



The crystal structure of mammalian inositol 1,3,4,5,6-pentakisphosphate 2-kinase reveals a new zinc-binding site and key features for protein function

Received for publication, February 7, 2017, and in revised form, April 25, 2017. Published, Papers in Press, April 27, 2017, DOI 10.1074/jbc.M117.780395

Elsa Franco-Echevarría[‡], Julia Sanz-Aparicio[‡], Charles A. Brearley^{§1}, Juana M. González-Rubio[‡], and Beatriz González^{‡2}

From the [‡]Departamento de Cristalografía y Biología Estructural, Instituto de Química-Física "Rocasolano," Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain and the [§]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, United Kingdom

Edited by F. Peter Guengerich

Inositol 1,3,4,5,6-pentakisphosphate 2-kinases (IP₅ 2-Ks) are part of a family of enzymes in charge of synthesizing inositol hexakisphosphate (IP₆) in eukaryotic cells. This protein and its product IP₆ present many roles in cells, participating in mRNA export, embryonic development, and apoptosis. We reported previously that the full-length IP₅ 2-K from *Arabidopsis thaliana* is a zinc metallo-enzyme, including two separated lobes (the N- and C-lobes). We have also shown conformational changes in IP₅ 2-K and have identified the residues involved in substrate recognition and catalysis. However, the specific features of mammalian IP₅ 2-Ks remain unknown. To this end, we report here the first structure for a murine IP₅ 2-K in complex with ATP/IP₅ or IP₆. Our structural findings indicated that the general folding in N- and C-lobes is conserved with *A. thaliana* IP₅ 2-K. A helical scaffold in the C-lobe constitutes the inositol phosphate-binding site, which, along with the participation of the N-lobe, endows high specificity to this protein. However, we also noted large structural differences between the orthologues from these two eukaryotic kingdoms. These differences include a novel zinc-binding site and regions unique to the mammalian IP₅ 2-K, as an unexpected basic patch on the protein surface. In conclusion, our findings have uncovered distinct features of a mammalian IP₅ 2-K and set the stage for investigations into protein-protein or protein-RNA interactions important for IP₅ 2-K function and activity.

Inositol 1,3,4,5,6-pentakisphosphate 2-kinase (IP₅ 2-K)³ is a key enzyme of higher inositol phosphate (inositide, IP) metab-

olism. IP₅ 2-K is present in yeast to mammals and catalyzes the synthesis of phytic acid (inositol hexakisphosphate or IP₆) from IP₅ and ATP (1). A variety of roles have been proposed for this enzyme and its product IP₆ (2) in DNA repair (3), mRNA editing, export and degradation (4, 5), vesicle trafficking (6) and protein ubiquitylation (7). At a molecular level, IP₆ acts as a cofactor for proteins with DNA-dependent protein kinase activity in non-homologous end joining (8) and with GLE1 in mRNA export (4). IP₆ also can act as a folding factor as in the case of adenosine deaminase that participates in editing of mRNA and tRNA (5). In addition, IP₆ is the precursor of inositol pyrophosphates, essential for cellular energy homeostasis, signal transduction control, and apoptosis (9, 10). Mice embryos with IP₅ 2-K deletion do not survive more than a few weeks (11). More recently, a role in ribosomal rRNA synthesis independent of IP₅ 2-K catalytic function has been proposed for the human enzyme (12), which has been shown to colocalize with mRNA either in the nucleus or cytoplasm (13). All these findings increase the potential of this enzyme as an attractive target.

IP₅ 2-K belongs to the inositol polyphosphate (IPK) structural family that include enzymes capable of phosphorylating hydroxyls at different positions of the inositol ring starting from inositol 1,4,5-trisphosphate (IP₃), a well known second messenger responsible for calcium mobilization (14). These phosphorylation events occur in combination with another family of inositol kinases that adopts an "ATP grasp-like" fold (15, 16). Both families together cover a great range of phosphorylation reactions on the six -OH positions of the *myo*-inositol ring and even on those already phosphorylated (15, 17, 18). Many of these enzymes present redundant abilities acting on similar substrates or they bind an inositide in different orientations that is thus phosphorylated in different positions (19). In contrast, IP₅ 2-K is the unique IPK whose physiological role is the phosphorylation of the axial 2-OH position of *myo*-inositol, the other five hydroxyls being in equatorial positions. It is also a

This work was supported in part by Grants BFU2011-24982 and BFU2014-53762-P from the Spanish Ministry of Economy and Competitiveness (to E. F.-E.). The authors declare that they have no conflicts of interest with the contents of this article.

✂ Author's Choice—Final version free via Creative Commons CC-BY license.

This article contains supplemental Tables S1–S3 and Figs. S1 and S2.

The atomic coordinates and structure factors (codes 5MW8, 5MWL, and 5MWM) have been deposited in the Protein Data Bank (<http://www.pdb.org/>).

¹ Supported by Biotechnology and Biological Sciences Research Council (United Kingdom) Grants BB/N002024/1 and BB/M022978/1.

² To whom correspondence should be addressed. Tel.: 34-91-5619400; Fax: 34-91-564-24-31; E-mail: xbeatriz@iqfr.csic.es.

³ The abbreviations used are: IP₅ 2-K, inositol 1,3,4,5,6-pentakisphosphate 2-kinase; IP, inositol phosphate; IP₅, inositol 1,3,4,5,6-pentakisphosphate; IP₆, inositol hexakisphosphate; IPK, inositol polyphosphate kinase; LSL,

Laetiporus sulphureus lectin; PEG, polyethylene glycol; TEV, tobacco etch virus; UBF, upstream binding factor; IP₃, inositol 1,4,5-trisphosphate; ICP-OES, inductively coupled plasma-optical emission spectroscopy; mIP, mouse IP; hIP, human IP; r.m.s.d., root mean square deviation; CL, connecting loop; Ins(1,2,4,5,6)P₅, 1,2,4,5,6-inositol pentakisphosphate; Ins(1,3,4,5,6)P₅, inositol 1,3,4,5,6-pentakisphosphate.

very specific enzyme and, together with IP₃ 3-K, phosphorylates just one position of inositol (20, 21).

IPK enzymes are classified as a structural subgroup of the protein kinase (PK) family, because they conserve a few features including a fold in two separated lobes (N- and C-lobe) and similar nucleotide recognition mode (22). In addition, the core of the N-lobe and a few involved in catalysis are conserved. In 2004, the first structure of an IPK was described from IP₃ 3-K (23). Since then, the structure of at least one member of each IPK subfamily has been reported (23–26). In summary, the IPK family presents a specific and characteristic fold in the C-lobe different from PKs, having a β -sheet core with helical insertions showing a great range of sizes depending on the IPK class. These helical regions are involved in substrate binding and have been named by us as IP-lobe or CIP-lobe (19). Thus, promiscuous IPKs, such as the IP multikinases, have a single helix inserted showing an open active site able to cope with various substrates and products, whereas more specific enzymes as IP₃ 3-K and IP₅ 2-K show larger helical scaffolds. In fact, IP₅ 2-K shows the most elaborated helical region.

The full-length structure of IP₅ 2-K from *Arabidopsis thaliana* (*AtIP₅ 2-K*) is the only one known for this subfamily (25). It shows the features described above and shows that it is a zinc metallo-enzyme, for which a structural role has been proposed (25). We subsequently captured different IP₅ 2-K conformations by X-ray crystallography (27) displaying open, half-closed, or closed conformations, as the nucleotide and/or the inositide are absent or present in the active site. Extensive work by others and us has identified the residues involved in substrate recognition and catalysis and has depicted the participation of the N-lobe in achievement of a productive conformation (25, 27–30). The structure of *AtIP₅ 2-K* was an important advance for this field, but nevertheless, the specific features of mammal IP₅ 2-K remain unknown. The IP₅ 2-K family shows moderate sequence conservation across the species. In particular, mammalian enzymes show different insertions and do not conserve the zinc site found in the plant enzyme, making it difficult to obtain a good sequence alignment. From its sequence motifs, putative zinc-binding residues have been proposed in human IP₅ 2-K (13). However, if mammal IP₅ 2-K is a zinc metallo-enzyme, the putative role for this metal awaits further studies. Because of the significance of IP₅ 2-K to proper cell functioning and the many roles of its product IP₆, it is important to understand the molecular basis that underlies this enzyme function. In this work, we have determined the mouse IP₅ 2-K structure in the presence of inositide showing that although this enzyme conserves features with the plant enzyme, it differs significantly in many aspects. Our results define the specific features of mammal IP₅ 2-Ks. In addition, we present here valuable information that could help in understanding IP₅ 2-K functions beyond its catalytic activity as its role in ribosomal RNA synthesis (12).

Results

Structure of IP₅ 2-K mammalian isoform

We have solved the structure of *Mus musculus* IP₅ 2-K (*mIP₅ 2-K*) at 2.4 Å resolution (Table 1) from a truncated form lacking the 21 C-terminal residues (Δ C-*mIP₅ 2-K*). Noticeably, *mIP₅*

2-K crystals were not obtained in the absence of the inositide. The structure for *mIP₅ 2-K* in the presence of one or both ligands, forming binary complexes (IP₆) or ternary complexes (IP₅ + ATP), is presented (supplemental Fig. S1). In addition, we present two different crystal forms, including one or two molecules in the asymmetric unit, respectively, the first showing much better resolution (2.4 versus 3.2 Å). As mouse and human IP₅ 2-K isoforms share 91% of sequence identity, we propose the structure of the mouse enzyme as a template for the mammalian IP₅ 2-Ks.

