PI-PLC, is also present, although not exactly at the same location, but could still bind a calcium ion (Fig. 7 (a)).

While the calcium-binding residues and catalytic histidines in *San*PI-PLC superimpose reasonably well with corresponding residues in *Ps*PI-PLC, the ligand-binding site as a whole has considerable differences in architecture from *Ps*PI-PLC as well as the *Bacillus* and mammalian enzymes, with a wider opening, possibly explained by a different reaction mechanism. In *Ps*PI-PLC, Trp262 lines the binding pocket, making a stacking interaction with inositol. There are no equivalent residues

in *SanPI*-PLC at the corresponding location (Fig. 7 (a)). His236 could possibly take part in the lining of the binding pocket in *San*PI-PLC. The other bacterial PI-PLCs have tyrosines involved in stacking with inositol, close to, but not exactly superimposing on Trp262 in *Ps*PI-PLC (Fig. 7 (b)), and the rat enzyme also has a tyrosine at the corresponding location (Fig. 7c). In summary, the binding pocket of the *Streptomyces* enzyme has significant differences from those of *Ps*PI-PLC and the *Bacillus* and rat enzymes discussed below.

The Bacillus PI-PLCs

*Bt*PI-PLC is a bacterial enzyme in which a R69D variant was designed to engineer a calcium-binding site (Apiyo *et al.*, 2005, Kravchuk *et al.*, 2003), and was shown to be activated by calcium. The structure was determined only for the mutant, but the closely similar *Bc*PI-PLC (PDB ID 1ptg) can be used to compare the calcium-bound mutant to a wild-type enzyme, with an active site arginine Arg69. *Bc*PI-PLC has 96% sequence identity to *Bt*PI-PLC only differing by eight amino acids, and was used for comparisons of the mutant with the wild-type enzyme (Apiyo *et al.*, 2005). Here again, there is confusion with residue numbering – the authors refer to the *Bc*PI-PLC numbering, even though the amino acids are numbered differently in the PDB entry 1t6m. The *Bc*PI-PLC numbering is used here for both *Bacillus* structures. Fig. 7 (b) shows that the Arg69 guanidinium group from *Bc*PI-PLC superimposes well on the calcium ion from *Bt*PI-PLC, which in turn is quite close to the calcium ion in *Ps*PI-PLC. The engineered calcium-binding residue Asp69 (result of the R69D mutation) is positioned between two calcium-binding residues of *Ps*PI-PLC: Asp50 and Asp119. The other two residues involved in calcium coordination in *Bt*PI-PLC - Asp33 and Asp67 – superimpose well on Asn27 and Glu48 from *Ps*PI-PLC. The myoinositol ligand is in a similar location to that in *Ps*PI-PLC, further away from the calcium ion (or Arg69 guanidinium group) towards the exit from the binding pocket. Tyr200 is involved in stacking interactions with the ligand, similar to Trp262 in *Ps*PI-PLC coming in from a different part of the protein chain. Asp198 of *Bc*PI-PLC occupies a position very close to that of Trp262, making hydrogen bonds to OH3 and OH4 of the ligand. This again differs from *Ps*PI-PLC, where the OH3 and OH4 of myoinositol are coordinated by Asn189, Asn240 and Arg260, which are all far away from Asp198 (*Bc*PI-PLC).

Rat PI PLC

The ligand binding site of the rat PI-PLC catalytic domain (Essen *et al.*, 1997) superimposes on that of the *Ps*PI-PLC strikingly well, more closely than for the *Bacillus* enzymes or even *San*PI-PLC. The calcium binding residues all superimpose almost exactly, with the exception of Asp119 in *Ps*PI-PLC which is Glu390 in rat PI-PLC. Nevertheless, the calcium coordinating OD2 from Asp119 in *Ps*PI-PLC is in a very similar location to the Glu390 OE1 in rat PI-PLC. The calcium ions from the *Ps*PI-PLC and rat PI-PLC structures superimpose very well, much better than the engineered calcium in *Bt*PI-PLC (Fig. 7 (b), (c)). The ligand in rat PI-PLC is D-*myo*-inositol-1,4,5-trisphosphate (Ins1,4,5-P3), which differs from the inositol (Ins) in the *Ps*PI-PLC complex by three phosphates on positions 1, 4 and 5 of the inositol ring. This is a phosphatidyl-less substrate analogue for the mammalian enzymes, which cleave the head groups from more highly phosphorylated species than the bacterial ones. The order of substrate preference phosphatidylinositol-4,5-bisphosphate (PIP_2) > phosphatidylinositol-4-phosphate (PIP) > PI; bacterial enzymes are only able to cleave PI. There are several ligand complexes in the PDB, but only 1djx is used for the comparisons here, since the others are very similar. The inositol ring of Ins from *Ps*PI-PLC superimposes on that of Ins1,4,5-P3 even better than on Ins from *Bc*PI-PLC, (where the ligand is exactly the same). A stacking interaction is provided by Tyr551, which is located closer to Phe262 from *Ps*PI-PLC than the tyrosines in the *Bacillus* enzymes. There is no equivalent stacking interaction in *San*PI-PLC at this location, probably because it binds different ligands and cleaves them by a different mechanism.

Reaction mechanism

For PI-PLCs, both bacterial (*Bc*PI-PLC; (Heinz *et al.*, 1995, Heinz *et al.*, 1998), and mammalian (rat PI-PLC; (Essen *et al.*, 1996, Essen *et al.*, 1997), substrate-assisted catalysis has been proposed, in which the 2-OH group of phosphoinositide acts as a nucleophile. His32 in *Bc*PI-PLC is a general base, and His82 a general acid, donating the proton to the leaving group. For the mammalian (rat) PI-PLC the mechanism may differ in that His311 (analogous to His32) is too distant from 2-OH of inositol in all available structures of complexes with inositol phosphates, so a role in stabilizing a transition state has been suggested for this His, while the role of general base was

assigned to Glu341 (Ellis *et al.*, 1998). In *Ps*PI-PLC, His26 (analogous to His311 in rat PI-PLC and His32 in *Bc*PI-PLC) is close enough to the 2-OH group of the inositol (3.2Å) and its role as general base in the first step of the reaction cannot be excluded. Otherwise the mechanism for *Ps*PI-PLC should be similar to that suggested for the mammalian enzyme (Fig. 6 (b)), and the general base His26 (His311 in rat PI-PLC and His32 in *Bc*PI-PLC) or Glu 48 (Glu341 in rat PI-PLC; no direct equivalent in *Bc*PI-PLC) deprotonates 2-OH leading to its nucleophilic attack on to 1-phosphate group and formation of a cyclic intermediate. In this first step the oxygen of 1-phosphate group becomes transiently coordinated to calcium. His70 (His356, His82) then acts as a general acid, protonating the diacylglycerol leaving group. In the second step of the reaction, the cyclic intermediate undergoes hydrolysis. At this step, the acid/base roles were suggested to be reversed, so, in *Ps*PI-PLC, His70 (His356, His82) would activate a water molecule and His26 (His311, His32) or Glu48 (Glu341) would become a general acid.

