

Inositol 1,3,4,5,6-pentakisphosphate 2-kinase is a distant IPK member with a singular inositide binding site for axial 2-OH recognition

Beatriz González^{a,1}, Jose Ignacio Baños-Sanz^a, Maider Villate^a, Charles Alistair Brearley^b, and Julia Sanz-Aparicio^a

^aGrupo de Cristalografía Macromolecular y Biología Estructural, Instituto de Química-Física Rocasolano, Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain; and ^bSchool of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

Edited by Susan S. Taylor, University of California at San Diego, La Jolla, CA, and approved March 24, 2010 (received for review November 10, 2009)

Inositol phosphates (InsPs) are signaling molecules with multiple roles in cells. In particular Ins(1,2,3,4,5,6)P₆ (InsP₆) is involved in mRNA export and editing or chromatin remodeling among other events. InsP₆ accumulates as mixed salts (phytate) in storage tissues of plants and plays a key role in their physiology. Human diets that are exclusively grain-based provide an excess of InsP₆ that, through chelation of metal ions, may have a detrimental effect on human health. Ins(1,3,4,5,6)P₅ 2-kinase (InsP₅ 2-kinase or Ipk1) catalyses the synthesis of InsP₆ from InsP₅ and ATP, and is the only enzyme that transfers a phosphate group to the axial 2-OH of the myo-inositide. We present the first structure for an InsP₅ 2-kinase in complex with both substrates and products. This enzyme presents a singular structural region for inositide binding that encompasses almost half of the protein. The key residues in substrate binding are identified, with Asp368 being responsible for recognition of the axial 2-OH. This study sheds light on the unique molecular mechanism for the synthesis of the precursor of inositol pyrophosphates.

InsP5 2-kinase | IP5 | IP6 | ipk1 | crystal structure

Inositol phosphates (InsPs) are soluble signaling molecules that have diverse and important biological roles (1, 2). Inositol 1,2,3,4,5,6-hexakisphosphate (InsP₆ or IP₆), known as phytic acid, plays an essential role in cell biology. For example, IP₆ is involved in processes as diverse as mRNA export (3), DNA repair (4), maintenance of basal resistance to plant pathogens (5), apoptosis (6), and regulation of chromatin structure (7, 8). In addition, IP₆ is the precursor of pyrophosphate inositols IP₇ and IP₈ (9, 10). A primary aspect of IP₆ is the key role it plays in plant biology, reflected in considerable commercial interest in this compound (11, 12). IP₆ constitutes the major phosphorus reserve in plant seeds, being a source not only of phosphorus, but also of inositol and minerals during development and germination (12). Concerns arise from the fact that IP₆ can present detrimental effects on human health and on the environment (reviewed in ref. 13). Grain-based diets provide high levels of IP₆, a potent chelator of essential ions as Zn²⁺ and Fe²⁺, causing deficiencies that contribute to malnutrition. Monogastric animals are unable to digest IP₆, which is excreted as salts that contribute to eutrophication of waterways. There is, consequently, considerable interest in the manipulation of plant phytate, largely by genetic approaches, though the desirability of this is not without dispute (11–15). Indeed, there are several lines of evidence that show benefits of IP₆ as antioxidant or a potential antitumor agent (16, 17). The IP₆ precursor, inositol 1,3,4,5,6-pentakisphosphate (InsP₅ or IP₅), also presents a role in induction of apoptosis (18).

The family of enzymes that catalyzes the phosphorylation of IP₅ to form IP₆ are known as Ipk1, after the ortholog first identified in yeast (3) or InsP₅ 2-kinases (IP₅ 2-Ks), after the enzyme activity characterized from various sources (19–24). It is only recently, however, that important physiological roles other than a biosynthetic enzyme have been ascribed to IP₅ 2-K. Thus, disruption of the IP₅ 2-K gene (*Ipk1*) yields non viable murine embryos (25). IP₅ 2-K is also involved in the establishment of left–right

organ asymmetry in mammals (26). As a family, IP₅ 2-Ks present very low sequence homology from yeast to mammals, with only a few sequence motifs being conserved. IP₅ 2-Ks are unique among inositol phosphate kinases (InsP Ks) in that they phosphorylate the axial 2-position of the myo-inositide ring, whereas the other enzymes act on equatorial positions of the ring.

The first structure of an InsP K became available in 2004 (27, 28), from IP₃ 3-Kinase (IP₃ 3-K), a family not represented in yeast or plant genomes. Based on sequence similarity, the InsP K enzymes encompass three different structural families. The first family is formed by the inositol polyphosphate kinases (IPKs), to which IP₃ 3-Ks belong, and this family displays some degree of structural homology with protein kinases (PKs) (27). Another family of InsP Ks have been shown to display the ATP-grasp fold (10, 29). A third structural family is formed by IP₅ 2-Ks, their structure remaining unknown, although Cheek et al. (30) have postulated that IP₅ 2-K could display some structural similarity to IPKs.

Herein we present the crystal structure of *Arabidopsis thaliana* IP₅ 2-K (*At-IP₅ 2-K*, *AtIpk1*), as well as a detailed description of the substrate and product interactions that have allowed us an in-depth characterization of this family fold and active site. The structure of IP₅ 2-K demonstrates membership, at a distance, of the IPK family of proteins. However, the IP₅ 2-K structure shows a unique and large inositide-binding pocket, which explains how the myo-inositidic 2-OH axial position that accepts the phosphate is recognised. In addition, IP₅ 2-K structure reveals key features of IPKs that elucidate the functionality of this essential family of enzymes.

Results

IP₅ 2-K Structure. The structure of *At-IP₅ 2-K* has been solved by multiwavelength anomalous dispersion (MAD) techniques and the structures of different IP₅ 2-K crystal complexes have been obtained: two binary complexes (IP₅ 2-K/IP₅ and IP₅ 2-K/IP₆), and two ternary complexes (IP₅ 2-K/IP₅/AMPPNP and IP₅ 2-K/IP₆/ADP) (Table S1 and Fig. S1). IP₅ 2-K fold consists of two lobes connected by a hinge, which we will refer as N-terminal lobe (residues 1–152) and C-terminal lobe (161–432) (Fig. 1 A and B). Within the C-lobe there are three noncontiguous regions folded in a unique assembly involved in the inositide binding (C_{IP}-lobe).