Mouse IP₅ 2-K folds in two lobes, N- and C-terminal lobes, connected by a hinge, thereby conserving the general fold scheme of PKs and IPKs, and in a similar way, both lobes coordinate the nucleotide between them (Fig. 1A). The N-lobe core forms a β -sheet formed by five antiparallel β -strands (β 1– β 5) showing two helical segments. The first helical segment (N-I) harbors α 1, equivalent to the helix α C characterized in all protein kinases, whereas the second one (N-II) is a specific insertion different in every IPK subfamily. A role of this region for substrate binding in the IP₅ 2-K subfamily has been previously reported by others (29) and by us (25, 27). Regarding the C-lobe, it also presents a β -sheet core formed by five antiparallel β -strands (β 6– β 10). Three helical segments are inserted in the β -sheet core. These segments altogether form a large helical ensemble named the CIP lobe in the structure of *AtIP₅ 2-K* (25), and each of them is consequently named as CIP_I, CIP_{II}, and CIP_{III}. The CIP lobe represents more than half of the protein and is specific to IP₅ 2-K enzymes. It creates a scaffold that builds up most of the inositide substrate-binding site. The five loops (CL1–CL5) joining the CIP region to the C-lobe β -sheet core are essential because they play a key role in substrate binding and catalysis (Figs. 1B and 2).

mIP₅ 2-K active site and substrate recognition

A general view of IP₅ 2-K substrate recognition is shown in Fig. 2A and detailed in supplemental Table S1. The adenine is strongly recognized through polar and hydrophobic interactions with both protein lobes and the hinge connecting them (Fig. 2B). In particular, it forms polar interactions with His-14 and the backbones of Pro-116 and Leu-118. The ribose OHs interact with the C-lobe residues Glu-136 and Arg-209. The triphosphate moiety is tightly bound to the N-lobe of the enzyme through polar interactions and to the C-lobe through two magnesium ions. In particular, phosphate interaction with residue Arg-33, with a flexible loop (G-loop, residues Gly-15–Ser-20) and with an acidic residue (Asp-437) through the magnesium ions, is conserved throughout the PK superfamily and is essential for nucleotide binding and kinase activity.

Regarding the inositide substrate, IP₅ is tightly bound to the enzyme through its five phosphates (Fig. 2C). A total of 14 residues coordinate the phosphate groups, two of them through water molecules and eight of them being lysine or arginine. P1 and P3 are coordinated by both lobes, whereas P4 and P5 are coordinated exclusively by the C-lobe. Residues from the N-lobe involved in P1 and P3 binding come from the segment N-II (Arg-100) and the G-loop (Asn-18 and Lys-19) mentioned above. Residues from the C-lobe involved in coordination of the five phosphates come from the CIP lobe and its CLs. Fig. 2C and supplemental Table S1 show all the polar interactions produced

Table 1
Crystallographic data statistics and refinement

	ΔC -mIP ₅ 2-K	ΔC -mIP ₅ 2-K	ΔC -mIP ₅ 2-K
Ligand modeled	+IP ₅ + ATP	+IP ₆	+IP ₅ + ATP
Crystallization pH	6.25	5.50	5.50
Data collection and processing			
Space group	P2 ₁	P2 ₁	P2 ₁
Unit cell <i>a</i> , <i>b</i> , <i>c</i> (Å)	64.62, 140.76, 68.66	60.55, 71.64, 61.82	60.16, 71.50, 61.20
Unit cell α , β , γ (°)	90.0, 106.5, 90.0	90.0, 111.7, 90.0	90.0, 111.4, 90.0
Temperature (K)	100	100	100
Radiation source	Synchrotron	Synchrotron	Synchrotron
Wavelength (Å)	0.979490	0.979260	0.979260
Resolution range (Å)	49.86–3.20 (3.42–3.20)	44.82–2.54 (2.65–2.54)	71.50–2.40 (2.49–2.40)
No. of observed reflections	134,008 (24,350)	106,382 (13,008)	123,314 (13,608)
No. of unique reflections	19,318 (3482)	16,270 (1970)	18,528 (1961)
Multiplicity	6.9 (7.0)	6.5 (6.6)	6.7 (6.9)
Data completeness (%)	99.3 (99.0)	99.6 (99.7)	97.3 (99.8)
Matthews coefficient (Å ³ Da ⁻¹)	2.32	2.32	2.30
No. of molecules in a.u.	2	1	1
Wilson <i>B</i> -factor (Å ²)	83.26	54.65	45.72
Mean <i>I</i> / σ (<i>I</i>)	14.7 (3.3)	12.5 (3.1)	14.9 (3.5)
<i>R</i> _{merge} (%) ^a	10.4 (63.5)	7.6 (58.8)	7.3 (52.8)
<i>R</i> _{pim} (%) ^b	4.3 (25.8)	3.2 (24.6)	3.1 (21.5)
<i>CC</i> _{1/2}	0.99 (0.89)	0.99 (0.99)	0.99 (0.97)
Refinement			
Resolution range (Å)	70.37–3.20	57.45–2.60	56.99–2.40
<i>R</i> _{work} / <i>R</i> _{free} (%)	22.51/24.53	25.34/28.95	24.10/27.12
No. of atoms/Bav (Å ²)	6647/103.43	3348/69.36	3451/52.81
Protein	6575/103.15	3285/69.65	3342/53.178
Ligand	63/153.38	36/64.72	63/44.77
Zinc	4/79.47	1/46.91	1/49.87
Magnesium	2/102.48		2/38.56
Water molecules	3/51.26	26/39.60	39/35.34
Ramachandran plot (%)			
Favored/outliers	88.8/0.0	92.8/0.3	91.7/0.3
r.m.s.d.			
Bonds/angles (Å/°)	0.007/1.18	0.008/1.29	0.006/1.22
Protein Data Bank codes	5MWL	5MWM	5MW8
Missing residues in Protein Data Bank	A: 1–4/99–102/221–225/296–311/ 413–426/465–468	1–9/40–41/122–128/221–228/ 244–248/298–310/411–419/465–468	1–4/41–42/122–128/221–228/ 244–245/296–310/411–419/465–468
	B: 1–7/36–43/97–105/221–226/ 244–251/277–278/298–311/ 412–426/464–468		

^a $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $I_i(hkl)$ is the measurement of reflection hkl and $\langle I(hkl) \rangle$ is the weighted mean of all measurements.

^b $R_{\text{pim}} = \frac{\sum_{hkl} (1/(N-1)) |2 \sum_i I_i(hkl) - (I(hkl))|}{\sum_{hkl} \sum_i I_i(hkl)}$, where N is the redundancy for the hkl reflection.

^c $R_{\text{work}}/R_{\text{free}} = \frac{\sum_{hkl} |F_o - F_c|}{\sum_{hkl} |F_o|}$, where F_c is the calculated and F_o is the observed structure factor amplitude of reflection hkl for the working/free (5%) set, respectively.

with the inositol phosphates (P1: Asn-18, Arg-100, Lys-138, Lys-173, and Asn-206; P3: Asn-18, Lys-19, and Lys-441; P4: Gln-449; P5: with Lys-140, Arg-160, and His-164 through water molecules; and P6: Lys-138, Lys-140, Lys-168, Asn-206, and water mediated with Asn-207). The tight and extensive recognition explains the high specificity of this enzyme. Human IP₅ 2-K displays a 0.43 μM K_m value for IP₅ (31) in good agreement with the value obtained for mIP₅ 2-K by us (0.29 μM) (Fig. 3A) and in contrast with the low tens of micromolar values reported for the AtIP₅ 2-K (22 μM) (32). A possible explanation for the different K_m values found in the plant enzyme might be related to some differences found in enzymes from both kingdoms either in the inositol recognition mode or in the constraints introduced by the zinc-binding site (see below).

Six residues from the CLs form the interface of substrate recognition, generating a net of interactions, including the magnesium ions (Fig. 2D). Residues Lys-138 and Asp-400 make direct interaction with P γ and/or the nucleophile 2-OH. The distance between the 2-OH and P γ oxygen is 3.2 Å suggesting an in-line transference mechanism probably in agreement with an associative mechanism. P γ is oriented through a magnesium ion (Mg1) coordinated to Asp-437 and to Asp-439 in a second sphere. Asp-437 also coordinates a second magnesium ion (Mg2) together with Ser-402 through a water molecule.

Comparison of the ternary and binary complexes (IP₅ 2-K/IP₅/ATP versus IP₅ 2-K/IP₆) shows that there is no significant structural variation among them (r.m.s.d. is 0.377 Å for 403 Ca atoms). The IP₆ shows similar interactions, with P2 remaining at the substrate interface regions described above. A BLAST search using mIP₅ 2-K sequence and limited to mammals (taxid: 40674) shows that all residues coordinating the nucleotide and inositol are absolutely conserved, with just a couple of residues showing a conservative change in some species.

Unexpectedly, the formation of the ternary complex (mIP₅ 2-K + IP₅ + ATP) was achieved by protein incubation with IP₆ and ADP (supplemental Fig. S1). Therefore, we checked that our crystallized mIP₅ 2-K samples are able to catalyze both the forward (Fig. 3B) and reverse (Fig. 3C) reactions in solution. In agreement, the plant enzyme is highly reversible with an equilibrium constant in the forward “kinase” direction of ~14 (20). By ion-pair reverse-phase HPLC, we were also able to confirm the production of ATP from IP₆ and ADP (data not shown).