CONCLUSIONS

A new member of the bacterial PI-PLC family, the first of its kind for a PI-PLC from a gram-negative bacterium, has been cloned, expressed, and its X-ray structure determined. It is only the second bacterial PI-PLC in which calcium binding has been reported, and the first where a calcium ion has been modelled in the structure. The presence of a catalytic calcium makes this new group of bacterial enzymes more suitable models for activity studies on catalytic domains of medically important mammalian PI-PLCs. However, for *San*PI-PLC, recent publications report a cleavage mechanism distinct from the "canonical" one, namely, trans-cyclization involving formation of inositol 1,6-cyclic phosphate intermediate rather than cis-cyclization to form 1nositol 1,2-cyclic phosphate occurring for all other known phospholipases C (Bai *et al.*, 2009, Bai *et al.*, 2010). In addition, comparisons show that the ligand-binding site of *Ps*PI-PLC is strikingly close to that of the mammalian enzymes and less similar to those from the closest bacterial structures. Taking into account these factors, the *Pseudomonas* enzyme seems a better model for eukaryotic catalytic domain than *San*Pi-PLC. Besides serving as a model system for mammalian PI-PLCs, the *Ps*PI-PLC structure provides insight into the gram-negative PI-PLCs.

Finally, phospholipases, including PLCs, are used in a number of industrial applications e.g. in the baking, egg yolk and dairy industries, and in plant oil degumming (De Maria *et al.*, 2007). Such enzyme-based procedures are much more environmentally friendly than chemical methods. The new structures give valuable information for further protein engineering to improve on the existing applications or create opportunities for the new ones.

ACKNOWLEDGEMENTS

We thank Diamond Light Source for access to beamline I03 (proposal number mx-9948) that contributed to the results presented here. We thank Johan Turkenburg and Sam Hart for support with the X-ray data collection.

FIGURE LEGENDS

Figure 1. Content of phosphatidic acid (PA), phosphatidylethanolamine (PE), phosphatidylinositol (PI) and phosphatidylcholine (PC) in crude soybean oil before (blank) and after incubation with 10 mg/kg and 100 mg/kg *Ps*PI-PLC, respectively, over 2 h at 50 °C at a range of pH.

Figure 2. Subgroup-supergroup relation between crystal Forms 1 and 2. (a) Black and orange colours designate crystallographic symmetry and pseudo-symmetry elements, respectively, in the Form 2 crystal. All these elements are crystallographic symmetry elements in Form 1, which has to be rotated by –45° about **c** and translated by **b**/ 2 from its standard setting for corresponding symmetry elements to overlay as shown. (Note that both diagonal pseudo-symmetry two-fold axes and diagonal crystallographic screw two-fold axes in Form 2 correspond to crystallographic coordinate two-fold axes in Form 1.) The black rectangle represents the unit cell in Form 2 and the orange rectangle represents the rotated and translated unit cell of Form 1. (*b*) Form 1 and (*c*) Form 2, the Patterson maps are coloured blue at 0.03 of the height of the origin peaks and the unit cells are shown as yellow rectangles. The full resolution range of the experimental data was used, and hence the Patterson peaks are narrower for Form 2 which diffracts to higher resolution. Peaks 1 and 1' are origin peaks and the heights of other peaks relative to corresponding origin peaks are 0.30 for peak 2, 0.21 for peak 2' and 0.10 for peak 3'. All non-origin peaks are split along **c** and their fractional coordinates along **c** are 0.5±0.017 for peak 2, 0.0±0.005 for peak 2' and 0.5±0.017 for peak 3'. Peaks that are not labelled are generated from equivalent peaks by crystallographic translations. Unlabelled images for (*b*, *c*) were obtained using *COOT*.

Figure 3. Ribbon representation of the *Ps*PI-PLC fold. The numbering of the β-strands in the $(\alpha\beta)_8$ barrel is as for bacterial PI-PLCs (I to VIII), but the numbering of the helices is different because of their different arrangement. The helices are instead numbered similarly to rat PI-PLC (1-7), with numbers following that of the previous β-strand (1 after I, etc.); when there are several helices between two strands, they have the same number followed by a letter. There is a 3_{10} helix between strands I and II (38-42), coloured in purple, with a number 1 in italics. Several very short 3_{10}

helices and β-strands are not shown. This and other structure figures were prepared using CCP4mg (McNicholas *et al.*, 2011).

Figure 4. Ribbon diagrams showing the superposition of *Ps*PI-PLC complex with myoinositol (gold) on (a) apo *San*PI-PLC (PDB code 3h4x; green, with the successful molecular replacement model highlighted in darker green, §2.3.1), (b) the *Bc*PI-PLC complex with myoinositol (1ptg, purple) and apo *Bt*PI-PLC with an engineered calcium site (1t6m, ice blue), (c) rat PI-PLC complex with D-myoinositol-1,4,5-triphosphate (1djx; pink), the zoom-in showing details of the active site. Superposition was carried out as detailed in Table 2.

Figure 5. Sequence alignment of *Ps*PI-PLC with *San*PI-PLC and rat PI-PLC, with calcium-binding residues (putative Ca binding for *San*PI-PLC) and catalytic histidines outlined in green. Secondary structure elements are shown as symbols for *Ps*PI-PLC and as background colours for the other two sequences. In both types of presentation, helices are red, 3_{10} helices yellow and β-strands violet. ClustalX (Larkin *et al.*, 2007) and structure superpositions described in Table 2 were used for initial alignment of X-boxes and its manual adjustment, respectively. The alignment for Y-boxes was obtained in a similar manner but required more substantial structure-based manual correction. The graphical representation of alignment shown here was prepared using the program ALINE (Bond & Schuttelkopf, 2009)

Figure 6. Active site and reaction mechanism of *Ps*PI-PLC. (*a*) Schematic representation of the hydrogen bonding in the active site. Dashed lines show potential hydrogen bonds made by the enzyme side chains to myoinositol, Ca^{2+} and water molecules; hydrogen bonds between the side chains are not shown. Bond distances (Å) are indicated for H-bonds involving myoinositol and Ca^{2+}. Trp262, shown as a red arch, is involved in a stacking interaction with the myoinositol ring. His70 is not shown as it makes no H-bonds to the phosphate-free inositol in the *Ps*PI-PLC structure, but it is shown in Fig. 7(c), where it occupies a very similar position to His356 of the superposed structure of the rat PI-PLC D-*myo*-inositol-1,4,5-trisphosphate complex. (*b*) The detailed mechanism shown here was initially demonstrated for rat PI-PLC (Essen *et al.*, 1996, Essen *et al.*, 1997) and adapted for *Ps*PI-PLC based on structural similarity between the mammalian and the bacterial enzymes. His26 (His311 in the mammalian enzyme) or Glu48 (Glu341) acts as a

general base and deprotonates the myoinositol OH-2 which then carries out a nucleophilic attack on the 1-phosphate group of the ligand resulting in a cyclic intermediate. The role of calcium is to stabilise the transition state, while His70 (His356) acts as a general acid and protonates the diacylglycerol leaving group. The cyclic intermediate undergoes hydrolysis in the reverse reaction with acid and base swapping their roles. LigPlot+ (Wallace *et al.*, 1995) with a distance threshold of 3.3 Å and ChemDraw (ChemDraw Prime 15.1, PerkinElmer) were used to generate (a) and (b) respectively.