Author contributions: B.G., C.A.B., and J.S.-A. designed research; B.G., J.I.B.-S., and M.V. performed research; B.G., J.I.B.-S., M.V., C.A.B., and J.S.-A. analyzed data; and B.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The structural coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2xal, 2xam, 2xan, 2xao and 2xar).

¹To whom correspondence should be addressed. E-mail: xbeatriz@iqfr.csic.es.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.0912979107/-DCSupplemental.

The N-terminal lobe (N-lobe) has an $\alpha + \beta$ fold (Fig. 1B and Fig. S2A), the core of this lobe being formed by an antiparallel and twisted main β -sheet comprised of five β -strands in 1-2-3-6-4 order. Strands $\beta 1$ – $\beta 2$ are connected by a glycine rich loop that is involved in nucleotide binding. Strands $\beta 3$ – $\beta 4$ and $\beta 4$ – $\beta 6$ are connected through long regions built up mainly of α -helices, named as “N-I” and “N-II,” respectively. N-I is comprised of a mobile loop and three antiparallel helices ($\alpha 2$ – $\alpha 4$). Finally, N-II starts with a helix ($\alpha 5$) that crosses the whole β -sheet, followed by a small and antiparallel helix and the short $\beta 5$.

The C-lobe is almost twice the size of the N-lobe and also has an $\alpha + \beta$ fold (Fig. 1B and Fig. S2B). The β -strands form an antiparallel twisted β -sheet connected in order 9-8-7-10-12-11. The α -helices in this lobe are distributed in three noncontinuous regions that emerge from the β -sheet and comprise almost half of the protein. Helices $\alpha 11$ and $\alpha 12$, preceded by the long loop L3, fold over the β -sheet. All other α -helices form a well-defined assembly with a very irregular fold that we have named as C_{IP} lobe, because it constitutes the main inositide-binding site. The first region (C_{IP-I}), connects strands $\beta 7$ – $\beta 8$ and is formed by $\alpha 7$ – $\alpha 10$. C_{IP-I} starts with a long loop (L1) followed by α -helices that come back to the β -sheet in a zigzag fashion, ending at the loop L2. The second region (C_{IP-II}) connects $\alpha 12$ – $\beta 10$ and is comprised of $\alpha 13$ – $\alpha 16$, which form a kind of coil joined to the β -sheet by the short L4. Finally, C_{IP-I} and C_{IP-II} are held together through L1, and the C-terminal helix. This helix and its preceding loop L5 constitute C_{IP-III} . The loops L1–L5, located at the junction of the C-lobe β -sheet and helices, are related to four IP₅ 2-K conserved sequence motifs (motif I–IV) (Fig. 1C). There are two additional motifs conserved along IP₅ 2-Ks, located in $\alpha 8$ and $\alpha 13$ – $\alpha 14$ of the C_{IP} -lobe as will be seen later.

IP₅ 2-K crystallizes with two molecules in the asymmetric unit. There are only a few hydrogen bonds between the two molecules, with 1, 101 Å² buried surface. This suggests that it is not likely to be a physiological dimer. No significant differences are seen in the active sites of the two molecules.

IP₅ 2-K Substrate and Product Recognition.

Nucleotide-binding site. The ATP analogue is bound between the N- and C-lobes, at the interface created by both β -sheets of the enzyme (Figs. 1A and 2A, and Fig. S3). The adenine N1 and N6 are linked to hinge residues (Asn147 and His149, respectively). Several residues from both lobes provide a hydrophobic environment for the adenine and ribose: Arg16, Val24, Val38, Leu146, His149, Met372, and Ile406. The oxygen atoms of the ribose are coordinated to Glu166. A weaker interaction is found between the ribose O3* and Arg241. P α is coordinated to Arg40, whereas the P β is coordinated to several residues from the G-loop (Asn22, and the main chain of Gly19, Gly20, Ala21, and Asn22) and Arg40. Finally, P γ is recognised by the G-loop Gly20, Lys168, Asp407, and Ser409. Interactions with the latter two residues occur via Mg1, which in turn is coordinated to P β . P γ is further coordinated by Asp 368.

Interestingly, the conserved motifs I–IV are all involved in nucleotide recognition and in particular, residues located at L1, L4, and L5 surround the phosphoryl transference region (Fig. 2D).

Inositide-binding site. IP₅ 2-K substrate (IP₅) is tightly bound to the enzyme through all phosphates groups (Fig. 2B and Fig. S3). P1 makes two hydrogen bonds with Arg130 from the N-lobe, whereas P3 is bound to both the N-lobe (Arg45) and the C-lobe (Lys411 and Arg415). Interestingly, P4, P5, and P6 are coordinated only by residues from the C_{IP} -lobe. P4 coordinates to Lys411, Arg415, and Tyr419, P5 coordinates to Lys170 (and to His196 and Arg192 through water molecules), and P6 coordinates to Lys170, Lys200, and Asn238. A water-mediated link is also produced with Asn238 and Asn239. Finally, the acceptor 2-OH interacts to Asp368, and Lys168. Asp368 seems to be