Mammal IP₅ 2-Ks share a zinc-binding site with a novel structure

mIP₅ 2-K presents two zinc ions in its structure, one in the CIP-lobe (Zn1) and the other close to the hinge region (Zn2) (Figs. 1 and 4). In fact, Zn1 is present in all the crystals obtained,

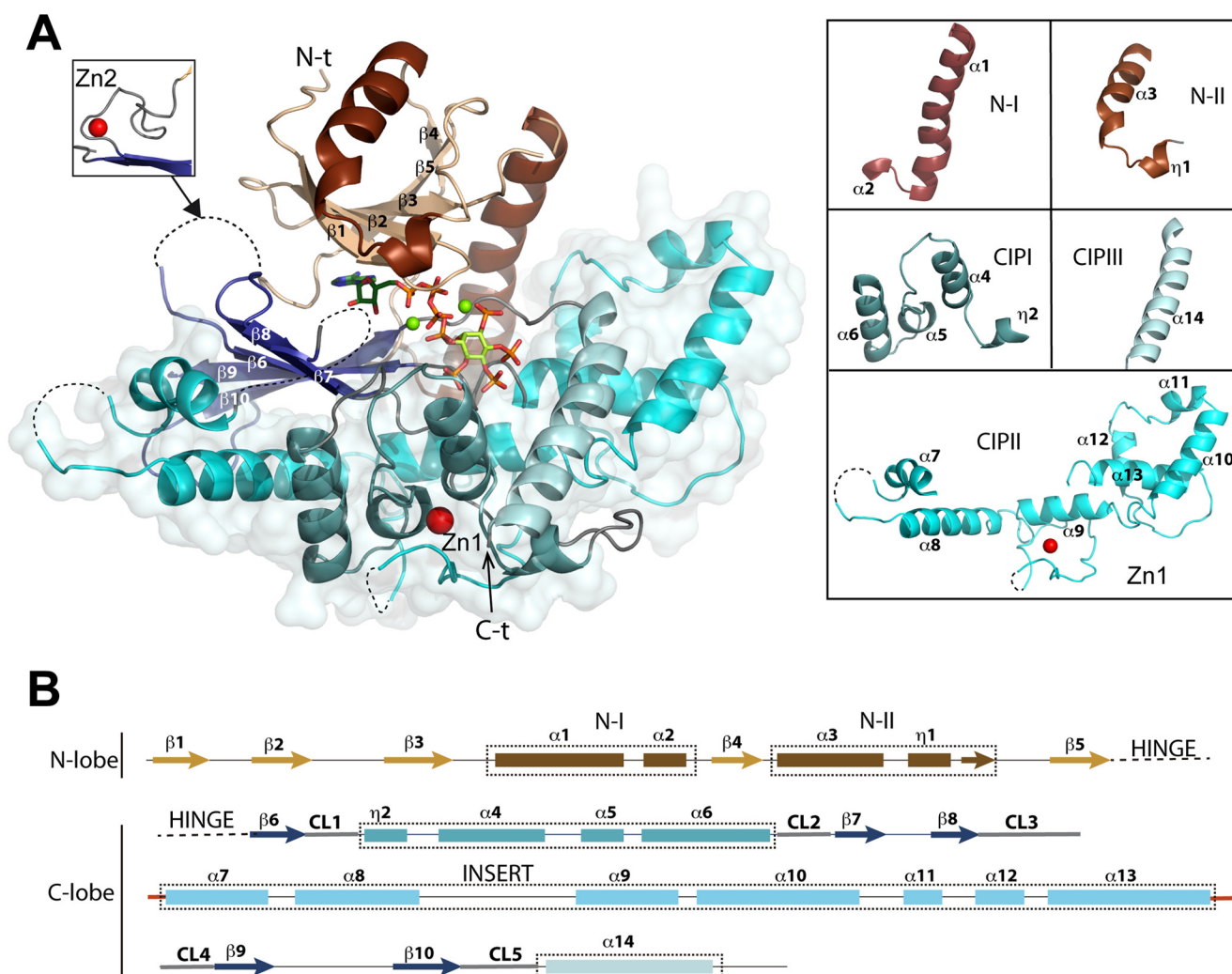


Figure 1. Structure of mIP_5 2-K. *A*, schematic representation of the structure of mIP_5 2-K. The N-lobe and C-lobe are shown in orange-brown and blue colors, respectively. ATP and IP_5 are shown as green sticks, highlighting the oxygen, nitrogen, and phosphorus atoms in red, blue, and orange, respectively. The zinc and magnesium ions are shown as red and green spheres, respectively. The left inset shows a second zinc site found in one of the complexes. The dashed lines show disordered regions. On the right, the five helical segments found in mIP_5 2-K are detailed. *B*, arrows indicate β -strands and rectangles the α -helices. The connecting loops between C-lobe β -sheet and CIP-lobe are highlighted in gray. Conserved sequence motifs within the IPK family are concentrated in the CL β -strands connections (CL1, ¹³⁶EIKPK; CL2, ²⁰⁶QNNXRFX; CL3, variable in sequence and length; CL4, ⁴⁰⁰DCSIMI; and CL5, ⁴³⁶LDLCLK).

whereas Zn2 was only detected in the low resolution ΔC - mIP_5 2-K crystals in which the two molecules in the asymmetric unit interact through the hinge region probably fixing a conformation captured by crystallography (supplemental Fig. S2).

The Zn1 site is formed by residues from two CIP lobe elements: Cys-159, Cys-162, and Cys-181 from CIP-I and Cys-291 from CIP-II (Fig. 4A). This site presents the typical zinc geometry and coordination, although it has no homologues in the structural databases using the DALI server (33). The two first cysteines from CIP-I are located in a helix, separated by two residues, and could resemble a partial zinc-finger; however, the fourth ligand breaks any resemblance because it comes from a position very distant in sequence. In fact, Cys-291 comes from a very long loop inserted into two helices that cross over the back of the CIP-lobe (Fig. 5). We have selected Cys-181 and Cys-291 as candidates for the mutagenic study (Fig. 4C). However, there was no expression of the C181S mIP_5 2-K mutant in the conditions reported herein for the wild-type protein. This suggests that the Zn1 site formation could be essential for proper protein

folding. By contrast, C291S mIP_5 2-K mutant expression levels are in the same order as the wild type, being only 2-fold decreased. Despite this, the impact of this mutation on enzyme activity is very high, because the mutated sample retains less than 10% of enzyme activity (Fig. 4C). In this line, it is worth mentioning that the two first cysteines of this zinc site are in helix $\alpha 4$, an element that provides four residues for the inositide substrate binding (Fig. 2A). We also consider that this fact could have some effect in obtaining a lower K_m value for the substrate in mIP_5 2-K.

Regarding Zn2, it is created by two residues from the hinge (His-125 and His-129) and one residue from the C-lobe (Cys-410) (Fig. 4B). Mutation of residues His-129 and Cys-410 to serine produces samples with a moderate decrease in enzymatic activity revealing that Zn2 is not critical for protein function (Fig. 4C). A possible explanation for the greater decrease found in the H129S mutant could be its key location in the hinge, an element important for enzyme flexibility.

We have subsequently evaluated the metal content in mIP_5 2-K samples by inductively coupled plasma-optical emission

Structure of a mammal IP_5 2-K

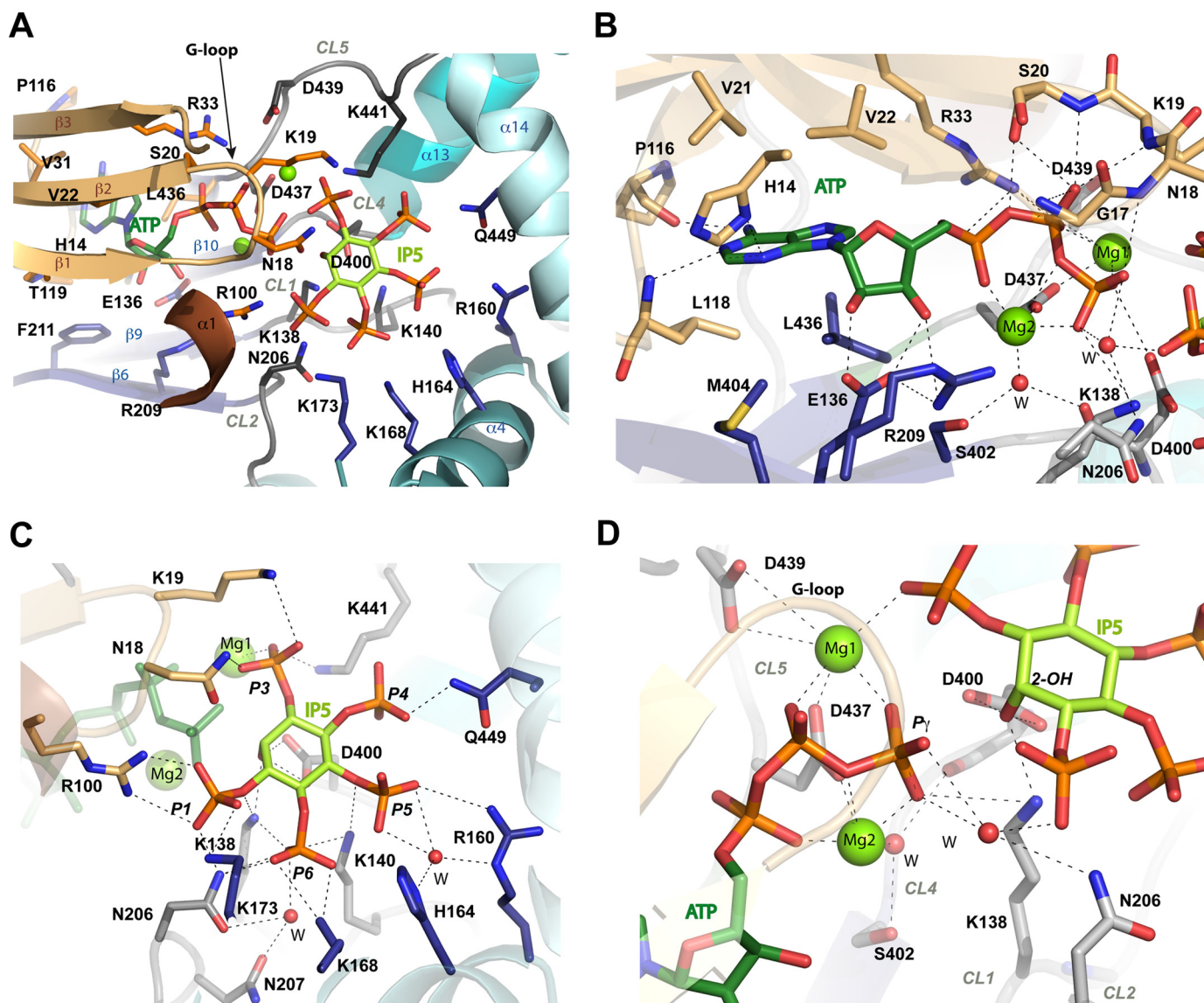


Figure 2. Substrate recognition by mIP_5 2-K. *A*, zoom of Fig. 1A showing the substrates ATP and IP_5 and the residues involved in their recognition as sticks in light orange (N-lobe), blue (C-lobe), and gray (CLs in C-lobe). *B*, zoom showing the nucleotide site. Water molecules are shown as red spheres. *C*, zoom showing the inositide site. *D*, zoom showing inositide and nucleotide interaction.

spectroscopy (ICP-OES) (Fig. 4D). We confirmed that mIP_5 2-K is a zinc metalloenzyme (supplemental Table S2). Unexpectedly, wild-type samples purified as LSLt-tagged protein exhibit a zinc/protein molar ratio of 0.6:1. An insufficient supply of zinc could also explain the difficulties encountered in the crystallization process of this enzyme due to structural inhomogeneity. However, attempts to add zinc from different salts to the crystallization drops failed because the protein precipitated. We found a zinc/protein molar ratio of 0.8:1, 0.4:1, and 0.7:1 for H129S, C291S, and C410S, mutants, respectively. As expected, mutation on residue Cys-291 (Zn1 site) yields an enzyme with less zinc content than the wild type (Fig. 4D and supplemental Table S2). By contrast, mutation of Zn2-coordinating residues (His-129 and Cys-410) yields samples with no reduction in the zinc content (Fig. 4D and supplemental Table S2). Therefore, we further analyzed the features of the C291S mutant by circular dichroism (CD) (Fig. 4E). We found that both WT (wild type) and C291S mIP_5 2-K samples display a similar far-UV CD

spectra suggesting that they share similar secondary structural elements. However, thermal denaturation followed by CD revealed that the mutated sample exerts an apparent T_m (32 °C) drastically reduced compared with that of the WT sample (43 °C).