Figure 7. Comparison of the binding sites for the PI-PLCs. Stereo diagrams show superposition of the binding site of the *Ps*PI-PLC complex with MI (gold) on the corresponding regions of (a) *San*PI-PLC (3h4x, green), (b) the *Bc*PI-PLC complex with MI (1ptg, purple) and the *Bt*PI-PLC R69D mutant (1t6m, ice blue), and (c) the rat PI-PLC complex with D-*myo*-inositol-1,4,5-trisphosphate (1djx, pink). Protein residues are shown as cylinders, ligands as ball and stick, and calcium ions as spheres. Residues used in the superpositions belong to the X-box domains (Table 2). In all panels (a-c), the catalytic His26 and His70 of *Ps*PI-PLC are in similar locations to their counterparts in the matched structures. The best alignment for His70 (general acid), is achieved in (c), in the superposition with the rat PI-PLC. The planes of His26 (possible general base) in the *Ps*PI-PLC structure and its counterpart (His311) in rat PI-PLC have different orientations, possibly because inositol is not phosphorylated in the former structure, while 1-phosphate is present and makes H-bond to His311 in the latter. In (c) Glu48 in *Ps*PI-PLC is seen to align perfectly with Glu341 in rat PI-PLC (a candidate for the general base in the mammalian enzymes; Essen et al., 1997), but has no clear counterpart in the wild type bacterial enzymes (a, b). The difference in the position of the inositol O2 between *Ps*PI-PLC and rat PI-PLC can be attributed to the lack of a 1-phosphate in the former. Ca^{2+} is present only in *Ps*PI-PLC, rat PI-PLC and the *Bt*PI-PLC R69D mutant and superposes most closely in the first two. Overall, the *Ps*PI-PLC active site has more similarity to the active site of the rat enzyme (c) than the bacterial enzymes (a, b) suggesting a mammalian-like catalytic mechanism for *Ps*PI-PLC.

TABLE 1.

Crystallographic statistics.

Data collection [a]			
Data set	Native Form 1	Native Form 2	Myoinositol complex
Beamline	Diamond I03	Diamond I03	Diamond I03
Wavelength (Å)	0.98	0.98	0.98
PDB code	5fyo	5fyp	5fyr
Space group	$P4_322$	$P4_32_12$	$P4_32_12$
Cell parameters (Å)	a=96.08 b=96.08 c=113.3	a=135.21 b=135.21 c=112.47	a=135.59 b=135.59 c=113.79
Resolution range (Å)	48.81-1.5 (1.53-1.50)	67.61-1.17 (1.20-1.17)	48.93-1.45 (1.47-1.45)
Number of reflections	1097060	4032897	1334118
Unique reflections	85192	346088	184175
Monomers in asymmetric unit	2	4	4
Completeness (%)	99.9 (100)	99.9 (99.8)	99.7 (95.9)
$\langle I / \sigma(I) \rangle$	27.5 (4.8)	17.2 (2.5)	14.5 (2.5)
$CC_{1/2}$ [b]	0.999 (0.938)	0.999 (0.733)	0.999 (0.786)
Multiplicity	12.9 (12.4)	11.7 (4.8)	7.2 (5.7)
R_{merge} [c]	0.049 (0.535)	0.075 (0.531)	0.067 (0.595)
Refinement statistics			
Fraction of free reflections	0.051	0.050	0.050
Final R_{cryst}	0.125	0.116	0.144
Final R_{freet}	0.166	0.141	0.192
R.m.s. deviations from ideal geometry (target values are given in parentheses)			
Bond distances (Å)	0.015(0.019)	0.013(0.019)	0.015(0.019)

Bond angles (°)	1.57 (1.92)	1.58 (1.92)	1.623(1.923)
Chiral centres (Å^3)	0.132 (0.200)	0.104 (0.200)	0.113(0.200)
Planar groups (Å)	0.009 (0.021)	0.011 (0.021)	0.009(0.021)
Average main chain B values (Å^2)	21.9	10.1	14.6
Average side chain B values (Å^2)	25.1	12.6	17.6
Average B values for Ca (Å^2)	23.1	13.0 (occ=0.8)	16.1
Average B values for Ins (Å^2)	N/A	N/A	13.8
Molprobity score	0.83	0.87	1.1
Ramachandran favoured (%) [(d)]	97.5	97.5	97.4
Ramachandran outliers (%) [(d)]	0.2	0.1	0.1
Clashscore	1.22	0.86	2.17

[a] values in parentheses correspond to the highest resolution shell

[c] $CC_{1/2}$ values for Imean are calculated by splitting the data randomly in half

[b] R_{merge} is defined as $\Sigma \left| I - \langle I \rangle \right| / \Sigma I$, where I is the intensity of the reflection

[d] Ramachandran plot analysis was carried out by MOLPROBITY (Chen *et al.*, 2010)

TABLE 2

Sequence and structure comparisons of *Ps*PI-PLC with bacterial PI-PLCs and rat PI-PLC. One representative structure per organism (liganded if available) was chosen for comparison with *Ps*PI-PLC complex with Ins. Sequence identity with *Ps*PI-PLC is shown for the full length bacterial enzymes and for the region including only X-box (300-440) and Y-box (487-606) of the mammalian (rat) PI-PLC. Structural alignments were conducted using SSM (Krissinel & Henrick, 2004) as implemented in CCP4mg (McNicholas *et al.*, 2011) for the X-boxes only, which form the most conserved parts of the proteins and contain all the residues essential for catalysis. The table shows the input residue ranges and C^{α} r.m.s.d.s reported by the program. The superposed structures were also used for manual adjustments of ClustalX sequence alignments (Fig. 5) and for generation of Figs. 4 and 7 showing superposition of the complete molecules and their active sites, respectively

Species	PDB ID	ligand *)	sequence identity to *Ps*PI-PLC	C^{α} r.m.s.d. to *Ps*PI-PLC, Å for the X-boxes (and input residue ranges)
P. sp-62186	5fyr	Ins	1.000	0.00 (2-190)**)
S. antibioticus	3h4x	none	0.271	1.60 (25-223)**)
B. thuringiensis	1t6m	none	0.200	2.22 (5-162)**)
B. cereus	1ptg	Ins	0.185	2.25 (1-166)**)
S. aureus	3v1h	Ins	0.185	2.27 (7-166)
L. monocytogenes	1aod	Ins	0.158	2.15 (21-177)
R. norvegicus	1djx	I3P	0.206	2.10 (300-440)**)

*) abbreviations for ligands: Ins – myoinositol, I3P D-myoinositol-1,4,5-triphosphate.

**) structures shown in Fig. 4 and Fig.7

REFERENCES

Ames, G. F. (1968). *J Bacteriol* **95**, 833-843.

Andreeva, A., Howorth, D., Chandonia, J. M., Brenner, S. E., Hubbard, T. J. P., Chothia, C. & Murzin, A. G. (2008). *Nucleic Acids Res* **36**, D419-D425.

Apiyo, D., Zhao, L., Tsai, M. D. & Selby, T. L. (2005). *Biochemistry-Us* **44**, 9980-9989.

Bai, C., Zhao, L., Rebecchi, M., Tsai, M. D. & Bruzik, K. S. (2009). *J Am Chem Soc* **131**, 8362-8363.

Bai, C., Zhao, L., Tsai, M. D. & Bruzik, K. S. (2010). *J Am Chem Soc* **132**, 1210-1211.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D. & Waley, S. G. (1975). *Nature* **255**, 609-614.

Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. (2004). *J Mol Biol* **340**, 783-795.

Bishop, D. G., Rutberg, L. & Samuelsson, B. (1967). *Eur J Biochem* **2**, 448-453.

Bond, C. S. & Schuttelkopf, A. W. (2009). *Acta Crystallogr D* **65**, 510-512.

Bunkoczi, G. & Read, R. J. (2011). *Acta Crystallogr D* **67**, 303-312.

Bunney, T. D. & Katan, M. (2011). *Trends Biochem Sci* **36**, 88-96.

Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Crystallogr D Biol Crystallogr* **66**, 12-21.

Cocco, L., Follo, M. Y., Manzoli, L. & Suh, P. G. (2015). *J Lipid Res* **56**, 1853-1860.

De Maria, L., Vind, J., Oxenboll, K. M., Svendsen, A. & Patkar, S. (2007). *Appl Microbiol Biot* **76**, 235-235.

Diderichsen, B., Poulsen, G. B. & Jorgensen, S. T. (1993). *Plasmid* **30**, 312-315.

Ellis, M. V., James, S. R., Perisic, O., Downes, C. P., Williams, R. L. & Katan, M. (1998). *J Biol Chem* **273**, 11650-11659.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Crystallogr D* **66**, 486-501.

Essen, L. O., Perisic, O., Cheung, R., Katan, M. & Williams, R. L. (1996). *Nature* **380**, 595-602.

Essen, L. O., Perisic, O., Katan, M., Wu, Y. Q., Roberts, M. F. & Williams, R. L. (1997). *Biochemistry-Us* **36**, 1704-1718.

Fukami, K., Inanobe, S., Kanemaru, K. & Nakamura, Y. (2010). *Prog Lipid Res* **49**, 429-437.

Goldstein, R., Cheng, J. J., Stec, B. & Roberts, M. F. (2012). *Biochemistry-Us* **51**, 2579-2587.

Griffith, O. H. & Ryan, M. (1999). *Biochim Biophys Acta* **1441**, 237-254.

Heinz, D. W., Essen, L. O. & Williams, R. L. (1998). *J Mol Biol* **275**, 635-650.

Heinz, D. W., Ryan, M., Bullock, T. L. & Griffith, O. H. (1995). *Embo J* **14**, 3855-3863.

Hocker, B., Jurgens, C., Wilmanns, M. & Sterner, R. (2001). *Curr Opin Biotech* **12**, 376-381.

Iwasaki, Y., Niwa, S., Nakano, H., Nagasawa, T. & Yamane, T. (1994). *Bba-Lipid Lipid Met* **1214**, 221-228.

Iwasaki, Y., Tsubouchi, Y., Ichihashi, A., Nakano, H., Kobayashi, T., Ikezawa, H. & Yamane, T. (1998). *Bba-Lipid Lipid Met* **1391**, 52-66.
Jackson, M. R. & Selby, T. L. (2012). *Acta Crystallogr F* **68**, 1378-1386.
Kabsch, W. (2010). *Acta Crystallogr D* **66**, 125-132.
Katan, M. (1998). *Bba-Mol Cell Biol L* **1436**, 5-17.
Kravchuk, A. V., Zhao, L., Bruzik, K. S. & Tsai, M. D. (2003). *Biochemistry-Us* **42**, 2422-2430.
Krissinel, E. & Henrick, K. (2004). *Acta Crystallogr D Biol Crystallogr* **60**, 2256-2268.
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics* **23**, 2947-2948.
Lebedev, A. A. & Isupov, M. N. (2014). *Acta Crystallogr D* **70**, 2430-2443.
Majerus, P. W., Connolly, T. M., Deckmyn, H., Ross, T. S., Bross, T. E., Ishii, H., Bansal, V. S. & Wilson, D. B. (1986). *Science* **234**, 1519-1526.
Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C. & Bryant, S. H. (2015). *Nucleic Acids Res* **43**, D222-226.
Marques, M. B., Weller, P. F., Parsonnet, J., Ransil, B. J. & Nicholson-Weller, A. (1989). *J Clin Microbiol* **27**, 2451-2454.
Mccoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J Appl Crystallogr* **40**, 658-674.
McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. (2011). *Acta Crystallogr D Biol Crystallogr* **67**, 386-394.
Moser, J., Gerstel, B., Meyer, J. E. W., Chakraborty, T., Wehland, J. & Heinz, D. W. (1997). *J Mol Biol* **273**, 269-282.
Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Crystallogr D* **67**, 355-367.
Nagano, N., Orengo, C. A. & Thornton, J. M. (2002). *J Mol Biol* **321**, 741-765.
Rhee, S. G., Suh, P. G., Ryu, S. H. & Lee, S. Y. (1989). *Science* **244**, 546-550.
Shah, A. K., Liu, Z. J., Stewart, P. D., Schubot, F. D., Rose, J. P., Newton, M. G. & Wang, B. C. (2005). *Acta Crystallogr D Biol Crystallogr* **61**, 123-129.
Shaw Stewart, P. D., Kolek, S. A., Briggs, A. R., Chayen, N. E. & Baldock, P. F. M. (2011). *Crystal Growth & Design* **11**, 3432-3441.
Singh, A., Bhatnagar, N., Pandey, A. & Pandey, G. K. (2015). *Cell Calcium* **58**, 139-146.
Smith, G. A., Marquis, H., Jones, S., Johnston, N. C., Portnoy, D. A. & Goldfine, H. (1995). *Infect Immun* **63**, 4231-4237.
Suh, P. G., Ryu, S. H., Moon, K. H., Suh, H. W. & Rhee, S. G. (1988). *Cell* **54**, 161-169.
Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). *Protein Eng* **8**, 127-134.
White, M. J., Boyd, J. M., Horswill, A. R. & Nauseef, W. M. (2014). *Infect Immun* **82**, 1559-1571.
Widner, B., Thomas, M., Sternberg, D., Lammon, D., Behr, R. & Sloma, A. (2000). *J Ind Microbiol Biot* **25**, 204-212.
Williams, R. L. (1999). *Bba-Mol Cell Biol L* **1441**, 255-267.
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., McCoy, A., McNicholas, S. J.,

Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Crystallogr D Biol Crystallogr* **67**, 235-242.

Winter, G., Lobley, C. M. C. & Prince, S. M. (2013). *Acta Crystallogr D* **69**, 1260-1273.

Yao, L. X. & Jung, S. (2010). *J Agr Food Chem* **58**, 4866-4872.

Zhao, J. F., Chen, H. H., Ojcius, D. M., Zhao, X., Sun, D., Ge, Y. M., Zheng, L. L., Lin, X. A., Li, L. J. & Yan, J. (2013). *Plos One* **8**.

Zhao, L., Liu, Y. H., Bruzik, K. S. & Tsai, M. D. (2003). *J Am Chem Soc* **125**, 22-23.