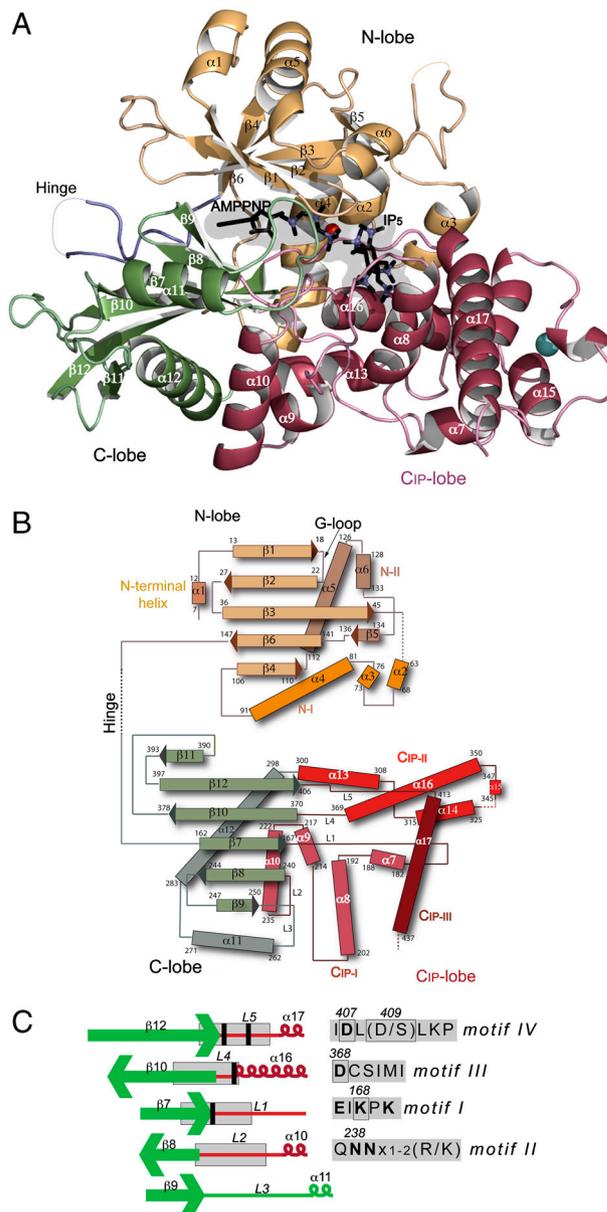


Fig. 1. The structure of IP₅ 2-K from *A. thaliana*. (A) Cartoon representation of *At*-IP₅ 2-K structure, in which the N-lobe is colored in light orange, the hinge in slate blue, the C-lobe β -sheet region in green, and the C_{IP} -lobe in raspberry. The substrates are shown as black sticks, with the phosphates as slate spheres. The magnesium and zinc ions are shown as red and cyan spheres, respectively. (B) IP₅ 2-K topology diagram. The color code is similar, applying different orange shades for the N-lobe regions (β -sheet, N-terminal helix, N-I, and N-II) and different red shades for the three C_{IP} elements (C_{IP-I} , C_{IP-II} and C_{IP-III}). (C) Scheme of the C-lobe β -sheet and α -helices interface showing loops L1 to L5. The gray shade squares show the position of four IP₅ 2-K's conserved motifs. Residues invariant along the species are in bold letter. Relevant residues for catalysis are squared and its position is represented as black rods.

essential in recognizing of the inositide 2-OH, likely conferring specificity for the substrate conformer.

In conclusion, the C_{IP} -lobe provides residues for binding all IP₆ phosphates, except P1, coming from motifs I–IV, motif ⁹²RxxMHQxLK in $\alpha 8$, and $\alpha 17$. The N-lobe helps to inositide binding by coordinating P1 and P3. The strong interaction that all IP₅ phosphates establish with the protein explains why other less substituted InsPs lacking phosphates on 1-, 3-, or 5- or 3,5-positions are substrates of IP₅ 2-K (20, 24, 31).

Finally, L3 seems to constrict the inositide-binding site by making interactions with $\alpha 6$ from N-lobe and motif II from