In conclusion, our results suggest that Zn1 is necessary for protein folding and stability. Furthermore, the reduced capacity of the C291S mutant to bind zinc correlates with a high reduction in protein activity (Fig. 4, C and D). In agreement, IP_5 2-Ks from mammals show conservation only in the Zn1 site, with its four cysteine ligands being fully conserved. Therefore, Zn1 is key in all mammal IP_5 2-K enzymes, whereas Zn2 could be an artifact of crystallization.

Mammal and plant IP_5 2-Ks show three large structural divergences

Until now, the other IP_5 2-Ks with known structure is that from *A. thaliana* (25). A $C\alpha$ superposition of mIP_5 2-K onto *AtIP_5* 2-K (Protein Data Bank code 2xan) overlays 327 resi-

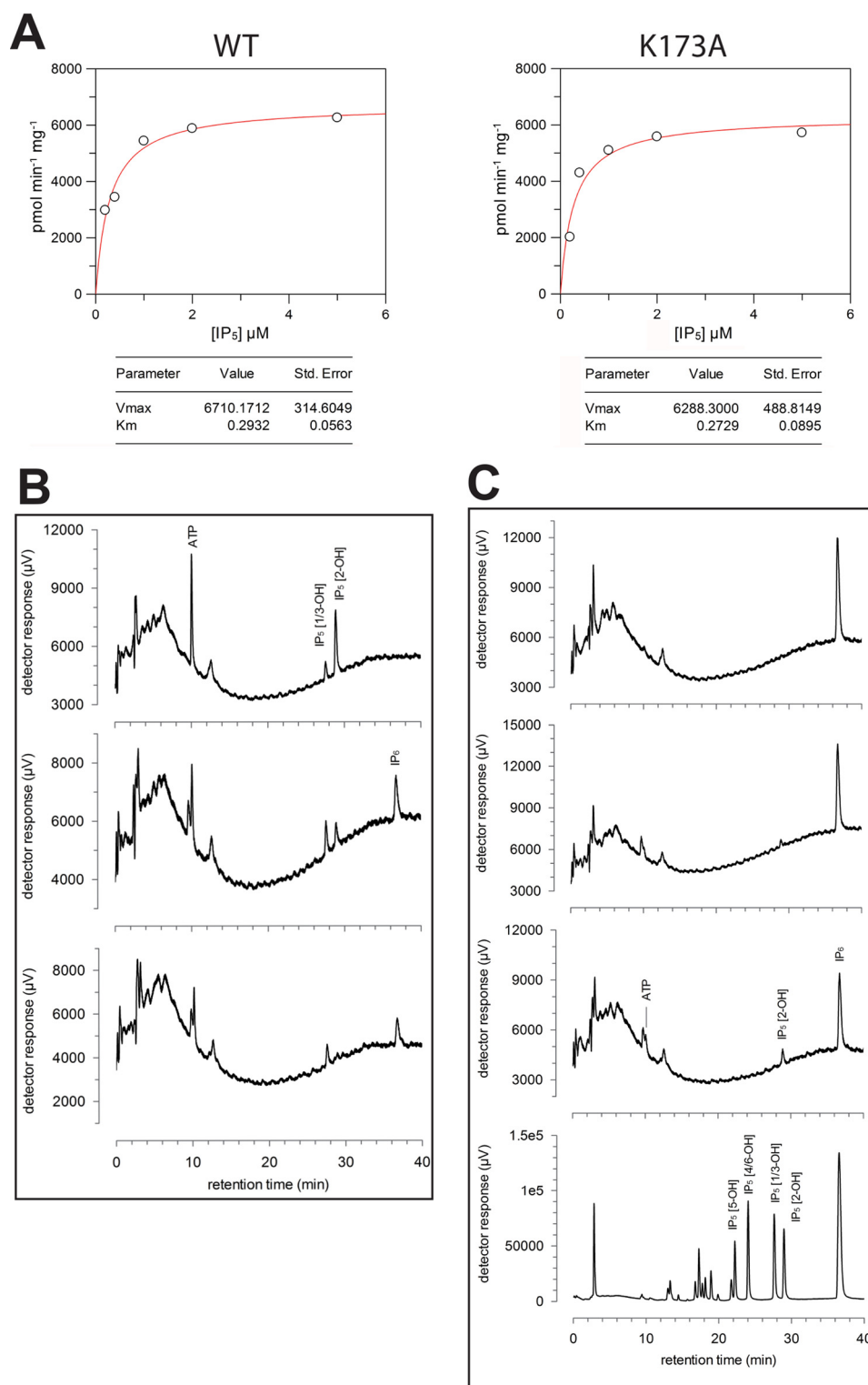


Figure 3. Enzyme assays and reaction catalyzed by mIP_5 2-K samples. *A*, kinetic parameters of LSLt- mIP_5 2-K WT and K173A variant. *B*, kinase reaction catalyzed by mIP_5 2-K. Products of enzyme assays resolved by ion-exchange HPLC: substrate, $Ins(1,3,4,5,6)P_5$ (IP_5) (upper panel); and reaction products after 60 min (middle panel) and 260 min (lower panel) of incubation of IP_5 and ATP with enzyme. *C*, reversibility of reaction catalyzed by mIP_5 2-K. Products of enzyme assays resolved by ion-exchange HPLC: substrate, IP_6 (upper panel), reaction products after 110 min (upper middle panel), reaction products after 110 min (upper middle panel) and 310 min (lower middle panel) of incubation of IP_6 and ADP with enzyme. The lower panel shows IP standards obtained by acid hydrolysis of IP_6 .

dues (out of 468 in mIP_5 2-K) with an r.m.s.d. of 1.2 Å. A good sequence alignment between both enzymes has remained elusive because their sequence homology is not very high (24% identity and 38% similarity), and they present different

insertions. A structural alignment of both IP_5 2-K isoforms is shown in Fig. 5. Although the topology of the N- and C-lobe cores and substrate-binding region is quite conserved, both enzymes present multiple dissimilarities.