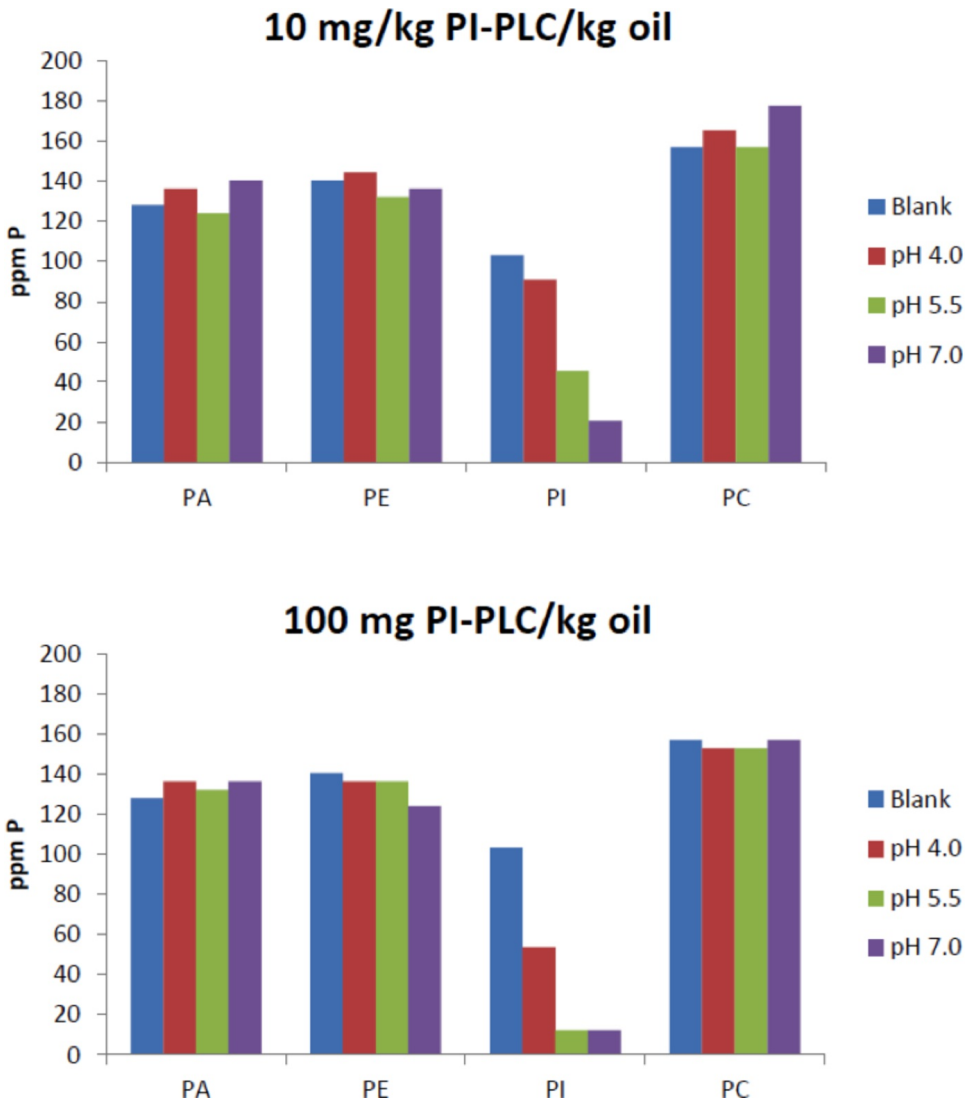
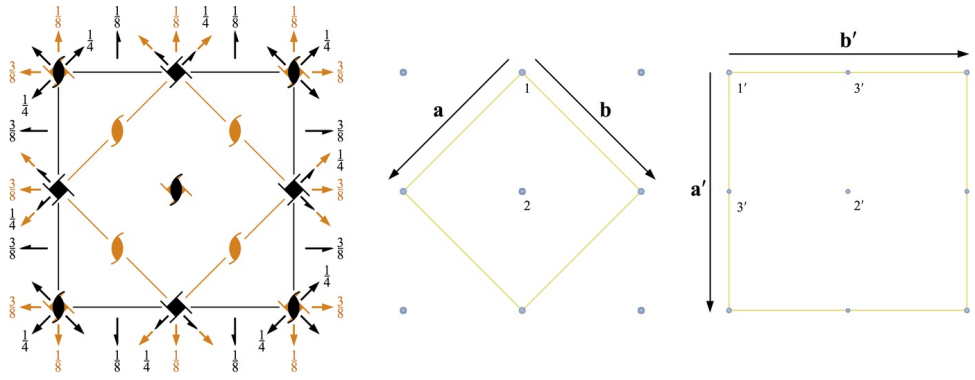

10 mg/kg PI-PLC/kg oil
ppm P
200
180
160
140
120
100
80
60
40
20
0
PA
PE
PI
PC
Blank
pH 4.0
pH 5.5
pH 7.0
100 mg PI-PLC/kg oil
ppm P
200
180
160
140
120
100
80
60
40
20
0
PA
PE
PI
PC
Blank
pH 4.0
pH 5.5
pH 7.0

Figure 1



(*a*) (*b*) (*c*)

Figure 2

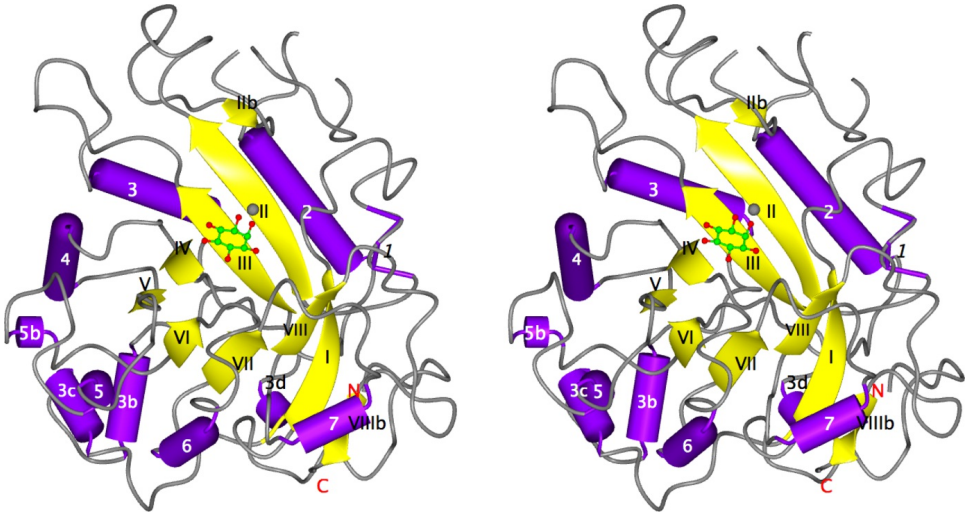

IIb
3
II
2
1
4
IV
III
V
5b
VI
VIII
I
VII
3d
3c
5
3b
N
7
VIIIb
6
C
IIb
3
II
2
1
4
IV
III
V
5b
VI
VIII
I
VII
3d
3c
5
3b
N
7
VIIIb
6
C