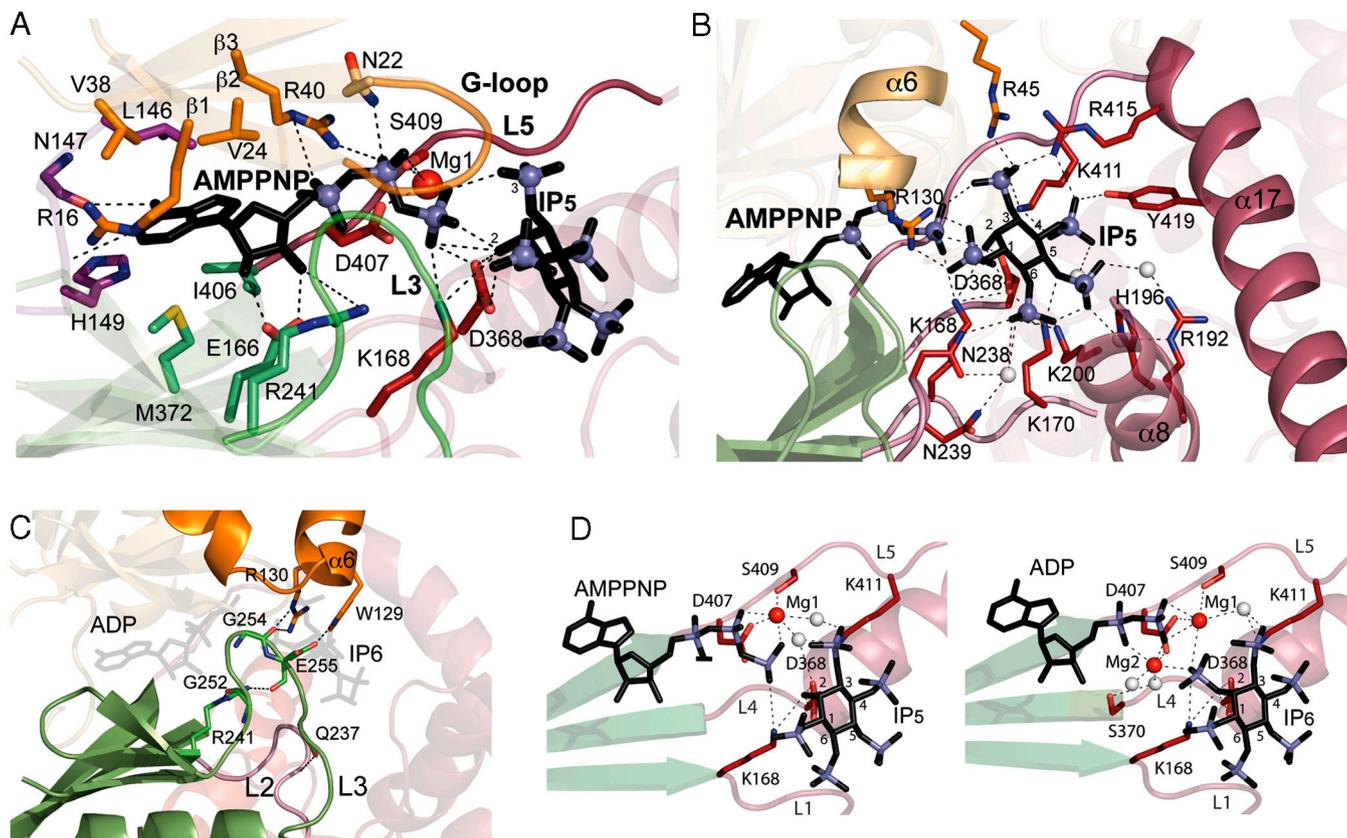


Fig. 2. IP₅ 2-K active site. AMPPNP (A) and IP₅ (B) interactions as found in the ternary substrates complex. Residues involved in AMPPNP or inosotide recognition are shown in sticks. Water molecules are represented in white and hydrogen bonds as dashed lines. (C) L3 atomic interactions as observed in the ternary products complex. This loop interacts with the side chain of Arg130 and Trp129 from $\alpha 6$, and residues in or near L2 at Gln237 and Arg241. L3 keeps together regions involved in inosotide recognition. (D) Substrates (Left) and products (Right) recognition and metal interactions in both ternary complexes. The substrates complex shows only one metal, Mg1, bound to the L5 residues Asp407 and Ser409, whereas the products complex shows two metals, Mg1 and Mg2, both bound to Asp407 and bridging inosotide and nucleotide ligands.

the C_{IP}-lobe (Fig. 2C). Interactions between L3 and G-loop through water molecules are also observed. We suspect that *At*-IP₅ 2-K undergoes a conformational change in presence or absence of the inosotide (for example, it only crystallizes in presence of inosotide), which could be attributed to changes in L3.

Different ligand recognition steps and metal binding. The complexes obtained do not display great structural differences (maximal rmsd of 0.479 Å), the inosotide interactions being essentially conserved. However, magnesium binding is only observed in ternary complexes with both substrates or products (Fig. 2D and Fig. S4A). In the complex with both substrates, the inosotide P3 and 2-OH interact with P γ , whereas Mg1 orients this P γ . In the complex with products, two metals (Mg1 and Mg2), are bridging the inosotide P2 to the nucleotide P α and P β . Detailed interactions of metal- and ligand- protein recognition are shown in Fig. 2D. The lead derivative crystal, obtained from a ternary ADP/IP₆ complex, showed two Pb²⁺ ions in similar positions to both Mg²⁺ ions, mimicking this cofactor.

There are no great structural changes among complexes upon nucleotide binding. A remarkable feature deduced from the binary complexes is that the inosotide apparently does not leave room for the nucleotide to enter (Fig. S4B), suggesting that either L3 moves from a putative closed to an opened conformation allowing nucleotide entrance, or the nucleotide must bind first into the active site. It remains to be determined whether IP₅ 2-K presents an ordered-ligand kinetics.

IP₅ 2-Ks from Plants Conserve a Zinc-Binding Motif. *At*-IP₅ 2-K presents a Zn motif in the C_{IP}-II ($\alpha 14$ – $\alpha 15$). The electron density

map shows a high density peak connected to His320, Cys330, Cys333, and His346 (Fig. S5A). This site has been attributed to a Zn²⁺ cation, although *At*-IP₅ 2-K was not previously identified as a zinc-binding enzyme. Inductively coupled plasma-optical emission spectroscopy (ICP-OES) analysis of *At*-IP₅ 2-K confirmed the presence of 1 mol of zinc per mol of protein. All IP₅ 2-Ks from plants conserve all the four Zn-binding residues (Fig. S5B). We have performed site-directed mutagenesis on the Zn-binding site, obtaining that C330S, C333S, and H346N single mutants are not stable under native conditions (Fig. S5C). In addition the incremental addition of EDTA to the wild-type protein reduced intrinsic tryptophan fluorescence by approximately 20% in an irreversible process (Fig. S5D). However incubation of the wild-type with EDTA has no effect on enzyme activity (Fig. S5E). Altogether, these observations suggest that Zn has a structural role in the proper folding of protein, but it seems not essential for enzymatic activity.

IP₅ 2-Ks Family and Mutagenesis Analysis. IP₅ 2-Ks from different sources have low sequence conservation (11%), *A. thaliana* and human enzymes being 40% homologous. However, small sequence motifs tend to be conserved in all IP₅ 2-Ks. We have performed a structural alignment among the 67 different IP₅ 2-Ks yielded by PFAM (<http://pfam.sanger.ac.uk/>), representatives being shown in Fig. 3A. The β -sheets are conserved in all species. The N-lobe helices are conserved, except $\alpha 2$ and $\alpha 3$ in region N-I, that seem specific for the plant enzymes, and conform a flexible region partially disordered in the crystal, with no apparent implications in substrate binding or catalysis. In the C-lobe, mammalian enzymes conserve all the α -helices described in *At*-IP₅ 2-K