Structure of a mammal IP_5 2-K

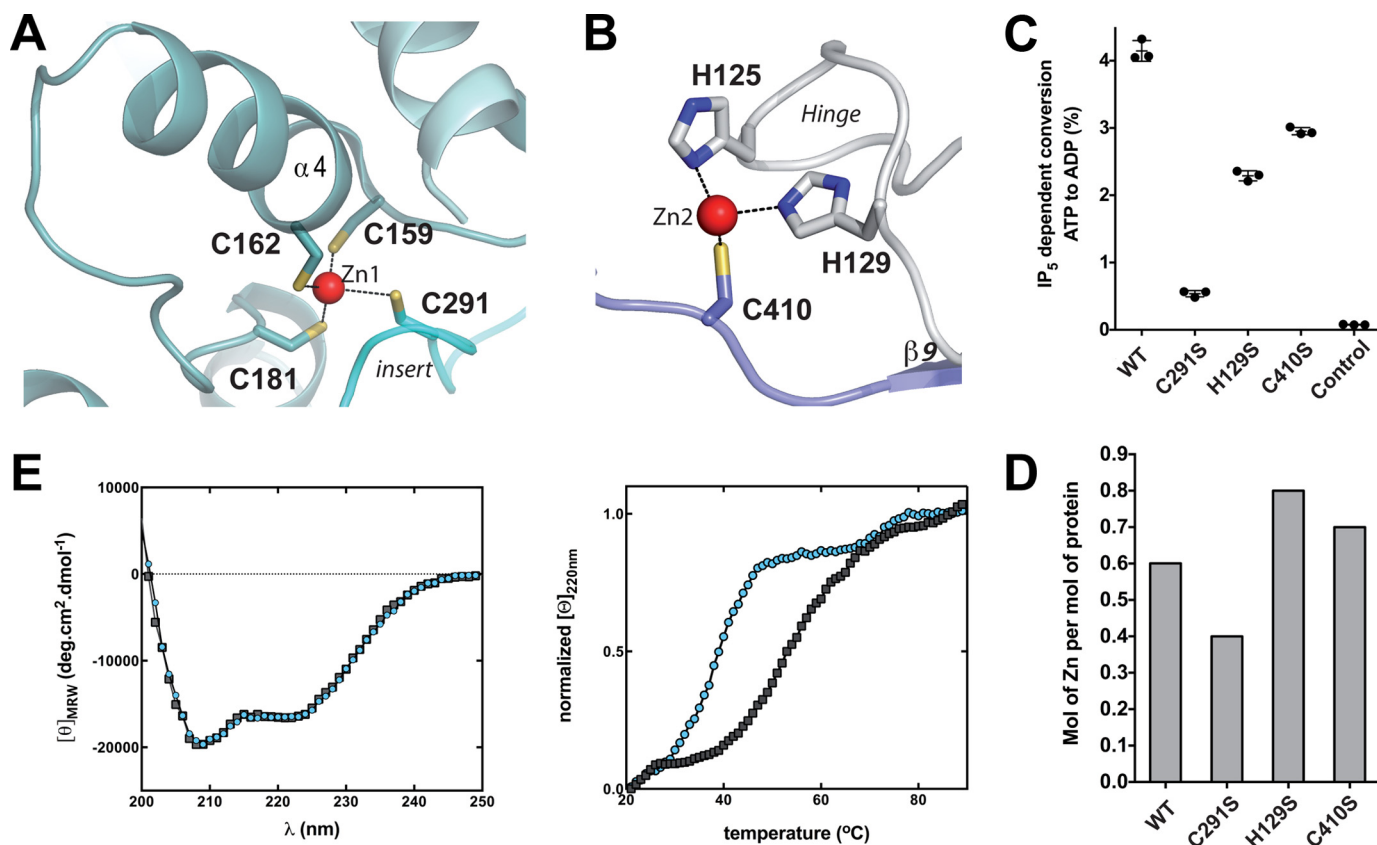


Figure 4. Zinc-binding sites located in mIP_5 2-K complexes. *A*, schematic representation of mIP_5 2-K Zn1 site, and *B*, of Zn2 site. *C*, plot of IP_5 -dependent conversion from ATP to ADP of mIP_5 2-K WT and mutants of the zinc-binding residues. Error bars show the standard deviation. *D*, zinc content of mutated mIP_5 2-K samples relative to WT mIP_5 2-K. *E*, far-UV CD spectra (*left*) and thermal denaturation followed by CD (*right*) of WT (black squares) and C291S (blue circles) ΔC - mIP_5 2-K.

The most predominant divergences are found within three regions, which show large differences in their sequence, length, and topology (Figs. 5 and 6, *A–D*). The first main difference (D1) is located in N-I region, *AtIP₅* 2-K having an insertion not present in the mammal enzyme (Fig. 6*B*). A possible function of this segment will be discussed below. The second clear difference (D2) is concentrated within the CIP-II region (Fig. 6*C*). Mouse IP_5 2-K presents a long and flexible loop (Ser-272–Gly-316, 45 residues) that crosses back from the whole CIP (Figs. 5 and 6*C*). Interestingly, this loop is the one that provides a cysteine residue (Cys-291) that completes the Zn1 site present in mIP_5 2-K (Fig. 4). In addition, this loop seems to stabilize the protein, because it packs with several regions of the enzyme, including both lobes, but mainly the CIP lobe. In particular, it makes hydrogen bonds through four residues (Leu-283, Arg-289, Glu-292, and Ser-294) and shows strong hydrophobic interactions through seven residues (Leu-281, Leu-283, Pro-288, Ala-293, Pro-295, Leu-313, and Pro-314) (Fig. 6*E*). These residues are fully or highly conserved along the mammal isoforms. Only residues Leu-283 and Arg-289 show a great variation, and both are involved in hydrogen bonds through their main chain atoms with other parts of the protein (Fig. 6*E*). Finally, a third difference (D3) is shown in a region also located in the CIP-II (Fig. 6*D*). Precisely, *AtIP₅* 2-K presents a zinc site located in this region (25) that is shorter than and completely different from that found in mIP_5 2-K (Fig. 6, *C* and *D*).

Noticeably, different regions of the two isoforms seem to share roles. On the one hand, D1 and D3, from *AtIP₅* 2-K and

mIP₅ 2-K, respectively (Fig. 6, *B* and *D*), are both stabilizing essential zones of the enzyme, as are $\alpha 1$ (Asn-54–Phe-66, analogous to protein kinases αC) and other CIP residues. Some relevant interactions of these regions are shown in Fig. 6, *F* and *G*. Among them, it is worth mentioning the central role of Tyr-363, Phe-367, and Tyr-368 in mIP_5 2-K (Fig. 6*F*), residues absolutely conserved in the mammal IP_5 2-K isoforms. *AtIP₅* 2-K presents residues with roles similar to Tyr-363 (Leu-75) and Phe-367 (Trp-69), which interact with equivalent regions of the protein (Fig. 6*G*). We have selected residue Tyr-363 for mutation, because it seems key in the network of interactions as proposed above. Y363A mutation halves enzymatic activity (Fig. 6*H*). Unexpectedly, this decrease in activity is moderate. An explanation is that it corresponds to a single mutation among a multiple net of interactions and in a residue far from the active site. On the other hand, mIP_5 2-K D2 and *AtIP₅* 2-K D3 (Fig. 6, *C* and *D*) are both involved in generation of the zinc sites, which are not conserved either in sequence, location, or structure between both enzymes. However, both sites seem to have a structural role, although we cannot discard any other additional function. Strikingly, the insertions found in mIP_5 2-K (D2 and D3) interact with each other (Fig. 6*E*). Thus, Tyr-368 in mIP_5 2-K and the following leucine residues (Leu-372 and Leu-374) interact with Leu-281 and Leu-283 located in D2 (Fig. 6*E*). A double mutation in this region (L281A/L283A) reduces slightly the activity in contrast with the >90% decreases caused by the other mentioned mutation on a zinc-binding residue (C291S)

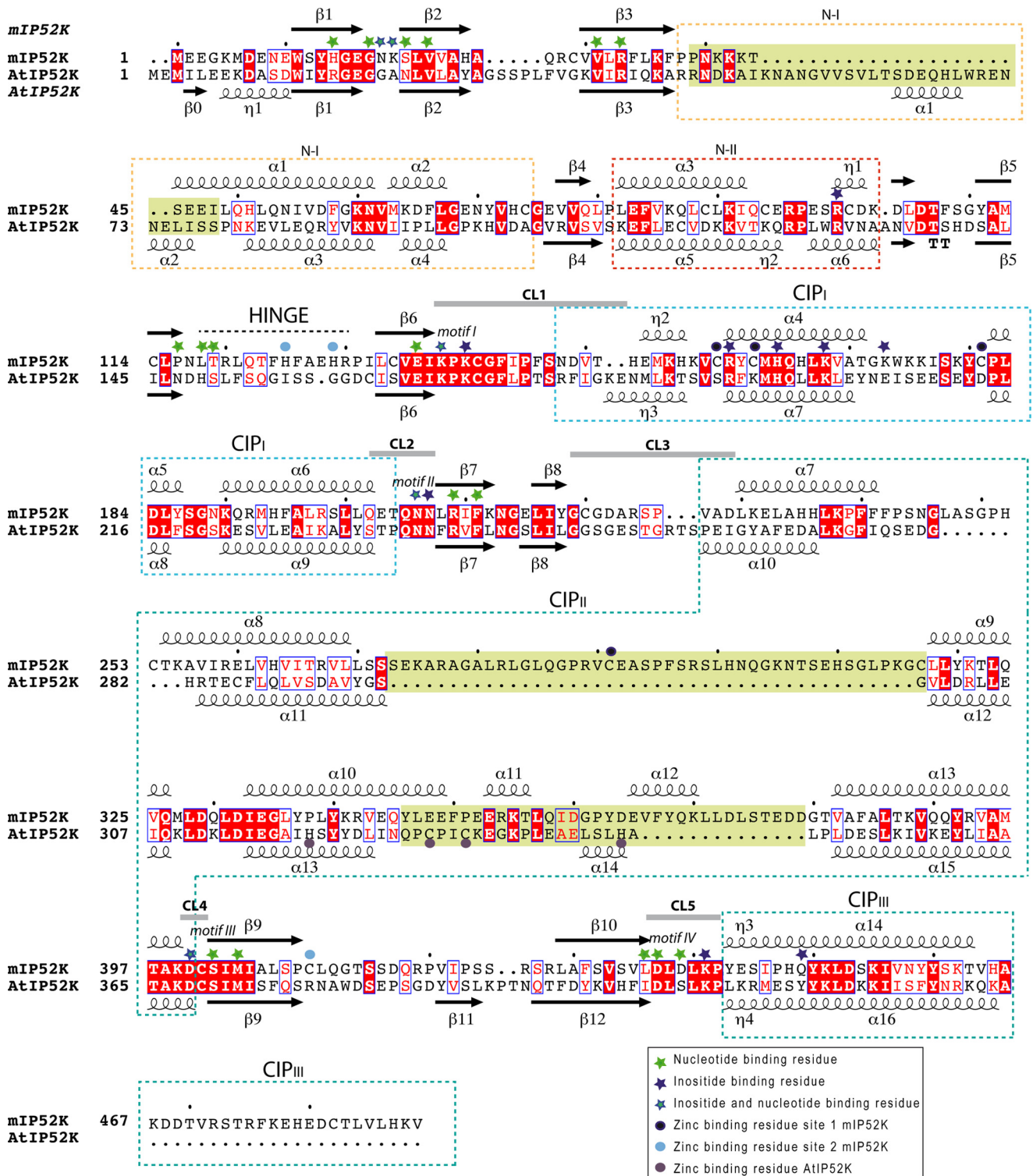


Figure 5. Structural alignment of mIP_5 2-K versus $AtIP_5$ 2-K. The secondary structure is shown at top (mammalian enzyme) and bottom (plant enzyme). Identical regions are shaded red, and similar regions are shown by red letters. Light-green shaded squares show regions with a divergent structure in both enzymes. Helical regions inserted into the β -sheet cores are marked with dashed lined boxes.

within this segment (Fig. 6H). All the mentioned leucine residues show very high conservation, except Leu-283, as mentioned previously. A mutation in this residue would keep the capacity of making hydrogen bonds through its main chain (Fig. 6E) and the slight structural destabilization introduced has no impact in the protein-active site.