Figure 3

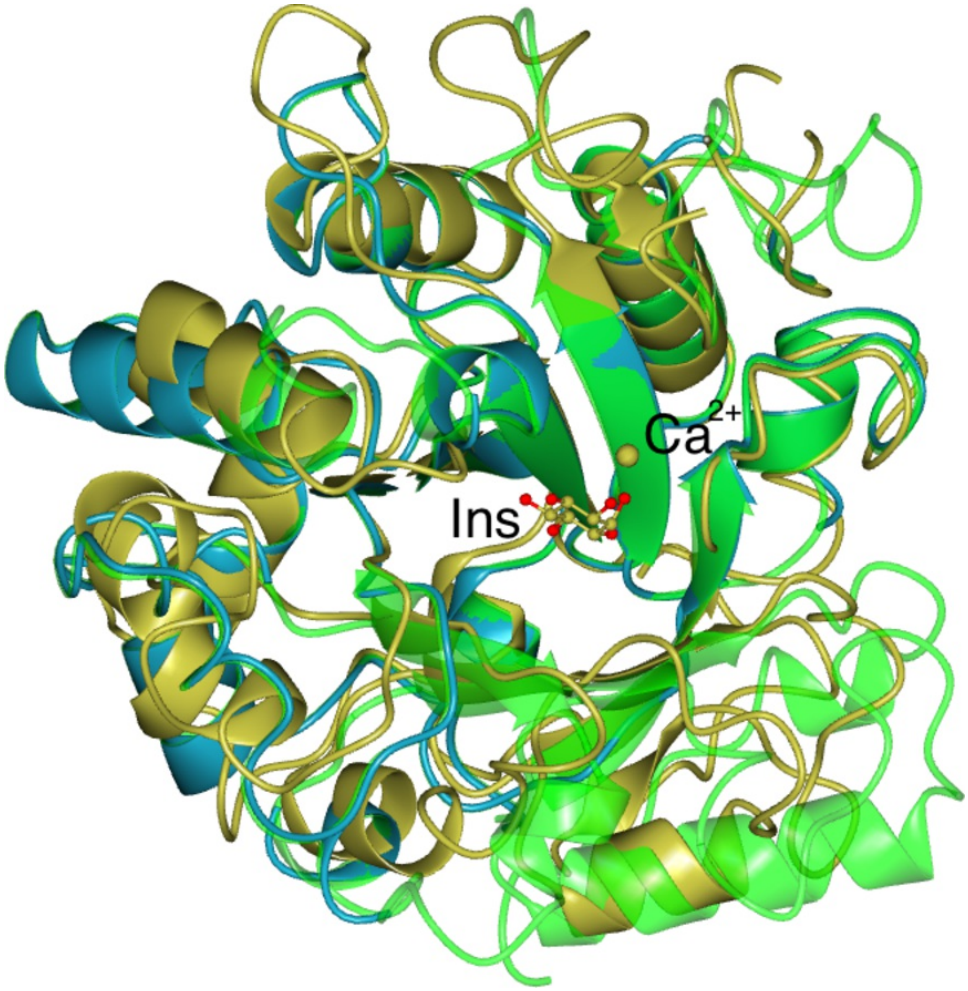


Figure 4a

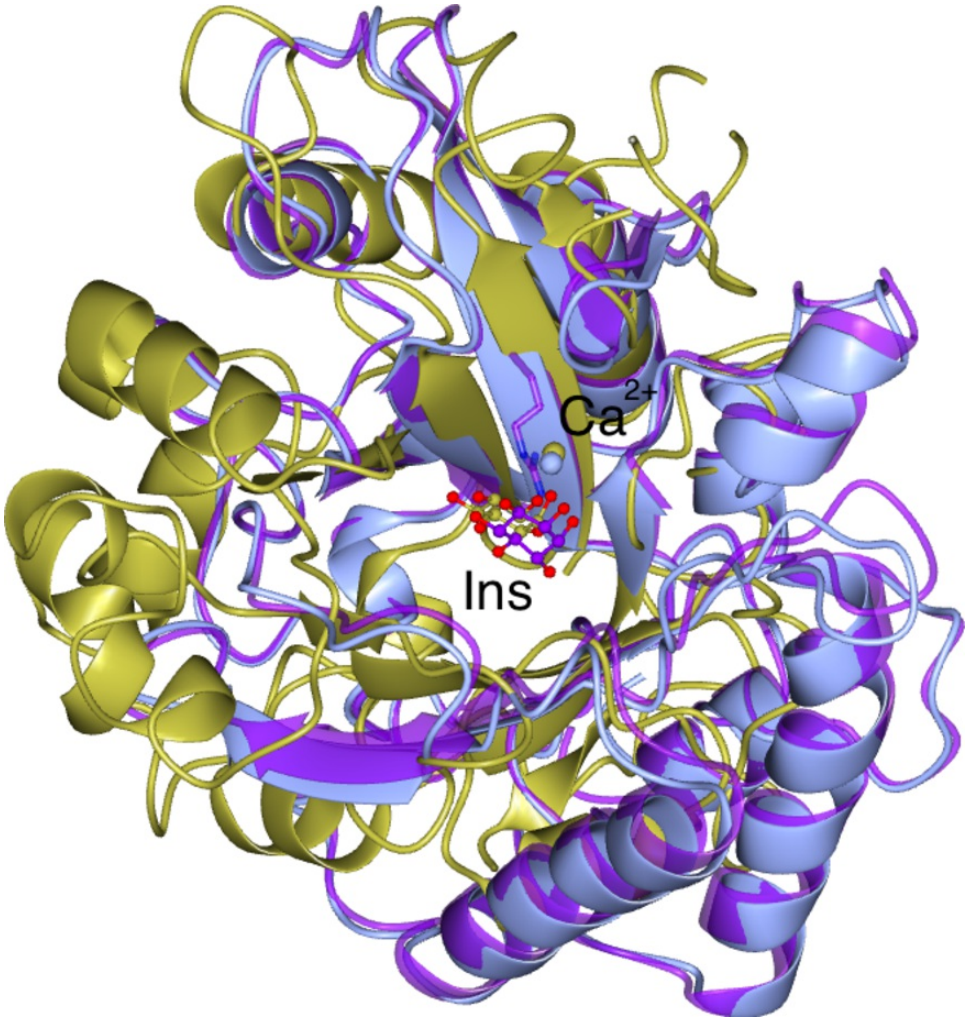


Figure 4b

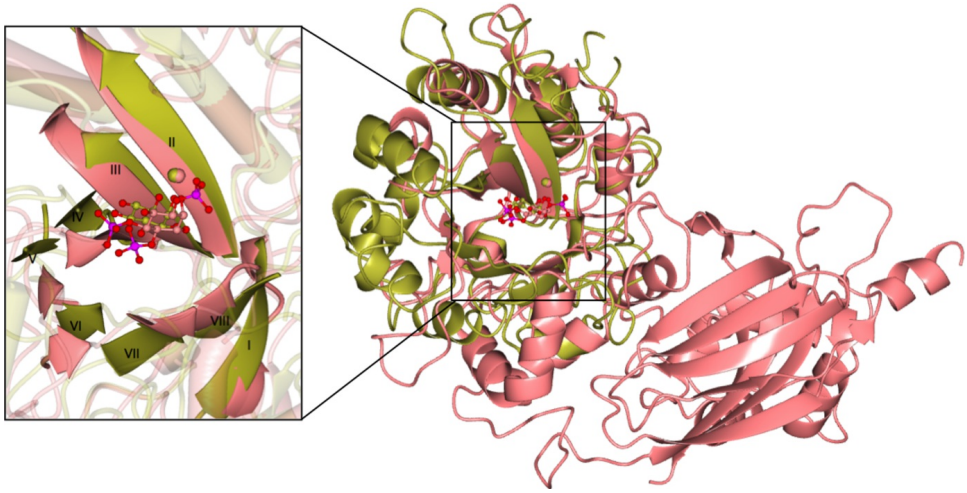


Figure 4c

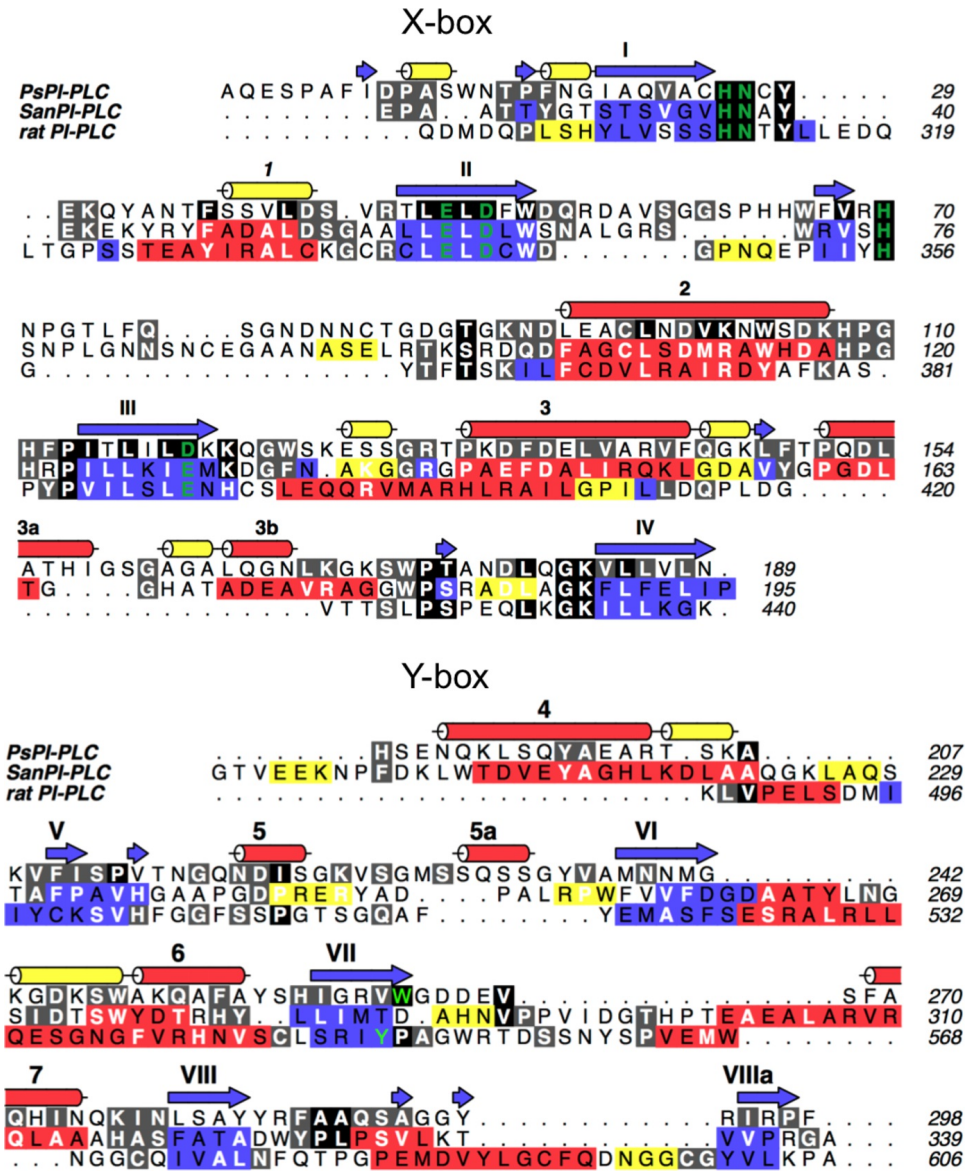

X-box
PsPI-PLC
SanPI-PLC
rat PI-PLC
Y-box
PsPI-PLC