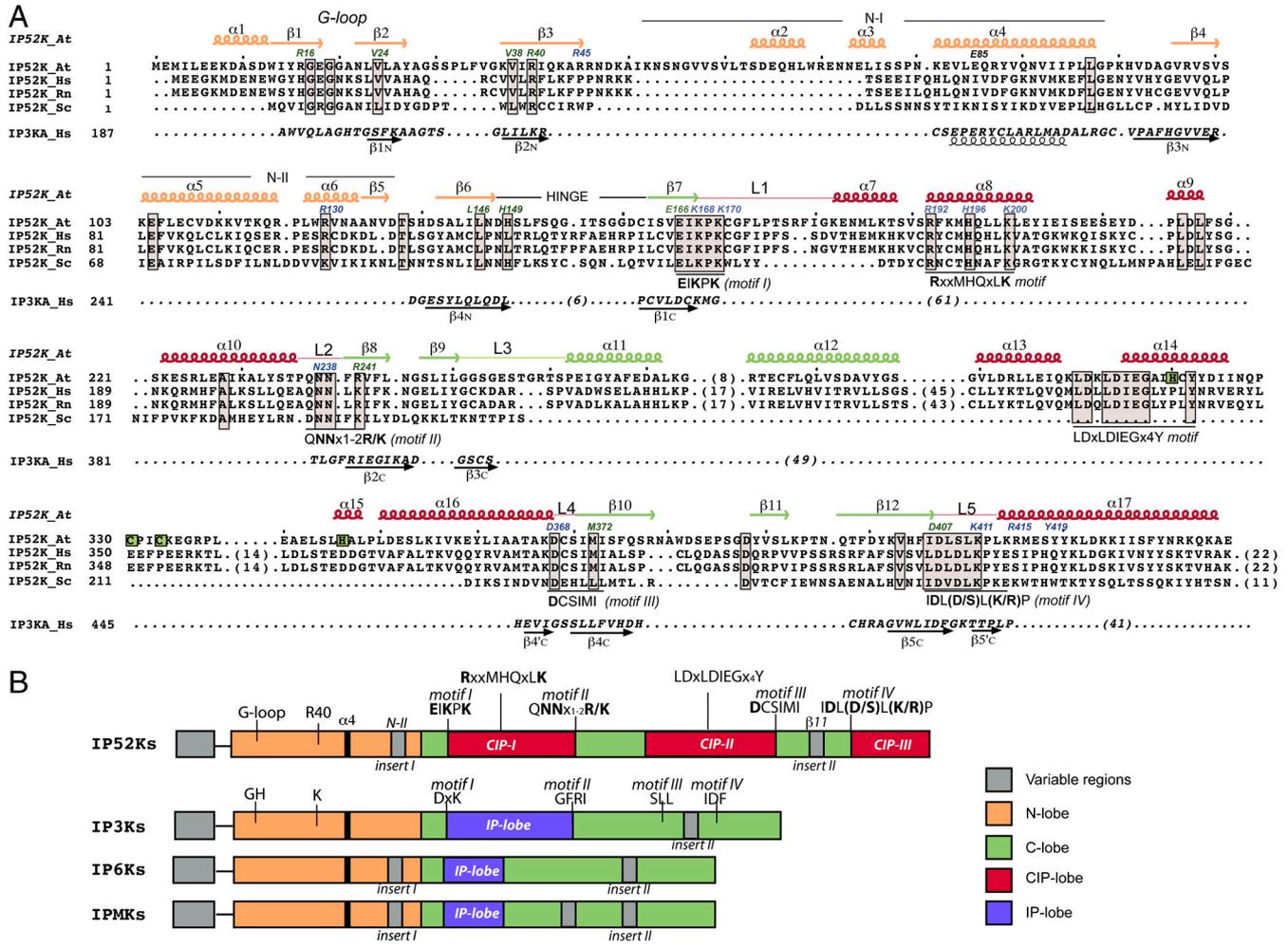


Fig. 3. The IP₅ 2-K family. (A) Structural alignment of At-IP₅ 2-K with IP₅ 2-Ks from other sources (Hs:human, Rn:rat, and Sc:yeast). Secondary elements are shown and named on top, and colored in the same code. Zn binding residues from At-IP₅ 2-K are inside green squares. Conserved residues are gray squared and motifs are shown below. Residues shown in dark green and blue are ATP and inosotide ligands, respectively. The IP₃ 3-K structural elements conserved with IP₅ 2-K and their sequence are shown below. (B) Schematic domain representation of IP₅ 2-Ks vs. the IPK family (IP₃ 3-K, IP₆ Ks and IPm Ks). The black rod represents alpha4, which is conserved in all IPKs, and is similar to PKs alphaC. The IP₅ 2-K G-loop, R40, and conserved motifs are marked on top, and the equivalent IPK regions are shown.

showing two additional main insertions at C_{IP-II}. The first 45 residues insertion is found between alpha12 and alpha13, with no secondary structure predicted. A second insertion, 14 residues long and between alpha14 and alpha15, is predicted to form an additional helix in mammals. This last insertion is located in the Zn-binding motif found in At-IP₅ 2-K, and interestingly the Zn ligands are not conserved in mammals or in other species. The yeast and fungi enzymes present a shorter C_{IP-lobe} lacking almost all region C_{IP-II}. Consequently, C_{IP-II} is the less conserved part among IP₅ 2-Ks and, in particular, the plant enzyme seems to have evolved to bind Zn that appears to have a structural role. Most of the residues implicated in ligand binding and activity are conserved (Fig. 3A) and are mainly enclosed in motifs I-IV, engaged in both substrates recognition, or in the ¹⁹²RxxMHQxLK motif, involved only in the inosotide recognition. There is a sixth conserved motif, ³¹⁰LDxLDIEGxIHxY not present in fungi, which seems to be essential for maintaining the C_{IP-lobe} properly folded. Because fungi have a shorter C_{IP-lobe}, this motif and its function are probably not required. Finally, L3 is not conserved among the IP₅ 2-K family, suggesting that L3 is characteristic of each IP₅ 2-K and could be essential in modulation of catalysis.

Two single mutations reported in yeast IP₅ 2-K (N10D and C162Y) yielded inactive enzymes (23). These residues are equivalent to Asn22 and Lys194, located in nucleotide- and inosotide-binding regions respectively. To confirm the role of other IP₅

2-K residues in catalysis, we have performed activity and kinetic analysis of selected mutants (Table 1). R40V activity is half the wild type, probably due to partial loss in phosphate coordination, which would also explain the increase in the Km for ATP observed in this mutant. As expected for catalytic residues (see later), mutants K168N and K168A yielded completely inactive enzymes and D407A has only residual activity. The almost complete inactivity of the R130I mutant illustrates the importance of P1 binding and the key role of the N-lobe in the enzymatic activity. Arg130 interacts with L3, and therefore it could also serve to stabilize the inosotide binding site conformation. Similarly, Lys170, which coordinates the inosotide P5 and P6, is essential for enzymatic function, whereas Gln238, also involved in P6 coordination, is not critical. The Km for ATP of N238A mutant is 2- to 3-fold above the wild type. We suspect that the nucleotide-binding pocket is affected by the interaction between L3 and motif II, therefore a change in Asn238 within this motif could have a direct effect in ATP affinity. Asp368, which coordinates 2-OH, yields a completely inactive enzyme when mutated even when assayed at high protein concentration. Asp368 is in turn coordinated to Lys411, a residue that interacts with P3 and P4 of the inosotide, and which is also critical for the enzymatic activity. In summary, all the mutagenesis assays support the conclusions deduced from the IP₅ 2-K structures presented here. Finally, it is worth noting that mutant

Table 1. Activity analysis of IP₅ 2-K mutants

IP ₅ 2-K mutant	Activity (% wild type)
Native	100.00
R40V*	56.06
E85A*	124.24
R130I	6.06 (9.09)
K168A	nd (2.27)
K168N	nd (4.54)
K170S	5.30 (3.78)
N238A*	85.60
D368A	nd (nd)
D407A	3.03 (7.57)
K411A	nd (3.53)

ND: No activity detected. Data in parentheses shows values obtained at 10-fold higher concentration of native.