Novel findings in mammal IP_5 2-K substrates binding and catalysis

Most residues involved in substrate binding and catalysis are conserved between mammal and plant IP_5 2-Ks (Fig. 5). Nevertheless, we can observe some differences in the inositide P1 and

Structure of a mammal IP_5 2-K

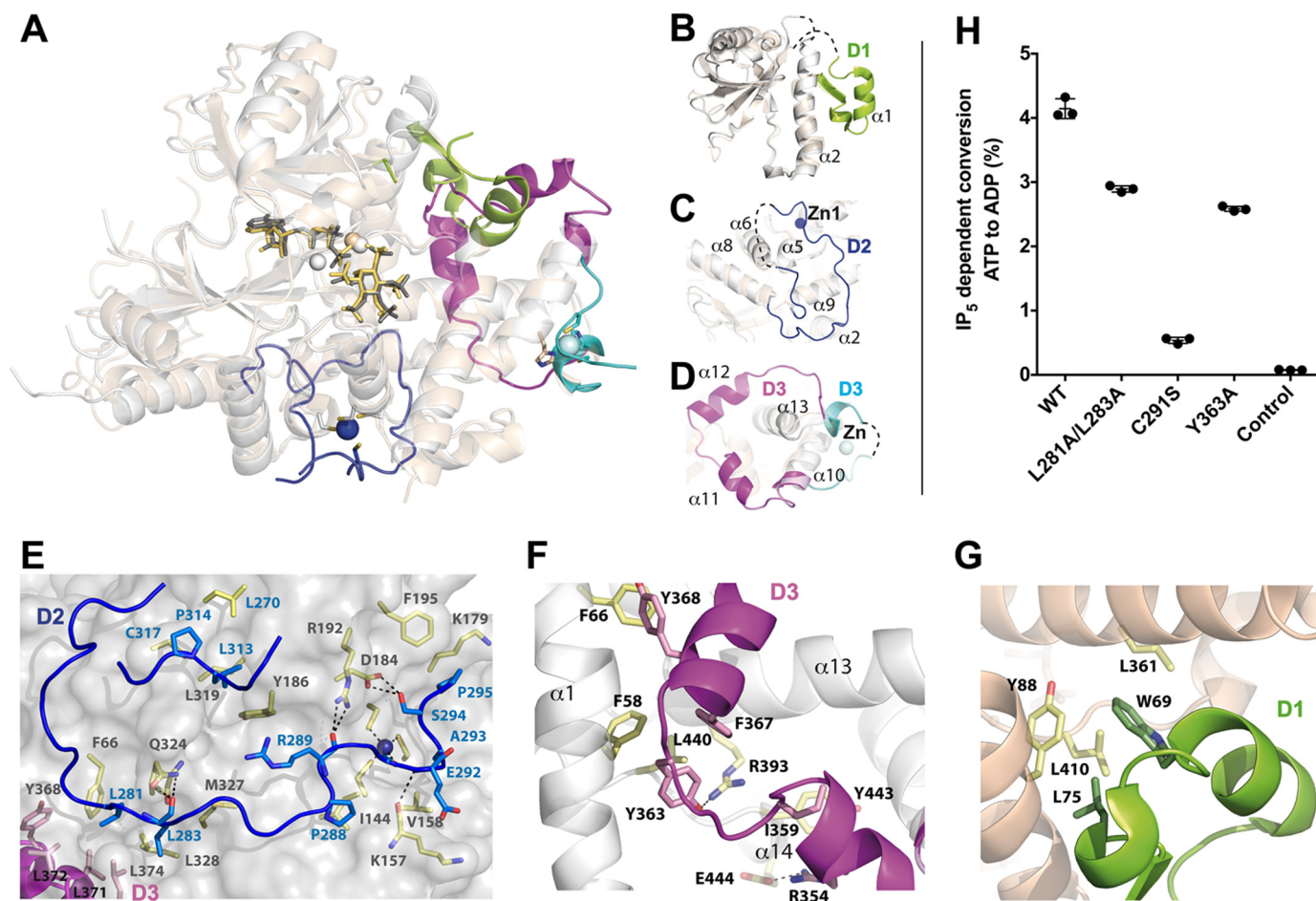


Figure 6. Novel regions of mIP_5 2-K and structural comparison with $AtIP_5$ 2-K. A, schematic representation of mIP_5 2-K (white) and $AtIP_5$ 2-K (wheat) superposed structures. Zinc ions are shown as dark blue (mIP_5 2-K) and cyan ($AtIP_5$ 2-K) spheres. Magnesium ions are shown as spheres in similar color to the isoform to which they belong. The main differences between both samples (D1, D2, and D3) are highlighted in different colors. B, representation of N-lobes from both isoforms highlighting D1 (green) insertion in $AtIP_5$ 2-K. C, piece of CIP1I lobe showing D2 in mIP_5 2-K (blue). D, view showing D3 (magenta in mIP_5 2-K and cyan in $AtIP_5$ 2-K). E, intramolecular interactions produced by D2 in mIP_5 2-K. Blue, pink, and yellow sticks show residues from D2, D3, and the rest of the protein, respectively. F, intramolecular interactions produced by D3 in mIP_5 2-K. G, intramolecular interactions produced by D1 in $AtIP_5$ 2-K. Green sticks show residues from D1. H, plot of IP_5 -dependent conversion of ATP to ADP of WT mIP_5 2-K and several mutants of residues located in the different regions. Error bars show the standard deviation.

P3 coordination. As shown, inositide P1 interacts extensively with Arg-100 of mIP_5 2-K (Fig. 2C). The role of the Arg-100 equivalent in $AtIP_5$ 2-K (Arg-130) has been largely argued, and an implication in substrate binding and triggering of a productive protein conformation has been proposed (27, 29, 30). Unfortunately, a construct prepared for the R100A mIP_5 2-K mutant did not show expression in the soluble fraction. However, we can conclude that this residue is structurally similar to Arg-130 in $AtIP_5$ 2-K. In mIP_5 2-K, additional interactions with P1 are produced through the side chain of Lys-173, a residue non-conserved with the plant IP_5 2-Ks but absolutely conserved in mammal enzymes, whereas conservative substitutions can be observed in other vertebrates (Fig. 7). However, mutation of Lys-173 produces an enzyme as active as the wild type (Fig. 7A) showing very similar kinetic parameters (Fig. 3A), suggesting that this interaction is dispensable for substrate binding and it probably might have other implications, as will be commented on later.

Inositide P3 coordinates with the main chain of the G-loop in both enzymes, but different additional interactions are made within the two enzymes. In mIP_5 2-K, P3 also interacts with the

side chains of Asn-18 and Lys-19 G-loop residues, whereas the plant enzyme provides two arginine residues from a different region (Arg-45 and Arg-415) to complete this binding. The double mutation on the G-loop residues Asn-18/Lys-19 (Fig. 7A) has a notable impact in the enzymatic activity, supporting a main role for these G-loop residues absolutely conserved in mammals (Fig. 7C).

We also can observe particular features in catalytic residues in the connecting loops (Fig. 1B). We have prepared mutations on relevant residues from these CLs (K138A, D437A, and D439A) (Fig. 2D), obtaining mutated enzymes with very low activities (Fig. 7B). The role of equivalent residues to Lys-138 and Asp-437 has been largely studied along the PK and IPK families. These residues are responsible for neutralizing the negative charge developed in the transition state and orienting the nucleotide P_{γ} through magnesium ions, respectively. We observe in mIP_5 2-K that Asp-439 displays a strategic position, helping with magnesium coordination and forming a bridge with Arg-33, a residue involved in the coordination of nucleotide phosphates. Therefore, it seems to provide a proper conformation for essential parts of the enzyme (Fig. 7B). No

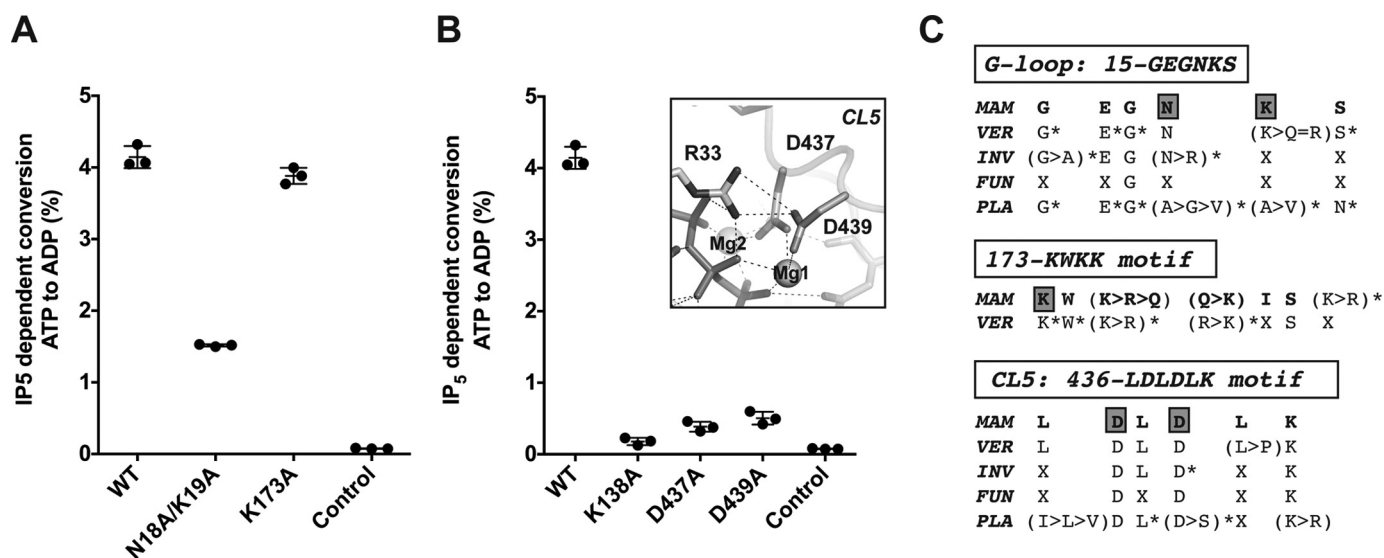


Figure 7. Analysis of relevant residues for *mIP*₅ 2-K function. *A*, plot of IP₅-dependent conversion of ATP to ADP by *mIP*₅ 2-K WT and several mutants to determine the effect of selected inositide-binding residues. *Error bars* show the standard deviation. *B*, same as *A* to check the effect of catalytic residues. The *inset* illustrates the salt bridge between Asp-439 and Arg-33. *C*, IP₅ 2-K sequence conservation across the kingdoms in regions containing the mutated residues (Asn-18 and Lys-19; Arg-173 and Asp-439). Letter *X* and *asterisks* denote residues non-conserved or mostly full conserved, respectively. Abbreviations: *MAM*, mammals; *VER*, vertebrates other than mammals; *INV*, invertebrates; *FUN*, fungi; *PLA*, plants.

mutagenesis data have been reported on the Asp-439 equivalent residues in other enzymes, because a homologous residue in this position is not present (neither in the IPK or PK families). Asp-439 shows a notable conservation in the whole IP₅ 2-K family across the species (Fig. 7C), showing only changes by a serine residue in some plant IP₅ 2-Ks, which in turn also coordinates the magnesium atom (25). We show here that the D439A mutation produces nearly inactive enzyme. Finally, the connecting loop CL3 does not show significant conservation between *mIP*₅ 2-K and *AtIP*₅ 2-K, either in length or sequence. This loop is responsible in *AtIP*₅ 2-K for interactions between the N- and C-lobe that might partially regulate the catalysis through opening and closing the active site (27). In addition, this loop makes interactions with two residues directly involved in inositide binding, one of them (Arg-130) shown to be essential for protein activity (27). We think that differences in this loop may also account for the differences observed in the *K_m* value from each protein. In the *mIP*₅ 2-K structure, this loop is disordered, and therefore interaction with the N-lobe has not been determined (Fig. 1). However, the flexibility found in the *mIP*₅ 2-K CL3 loop is consistent with a dynamic role and the previous proposed functions in catalysis regulation.