SanPI-PLC
rat PI-PLC

Figure 5

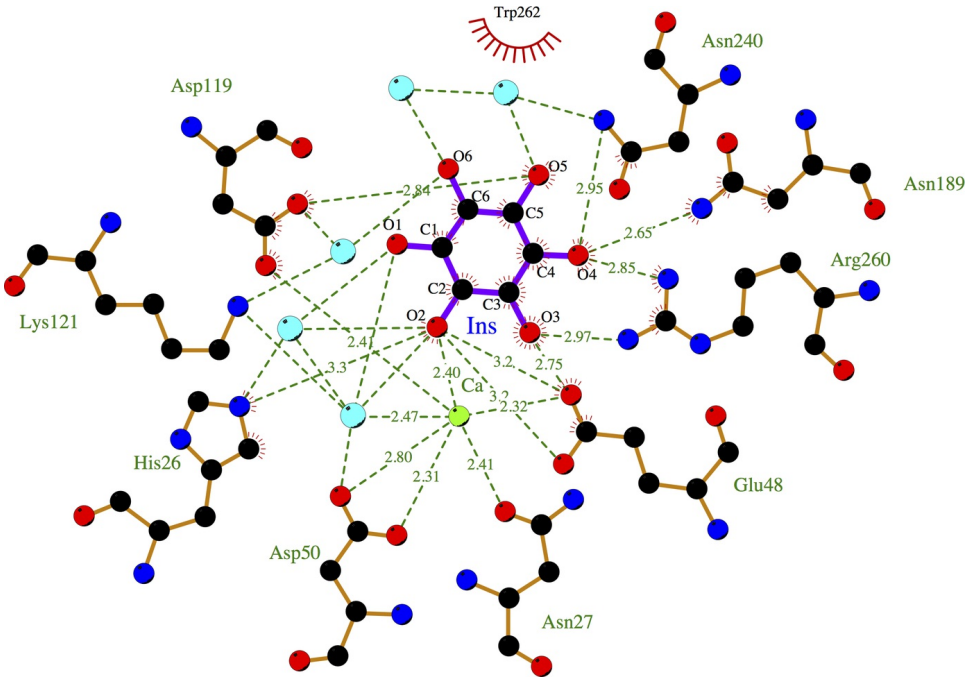

Trp262
Asn240
Asp119
Asn189
Arg260
Lys121
His26
Glu48
Asp50
Asn27
Ins
Ca
O1
O2
O3
O4
O5
O6
C1
C2
C3
C4
C5
C6
2.84
2.95
2.65
2.85
2.97
2.41
3.3
2.40
3.2
2.75
3.2
2.32
2.47
2.80
2.31
2.41