*Native ATP Km = 38 μM and Vmax = 1080 pmol/min/μg. The Km value increases 2- to 3-fold for the active mutants (R40V, E85A, and N238A), with <20% reduction in Vmax.

E85A (equivalent to Glu91 in PKAs), unexpectedly, has no detrimental effect on activity (see Fig. S6)

IP₅ 2-K Is a Distant Member of the IPK Family. A search performed with the DALI server (http://ekhidna.biocenter.helsinki.fi/dali_server/) shows that IP₃ 3-K isoform C is the third closest structural homologue of IP₅ 2-K (rmsd = 5.4 Å for 153 residues), the C_{IP} region not showing any significant structural homologue. We conclude that IP₅ 2-K belongs to the IPKs family, albeit as the most distant member with no structural similarity in the inositide binding region.

IPKs display structural homology with PKs, conserving the N-lobe fold and the nucleotide recognition pattern. However, the C-lobe is divergent and presents some structural relationship with the ATP-grasp fold. All these IPKs features are conserved in IP₅ 2-Ks. Fig. 3A shows how most of the IP₅ 2-K N- and C-lobe β-sheets, α4, and, partially, the helices that pack over the sheet, are conserved in IP₃ 3-K, chosen as IPK representative. Particularly, the four motifs I–IV of IP₅ 2-Ks are equivalent to the conserved motifs of IPKs (Fig. 3B and Fig. S7), preserving some of the key residues and functions. Notably, the catalytic residues Lys168 from motif I and Asp407 from motif IV are conserved not only in all IPKs, but also in PKs. In spite of the homology in this region, there are some active conformation hallmarks of PKs (32) present in IPKs, that can not be found in IP₅ 2-Ks (Fig. S6). The conformational changes governing IP₅ 2-K activation might be specific for this family of enzymes.

As mentioned before, the inositide-binding region of IP₅ 2-Ks (C_{IP-lobe}) is very different from that of IPKs (IP-lobe). The C_{IP-lobe} (190 residues) is much larger than the IP-lobe (30 residues in IPmKs or 60 in IP₃ 3-Ks). The IP₃ 3-Ks IP-lobe is formed by a four-helix insertion that emerges from the C-lobe β1c and β2c strands (Fig. 3, Fig. S7, and Fig. S8A). The equivalent region in IP₅ 2-K is the region C_{IP-I} inserted into β7–β8, but presenting a very different fold. The regions C_{IP-II} and C_{IP-III} have no equivalent in IP₃ 3-K. Curiously, the superposition of IP₅ 2-K and IP₃ 3-K structures reveals that the IP₅ axial and the IP₃ equatorial C-OH acceptor bonds have a similar orientation relative to the nucleotide (Fig. S8). Consequently, the two inositide ring planes are perpendicular to each other. The OH acceptor, is bound to a lysine from motif I in both enzymes (Lys 264 and Lys 168, respectively); however, in IP₅ 2-K it is additionally bound to Asp368 from motif III. In spite of the different ring orientation, the IP₅ P3 falls in a similar spatial disposition to IP₃ P4, both phosphates interacting with a residue from motif IV (Lys411 and Lys419, respectively) (Fig. S8B). No more homology is found in the inositide recognition mode.

Finally, most IPKs have nonconserved N-terminal domains with functions distinct from the kinase activity (Fig. 3B). Only *Schizo-*

saccharomyces pombe IP₅ 2-K, has such an N-terminal domain (23). Besides, IPKs have variable insertions in both β-sheets, the most common ones being insert I and insert II. The IP₆ K2 insert I has been involved in interactions with HSP90 (33), and yeast IPmk insert II has been also implied in protein–protein interactions (34), though other studies discard these interactions (35). The IP₅ 2-K region N-II is located at “insert I” and participates in recognition of the inositide substrate, therefore we postulate that insert I could play also a role in the inositide binding within the IPK family. The IP₅ 2-K region β11 is located at insert II, but to date, no work implies this β-strand in protein–protein interactions.

Discussion

We have structural information that shows different snapshots of the IP₅ 2-K catalytic mechanism, that allowed us to identify the residues that coordinate P_γ and 2-OH before and after phosphoryl transference. These residues are Lys168, Asp407, Ser409, and Asp368, together with some G-loop residues. Lys168 neutralizes the negative charge developed in the transition state, whereas the G-loop, and Asp407 through Mg1, orientates P_γ for the subsequent nucleophilic attack of the inositol OH. The roles proposed for Lys168 and Asp407 (motifs I and IV) are conserved with the roles of equivalent residues in PKs and IPKs. The IP₅ 2-K specific residues Ser409 (Asp in other species) and Asp368 also seem to participate in the catalytic mechanism. Ser409 coordinates Mg1, possibly assisting in P_γ orientation and in addition, Ser409 constrains the conformational mobility of L5 and positions Lys411, essential for inositide coordination. Asp 368 recognizes the acceptor 2-OH in axial position of the inositide. It is also linked to Lys168 and Lys411 assisting in the active site conformation (Fig. S3). This residue is topologically equivalent to the Asp262 of the PKA catalytic loop, which selects the correct conformation of the serine substrate. Other IPKs and lipid kinases always bind –OHs in equatorial conformation, and do not have an equivalent acidic residue, supporting the idea that the major role of Asp368 is to select the axial 2-OH substrate. Furthermore, two specific features, the Asp368 2-OH recognition and the whole IP₅ binding pocket, potentially discriminates myo-inositides from other conformers. It remains to be determined whether Asp368 participates in catalysis. After phosphoryl transference, Mg1 reorganizes its coordination, and together with Mg2, bridges nucleotide and inositide. Other potential roles of the metals include assistance in catalysis and stabilization of the transition state or products. The metal roles are conserved with IPKs. Finally, taking into account the short distance covered by the P_γ to reach its fate (P_γ-2-OH distance = 3.2 Å; OPβ – P2 distance = 3.1 Å), we predict that the IP₅ 2-K mechanism involves an “in line” direct transference, as was proposed for IP₃ 3-K and protein kinases.