Prominent basic patch on *mIP*₅ 2-K surface

The structure of *mIP*₅ 2-K shows a very notable basic patch on its surface, mainly concentrated down the active-site face of the enzyme (Fig. 8A). The enzyme regions that contribute to this patch are $\alpha 6$ and $\alpha 8$ in CIP-I and the large insert found in the CIP-II region. Interestingly, only a few of these basic residues are present or conserved in *AtIP*₅ 2-K (Fig. 8B), which apart from the basic pocket for the inositide binding does not show any significant accumulation of arginine or lysine residues on its surface. By contrast, most of these residues are conserved across the mammalian isoforms. This striking feature could be correlated with mammalian IP₅ 2-K localization and/or other

possible functions that this enzyme could present in cells. Particularly outstanding are two basic segments, one formed by residues Lys-175, Lys-176, and Lys-179, and the other containing the residues in the new insert (D2) found in mammals and its preceding helix (Lys-255, Arg-259, Arg-267, Lys-275, Lys-289, and Lys-315). The first segment overlaps in part with the motif ¹⁷³KW(K/R)(K/Q), which is highly conserved in mammals. The second segment encloses one of the main differences (D2) found with respect to plant enzymes. Based on the present finding, we propose that this insert, apart from supporting the zinc binding and stabilizing some protein regions, could be delineating a particular protein surface that could serve as an interface for other partners essential for IP₅ 2-K function.

Discussion

Here, we described the first structure of a mammalian IP₅ 2-K, a key enzyme in inositol metabolism with multiple impacts in diverse cellular events. Structural knowledge of this enzyme is essential to fully understand its function, although the fact that mammalian IP₅ 2-K shows extremely low bacterial expression and a very low tendency to produce suitable crystals has precluded it until now. The structure of *mIP*₅ 2-K presented here shows large structural differences with the *AtIP*₅ 2-K concentrated in the helical regions. Both enzymes are zinc metalloenzymes, the zinc sites showing a different location and structural features. Whereas the plant zinc site is exclusive for its kingdom, the site found in mammals (Zn1) seems to be conserved in all the species except plants. Point mutations of the Zn1 ligands present different effects, going from a null protein expression (Cys-181) to a decreased zinc amount (Cys-291) clearly correlated with a dramatic drop in protein activity and thermal stability. In particular, Cys-291, located in a long insertion (D2), is present in all vertebrates and most invertebrates (data not shown). This insertion is quite unusual because it is unstructured and placed between two contiguous helices (Fig.

Structure of a mammal IP_5 2-K

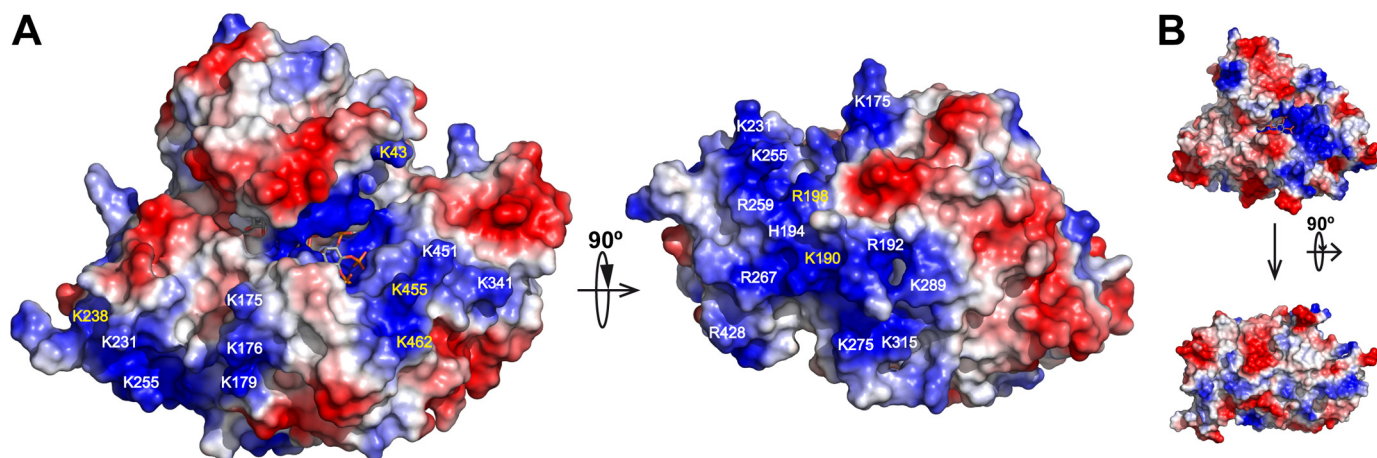


Figure 8. Basic patch on mIP_5 2-K surface. A, electrostatic surface representation of mIP_5 2-K showing its exposed basic patch. Residues conserved and non-conserved with the $AtIP_5$ 2-K are shown in yellow and white, respectively. B, same representation as in A for the $AtIP_5$ 2-K enzyme.

6C). Its role seems to support the architecture of the CIP lobe to complete the zinc-binding site and to configure a markedly basic protein surface.

Apart from these remarkable differences found, substrate recognition in both isoforms is quite similar. In previous works, $AtIP_5$ 2-K was shown to display conformational changes upon substrate binding that set the enzyme lobes together producing a closed conformation (27). Studies performed with $AtIP_5$ 2-K revealed that the IP_5 binding to the C-lobe is stronger, whereas the binding to the N-lobe is necessary for protein activation and conformational change (30). As the Zn1 is located in the CIP-lobe and far from the N-lobe, we suggest that it would affect the preliminary inositide binding events rather than the subsequent dynamic behavior of the protein. The structure solved for mIP_5 2-K probably corresponds to the close conformation in agreement with the fact that the inositide substrate is present in the structure and the good superposition of mIP_5 2-K reported here onto the $AtIP_5$ 2-K closed conformation. We do not know whether the changes reported for $AtIP_5$ 2-K also occur in the mammal isoforms and whether they are a general behavior of the IP_5 2-K family. However, the conservation of several elements involved in these changes, as the flexibility of CL3 or inositide binding by N-lobe through Arg-100 or G-loop, suggests that this open-close mechanism could be proposed for mammal IP_5 2-Ks. In relation to this, we made proteolysis experiments to check whether the substrates protect the enzyme digestion as happened in $AtIP_5$ 2-K (28). We observed no protection in mIP_5 2-K, probably due to the fact that its digestion sites are far from regions involved in the conformational change. Moreover, we did not get crystals in the absence of substrates, which could be pointing to some structural changes, although this is not conclusive either.

The structure of $AtIP_5$ 2-K revealed that IP_5 2-Ks are the most divergent among the IPK family, because it has the most elaborated CIP-lobe and binds the substrate in a very different orientation to face an axial OH to P γ of phosphate. Similarly, IP_5 2-Ks are the most divergent enzymes classified inside the PK structural superfamily, which is confirmed in the presented structure. Moreover, we observe that mIP_5 2-K lacks the N-lobe acidic residue (Glu-91, PKA nomenclature) reported to salt-

bridge to a basic residue (Lys-72, PKA nomenclature) and involved in ATP phosphate coordination. This bridge is a hallmark that identifies the active conformation of protein kinases, and equivalent residues are also present in other IPKs. In contrast, mIP_5 2-K has a different acidic residue, Asp-439, which stabilizes Arg-33 (equivalent to Lys-72 in PKAs). Therefore, in the IP_5 2-K family, this bridge could play a similar role to that described in PKs.

Brehm *et al.* (12, 13) have investigated the human (*h*) IP_5 2-K and, in particular, the possibility of additional functions apart from its catalytic role. They found that *hIP_5* 2-K colocalizes with mRNA, both in the nuclei and cytoplasm (13). The mutations on a region rich in basic residues, coincident with the $^{173}KW(K/R)(K/Q)$ motif conserved in mammals, show an altered enzyme ability to be exported out of the nuclei. This region is exposed and included within the basic patch of mIP_5 2-K identified in this work (Figs. 7C and 8). Interestingly, Lys-173 at the beginning of the above motif coordinates the inositide substrate. However, our mutagenesis experiments showed that this residue is dispensable for substrate binding, suggesting that the major role for this motif is played in the translocation process proposed by Brehm *et al.* (13). In addition, *hIP_5* 2-K was shown to be a structural component of the nucleolus acting as a molecular scaffold in nucleoli and influencing the degree of rRNA synthesis, thereby having a role in rRNA biogenesis. In relation to this, *hIP_5* 2-K interacts with three proteins (CK2, TCOF, and UBF) that regulate rRNA synthesis. In particular, UBF interacts with a basic region (^{41}RKK motif equivalent to ^{41}KKK motif in mIP_5 2-K) (12) which mutation prevents the UBF translocation out of nucleolus after *hIP_5* 2-K overexpression. As the authors predict, this region is completely exposed, and we observed that it is in a flexible loop as shown by its poor electron density. We now disclose the high and specific basic region found in the mammalian IP_5 2-K surface that could be used as a guide to find more target points.