Figure 6a

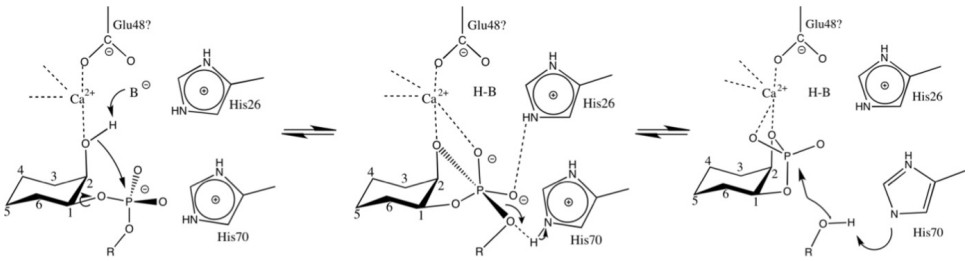


Figure 6b

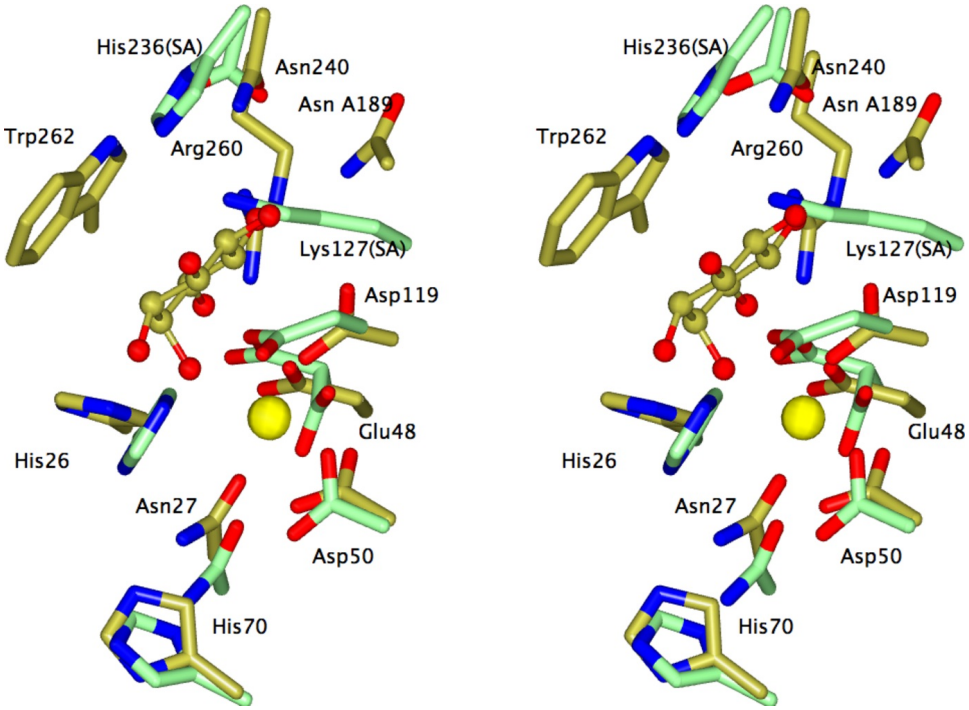

His236(SA)
Asn240
Asn A189
Trp262
Arg260
Lys127(SA)
Asp119
Glu48
His26
Asn27
Asp50
His70
His236(SA)
Asn240
Asn A189
Trp262
Arg260
Lys127(SA)
Asp119
Glu48
His26
Asn27
Asp50
His70

Figure 7a

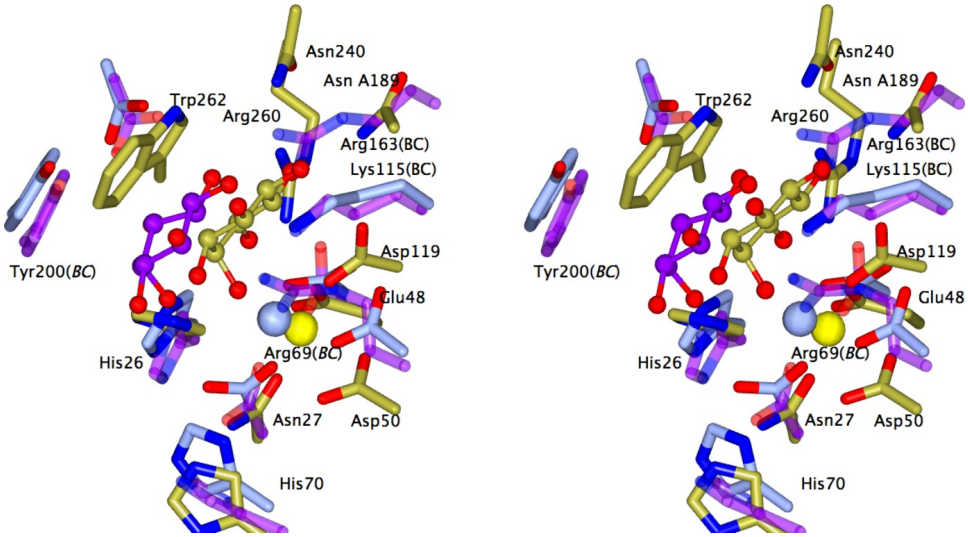

Asn240
Asn A189
Trp262
Arg260
Arg163(BC)
Lys115(BC)
Tyr200(BC)
Asp119
Glu48
His26
Arg69(BC)
Asn27
Asp50
His70
Asn240
Asn A189
Trp262
Arg260
Arg163(BC)
Lys115(BC)
Tyr200(BC)
Asp119
Glu48
His26
Arg69(BC)
Asn27
Asp50
His70

Figure 7b

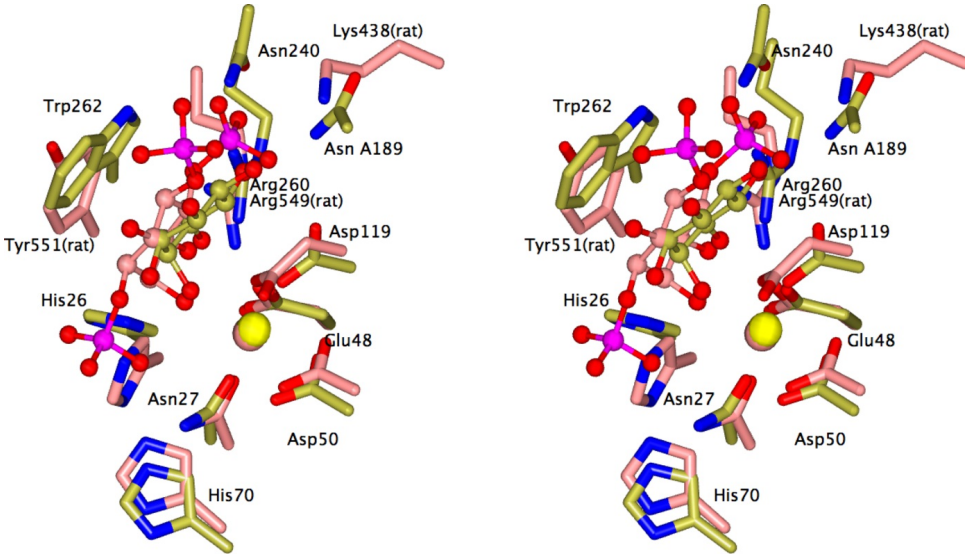

Lys438(rat)
Asn240
Trp262
Asn A189
Arg260
Arg549(rat)
Tyr551(rat)
Asp119
His26
Glu48
Asn27
Asp50
His70
Lys438(rat)
Asn240
Trp262
Asn A189
Arg260
Arg549(rat)
Tyr551(rat)
Asp119
His26
Glu48
Asn27
Asp50
His70

Figure 7c