In conclusion, the structure of an IP₅ 2-K, the only inositide kinase that phosphorylates the axial position of an inositol compound, has been solved. The reported fold serves as a template for the whole IP₅ 2-K family, allowing the identification of the key residues. The determined IP₅ 2-K structure allows its classification as a distant member of the IPK family, with a very different and sophisticated large inositide-binding region (C_{IP-lobe}). IP₅ 2-K C_{IP-lobe} snugly embraces the phosphates of the myo-inositol (1,3,4,5,6)-pentakisphosphate substrate with multiple residues that contribute to substrate orientation. The structure explains the specificity of the catalytic reaction, in which a phosphate group is transferred just to the inositol 2-OH axial position. Finally, a typical insert region of IPKs N-lobe has been implicated in recognition of inositide, allowing us to propose a similar role in other family members. Because IP₅ 2-K represents a key point in the metabolism of highly phosphorylated inositols, detailed knowledge of IP₅ 2-K structure gives very valuable information for cell biology and rationale design of enzyme inhibitors. These inhibitors could facilitate study of the effect of IP₅ 2-K inhibition in mam-

mals, because disruption of this gene yields an embryo lethal phenotype in mice. Furthermore, this structure could be a valuable tool in plant physiology, helping in the design of low phytate crops. Finally, IP₅ 2-K substrate and product, IP₅ and IP₆ have already been reported to confer antitumor effects through PI3K/Akt pathway inhibition. Nevertheless, the most interesting facets of IP₅ 2-K activity manipulation perhaps remain to be investigated, as the product IP₆ participates in multiple signalling processes.

Methods

Protein Sample Expression and Purification. At-IP₅ 2-K was amplified from At5G 42810 cDNA (21) and cloned into the pKLSL₁ vector (36). For the enzyme assays, IP₅ 2-K was cloned into the pOPTG vector. Details of protein expression and purification are found in *SI Methods*.

Protein Crystallization. Crystallization procedures were set up as described by us (37). The best crystals were obtained at 18 °C and 22% PEG 3350, 100 mM Bis-Tris pH 5.9, and always in presence of the inositol substrate or product. Details of crystal complexes formation are given in *SI Methods*.

Data Collection, Phasing, and Model Refinement. Cryoprotection and heavy atom formation were done as previously reported (37). A standard MAD experiment and native datasets for several complexes were collected (Table S1). Programs used are: MOSFLM and the CCP4 package for data processing and reduction (38, 39); SHELX for heavy atom location (40); autoSHARP for protein phasing (41); Buccaneer for preliminary model building (42); REFMAC for

refinement (43); and O for manual building (44). For model completion we used our best dataset (IP₅ 2-K with products) in combination with the experimental model. A more detailed description is shown in *SI Methods*.

Enzyme Assays. Wild-type and mutant enzymes assays and kinetic analysis were performed as shown in *SI Methods*. Products were resolved by reverse-phase HPLC (45). Peak areas for ADP and ATP were integrated after reverse-phase HPLC (40). The effect of EDTA treatment of IP₅ 2K was determined in a coupled assay containing 10 mM-MgCl₂ (46).

Zinc Quantification and Fluorescence. Zinc quantification by ICP-OES were performed on native enzyme and fluorescence analysis were performed on native vs. EDTA treated enzyme as shown in *SI Methods*.

ACKNOWLEDGMENTS. We thank R. Williams and O. Perisic for critical reading of the manuscript, O. Perisic and J.M. Mancheño for providing us with pOPTG and pKLSL₁ vectors, P. Green for preliminary studies of enzyme activity, Chris Hamilton for assistance with assays, and Graham Chilvers for performing ICP-OES analysis. We thank the European Synchrotron Radiation Facility for providing time and assistance for data collection. B.G. was supported by Ramon y Cajal Fellowship RYC-2006-002701. J.I.B. was supported by Formación de Profesorado Universitario (FPU) Fellowship AP2008-00916 from the Ministerio de Educación. C.A.B. and P. Green were supported by Biotechnology and Biological Sciences Research Council of the UK Grant BB/C514090/1. The work was supported by Comunidad de Madrid-CSIC Grants CCG07-CSIC/GEN-2232 and CCG08-CSIC/GEN-3490, Ministerio de Ciencia e Innovación Grant BFU2008-02897/BMC, and Ministerio de Educación y Ciencia Grant RYC-2006-002701.

1. Michell RH (2008) Inositol derivatives: Evolution and functions. *Nat Rev Mol Cell Biol* 9:151–161.
2. Irvine RF, Schell MJ (2001) Back in the water: The return of the inositol phosphates. *Nat Rev Mol Cell Biol* 2:327–338.
3. York JD, Odom AR, Murphy R, Ives EB, Wente SR (1999) A phospholipase C-dependent inositol polyphosphate kinase pathway required for efficient messenger RNA export. *Science* 285:96–100.
4. Hanakahi LA, Bartlett-Jones M, Chappell C, Pappin D, West SC (2000) Binding of inositol phosphate to DNA-PK and stimulation of double-strand break repair. *Cell* 102:721–729.
5. Murphy AM, Otto B, Brearley CA, Carr JP, Hanke DE (2008) A role for inositol hexakisphosphate in the maintenance of basal resistance to plant pathogens. *Plant J* 56:638–652.
6. Agarwal R, Mumtaz H, Ali N (2009) Role of inositol polyphosphates in programmed cell death. *Mol Cell Biochem* 328:155–165.
7. Shen X, Xiao H, Ranallo R, Wu WH, Wu C (2003) Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. *Science* 299:112–114.
8. Steger DJ, Haswell ES, Miller AL, Wente SR, O'Shea EK (2003) Regulation of chromatin remodeling by inositol polyphosphates. *Science* 299:114–116.
9. Saiardi A, Caffrey JJ, Snyder SH, Shears SB (2000) The inositol hexakisphosphate kinase family. Catalytic flexibility and function in yeast vacuole biogenesis. *J Biol Chem* 275:24686–24692.
10. Mulugu S, et al. (2007) A conserved family of enzymes that phosphorylate inositol hexakisphosphate. *Science* 316:106–109.
11. Shukla VK, et al. (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* 459:437–441.
12. Raboy V (2003) myo-Inositol-1,2,3,4,5,6-hexakisphosphate. *Phytochemistry* 64:1033–1043.
13. Raboy V (2001) Seeds for a better future: "Low phytate" grains help to overcome malnutrition and reduce pollution. *Trends Plant Sci* 6:458–462.
14. Raboy V (2007) The ABCs of low-phytate crops. *Nat Biotechnol* 25:874–875.
15. Shamsuddin AM (2008) Demoning phytate. *Nat Biotechnol* 26:496–497.
16. Bozsik A, Kokeny S, Olah E (2007) Molecular mechanisms for the antitumor activity of inositol hexakisphosphate (IP₆). *Cancer Genomics Proteomics* 4:43–52.
17. Graf E, Empson KL, Eaton JW (1987) Phytic acid. A natural antioxidant. *J Biol Chem* 262:11647–11650.
18. Maffucci T, et al. (2005) Inhibition of the phosphatidylinositol 3-kinase/Akt pathway by inositol pentakisphosphate results in antiangiogenic and antitumor effects. *Cancer Res* 65:8339–8349.
19. Stephens LR, et al. (1991) myo-inositol pentakisphosphates. Structure, biological occurrence and phosphorylation to myo-inositol hexakisphosphate. *Biochem J* 275:485–499.
20. Phillippy BQ, Ullah AH, Ehrlich KC (1994) Purification and some properties of inositol 1,3,4,5,6-pentakisphosphate 2-kinase from immature soybean seeds. *J Biol Chem* 269:28393–28399.
21. Sweetman D, Johnson S, Caddick SE, Hanke DE, Brearley CA (2006) Characterization of an *Arabidopsis* inositol 1,3,4,5,6-pentakisphosphate 2-kinase (AtIPK1). *Biochem J* 394:95–103.
22. Verbsky JW, Wilson MP, Kisseleva MV, Majerus PW, Wente SR (2002) The synthesis of inositol hexakisphosphate. Characterization of human inositol 1,3,4,5,6-pentakisphosphate 2-kinase. *J Biol Chem* 277:31857–31862.
23. Ives EB, Nichols J, Wente SR, York JD (2000) Biochemical and functional characterization of inositol 1,3,4,5,6-pentakisphosphate 2-kinases. *J Biol Chem* 275:36575–36583.
24. Sun Y, et al. (2007) Inositol 1,3,4,5,6-pentakisphosphate 2-kinase from maize: molecular and biochemical characterization. *Plant Physiol* 144:1278–1291.
25. Verbsky J, Lavine K, Majerus PW (2005) Disruption of the mouse inositol 1,3,4,5,6-pentakisphosphate 2-kinase gene, associated lethality, and tissue distribution of 2-kinase expression. *Proc Natl Acad Sci USA* 102:8448–8453.
26. Sarmah B, Latimer AJ, Appel B, Wente SR (2005) Inositol polyphosphates regulate zebrafish left-right asymmetry. *Dev Cell* 9:133–145.
27. Gonzalez B, et al. (2004) Structure of a human inositol 1,4,5-trisphosphate 3-kinase: substrate binding reveals why it is not a phosphoinositide 3-kinase. *Mol Cell* 15:689–701.
28. Miller GJ, Hurley JH (2004) Crystal structure of the catalytic core of inositol 1,4,5-trisphosphate 3-kinase. *Mol Cell* 15:703–711.
29. Miller GJ, Wilson MP, Majerus PW, Hurley JH (2005) Specificity determinants in inositol polyphosphate synthesis: crystal structure of inositol 1,3,4-trisphosphate 5/6-kinase. *Mol Cell* 18:201–212.
30. Cheek S, Ginalski K, Zhang H, Grishin NV (2005) A comprehensive update of the sequence and structure classification of kinases. *BMC Struct Biol* 5:6.
31. Stevenson-Paulik J, Bastidas RJ, Chiou ST, Frye RA, York JD (2005) Generation of phytate-free seeds in *Arabidopsis* through disruption of inositol polyphosphate kinases. *Proc Natl Acad Sci USA* 102:12612–12617.
32. Nolen B, Taylor S, Ghosh G (2004) Regulation of protein kinases; controlling activity through activation segment conformation. *Mol Cell* 15:661–675.
33. Chakraborty A, et al. (2008) HSP90 regulates cell survival via inositol hexakisphosphate kinase-2. *Proc Natl Acad Sci USA* 105:1134–1139.
34. Messenguy F, Dubois E, Boonchird C (1991) Determination of the DNA-binding sequences of ARGR proteins to arginine anabolic and catabolic promoters. *Mol Cell Biol* 11:2852–2863.
35. Caddick SE, Harrison CJ, Stavridou I, Johnson S, Brearley CA (2007) A lysine accumulation phenotype of Sc1pk2Delta mutant yeast is rescued by *Solanum tuberosum* inositol phosphate multikinase. *Biochem J* 403:381–389.
36. CSIC (2009) PCT Patent Application WO/2009/121994.
37. Baños-Sanz JI, Villate M, Sanz-Aparicio J, Brearley CA, Gonzalez B (2009) Crystallization and preliminary x-ray diffraction analysis of inositol 1,3,4,5,6-pentakisphosphate kinase from *Arabidopsis thaliana*. *Acta Crystallogr F* 66:102–106.
38. Leslie AGW (1992) *Joint CCP4 + ESF-EAMCB Newslett Protein Crystallogr*, (Daresbury Laboratory, Warrington, UK), 26.
39. CCP4 (Collaborative Computational Project Number 4) (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D* 50:760–763.
40. Sheldrick GM (2008) A short history of SHELX. *Acta Crystallogr A* 64:112–122.
41. Vonrhein C, Blanc E, Roversi P, Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215–230.
42. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D* 62:1002–1011.
43. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
44. Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47:110–119.
45. Caddick SE, et al. (2008) A *Solanum tuberosum* inositol phosphate kinase (StIPK1) displaying inositol phosphate-inositol phosphate and inositol phosphate-ADP phosphotransferase activities. *FEBS Lett* 582:1731–1737.
46. Rowan AS, et al. (2009) Nucleoside triphosphate mimicry: A sugar triazolyl nucleoside as an ATP-competitive inhibitor of *B. anthracis* pantothenate kinase. *Org Biomol Chem* 7:4029–4036.