In conclusion, our work provides novel features for the IP_5 2-K family and its mammalian isoforms. Unexpected protein regions and residues have been identified providing an illuminating picture of these enzymes. The findings comprise the characterization of subtle but important features for substrate

recognition, including unreported catalytic residues for this family, the identification of an unusual and exclusive zinc-binding site, and the conspicuous basic patch on the protein surface. Undoubtedly, the results obtained in this work provide a valuable tool for the design of therapeutics targeted at mammalian IP₅ 2-K with potential implications in health and also to perform IP₅ 2-K functional studies. Beyond the catalytic function, our work also suggests putative regions of interaction of mammalian IP₅ 2-Ks with the cognate partners necessary to accomplish their precise functions.

Experimental procedures

Protein expression and purification

Constructs for full-length IP₅ 2-K recombinant expression either in bacteria (*mipk1*/pKLSLl plasmid) or insect cells were obtained, as described by us (34), from a *mipk1* cDNA (commercial clone bc062167). To produce a truncated *mIP*₅ 2-K enzyme lacking the 21 C-terminal residues (Δ C-*mIP*₅ 2-K), a stop codon was introduced at a position coding for residue 469 of *mIP*₅ 2-K by site-directed mutagenesis and using as template the *mipk1* cDNA inserted into the pKLSLl vector (35). Point and double Δ C-*mIP*₅ 2-K mutants were obtained by site-directed mutagenesis using as template the Δ C-*mipk1*/pKLSLl plasmid. Primers used for construct preparations are shown in supplemental Table S3.

Expression and purification of Δ C-*mIP*₅ 2-K samples fused to LSL- was performed similarly to the full-length samples (34). Briefly, the protein was expressed in *Escherichia coli* BL21 Star (DE3) cells in 2TY medium supplemented with kanamycin (50 μ g ml⁻¹) at 310 K until an A_{600} of 0.9 was reached. Expression was induced with 0.3 mM isopropyl 1-thio- β -D-galactopyranoside for 96 h at 283 K. Pellets were resuspended and sonicated in buffer A (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM DTT) plus 0.2 mM PMSF and 0.05% Triton X-100. The filtrated lysate was diluted 3-fold, loaded onto a heparin column, washed with buffer B (20 mM Tris-HCl, pH 8, 50 mM NaCl, 1 mM DTT), and eluted with a 1 M NaCl gradient. The fusion protein was applied to a Sepharose CL-6B column equilibrated in buffer A and eluted using 200 mM lactose, followed by overnight cleavage with TEV protease (protease/protein mass ratio 1:40) gently rolling at 278 K. Our protein was separated from LSLt and TEV protease by a second heparin column and further purified by size-exclusion chromatography (HiLoad 16/600 Superdex 200 column) equilibrated in buffer A plus 2 mM IP₆, which was included to avoid protein precipitation. All *mIP*₅ 2-K samples used for crystallization were concentrated to around 5–6 mg ml⁻¹ and stored at 193 K. We obtained 1 mg of pure Δ C-*mIP*₅ 2-K per liter of bacteria culture. The purity of all the samples was confirmed by SDS-PAGE. For crystallization and CD analysis, WT and C291S Δ C-IP₅ 2-K samples were purified using this protocol.

Finally, for activity assays, wild-type LSLt- Δ C-*mIP*₅ 2-K and mutants were purified as follows. Clarified and filtrated cell lysate in buffer A was applied to a Sepharose CL-6B column equilibrated in buffer A. After washing with buffer A, the protein was eluted with 200 mM lactose. The sample was diluted 3-fold with 20 mM Tris-HCl, pH 8.0, loaded onto a heparin

column, and washed with buffer B and eluted with a salt gradient. The protein in final buffer C (20 mM Tris-HCl, pH 8.0, 700 mM NaCl, 1 mM DTT) was concentrated to 1–3 mg ml⁻¹ and stored at -80 °C.

Crystallization

All IP₅ 2-K samples used for crystallization were obtained in the presence of 2 mM IP₆. Best crystals obtained for the full-length *mIP*₅ 2-K (expressed either in bacteria or in insect cells) diffracted to 4–4.3 Å (34), and they did not allow the structure solution. Finally, the Δ C-*mIP*₅ 2-K construct allowed us to improve resolution to 3.2 Å from crystals grown in 0.2 M magnesium chloride, 0.1 M MES, pH 6.25, 10% (v/v) PEG 6000 and included 2 mM IP₆ and 2 mM ADP in the protein buffer. A new pH grid screen using the sample in the presence of 2 mM IP₆ allowed us to get better crystals grown in 0.2 M magnesium chloride, 0.1 M sodium acetate, pH 5.5, 16% (v/v) PEG 6000. Soaking experiments in precipitant solutions containing 10 mM IP₆ or 10 mM IP₆/ADP during 3 h yielded the complexes *mIP*₅ 2-K/IP₆ and *mIP*₅ 2-K/IP₅/ATP diffracting to 2.4 Å in the last case. Microseeding technique was necessary to improve the quality of all these crystals. For this purpose, we selected our best crystals and introduced them into 50 μ l of crystallization solution plus a seeding bead. After two cycles of 30 s vortexing and 30 s on ice, we made a seed stock. We streak-seeded the crystallization drops with a whisker using this seed stock. All IP₅ 2-K crystals appeared in a few hours after setting up the crystallization trials, and we observed that the protein is degraded in the crystallization conditions very quickly thus making crystal optimization extremely difficult.

Data collection and structural determination

Crystals were transferred for a few seconds into precipitant solution plus 20% (v/v) glycerol and then flash-cooled in liquid nitrogen. Data from IP₅ 2-K crystals were collected at 100 K in beam line BL13-XALOC of the ALBA Synchrotron (36). Δ C-*mIP*₅ 2-K crystallizes in monoclinic P2₁ space group in two different forms having one (pH 5.5) or two molecules (pH 6.25) in the asymmetric unit (Table 1). Diffraction data were indexed, integrated, and scaled using XDS (37) and merged using Aimless (38) from CCP4 suite (39, 40). Initially, Δ C-*mIP*₅ 2-K monoclinic crystals grown at pH 6.5 (3.2 Å) allowed us to get a partial model using molecular replacement with MOLREP (41) and the structure of *A. thaliana* IP₅ 2-K as a search model (Protein Data Bank code 2XAN). However, the preliminary electron density maps presented many ambiguities, although clearly showed high positive difference peaks for two possible zinc ions. An anomalous map computed with PHENIX (42) showed a strong anomalous signal in those positions (supplemental Fig. S2). Therefore, we tried SAD phasing in combination with MR (MRSAD-Auto-Rickshaw) (43). The heavy atom positions were located using PHASER (44) and refined with MLPHARE (39). The phases obtained were then combined, and density modification was performed with RESOLVE (45, 46) and PIRATE (47). Final electron density maps allowed the building of the whole chain except some exposed loops indicating the flexibility of these regions.

Structure of a mammal IP₅ 2-K

Later, the ΔC -*mIP*₅ 2-K monoclinic crystals grown at pH 5.5 allowed us to refine the structure of protein complexes with ligands IP₆ and IP₅/ATP to 2.6 and 2.4 Å maximum resolution, respectively. The structures were solved by molecular replacement using MOLREP (41) and the coordinates of ΔC -*mIP*₅ 2-K described above as a search model. The substrates/products were manually fit into the electron density maps. Although we soaked the crystals with the products IP₆/ADP, the initial electron density maps showed clear density consistent with IP₅ and ATP (supplemental Fig. S1). Then, we checked that our crystallized *mIP*₅ 2-K samples are able to catalyze both the forward and reverse reactions in solution (Fig. 3, B and C), and, therefore, we modeled the substrates IP₅/ATP in the active site. Model refinement was performed with REFMAC (48) alternating with manual model building using COOT (49). The stereochemistry of the model has been checked with PROCHECK (50). Statistics for all data processing and refinement are summarized in Table 1. Figures of the models were generated with PyMOL (51).

Circular dichroism

CD spectra were recorded using a Jasco-810 spectropolarimeter equipped with a Peltier-thermostatted cell holder. Measurements in the far-UV region (250–200 nm) were performed using the samples WT and C291S ΔC -IP₅ 2-K after protein buffer exchange to 25 mM sodium phosphate, pH 8, and at protein concentrations of 0.1 mg ml⁻¹ (10-mm path length quartz cells; bandwidth, 1 nm; response, 4 s; scan speed, 20 nm min⁻¹). Collected spectra were the average of four accumulations. The data were converted to molar ellipticities after subtraction of the buffer contribution using the average molecular mass per residue (114 Da). Thermal denaturation was monitored by CD measuring the ellipticity changes at 220 nm as the temperature was raised (20–90 °C) at 60 °C h⁻¹. The normalized ellipticity value at each temperature was calculated as $([\theta]_t - [\theta]_{25})/([\theta]_{90} - [\theta]_{25})$, where $[\theta]_t$ is the ellipticity value at temperature t , and $[\theta]_{25}$ and $[\theta]_{90}$ are the ellipticity values at 25 and 90 °C, respectively.

Protein sequence alignments and bioinformatics

IP₅ 2-Ks sequences of all kingdoms have been retrieved using BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) and Pfam (52) searches. After removing all incomplete sequences or lacking essential hallmarks for kinase function, we had a collection of the following sequences: 102 for mammals, 142 for non-mammal vertebrates, 28 for invertebrates, 158 for fungi, and 203 for plants. Sequence alignments have been performed with the Clustal Omega server (53) and corrected manually with SeaView (54). A structural alignment between *AtIP*₅ 2-K (code 1xan) and *mIP*₅ 2-K has been performed with EPSPRIT (55).

Enzyme assays

LSL- ΔC -IP₅ 2-K and ΔC -IP₅ 2-K samples showed comparable activity (data not shown); therefore, we used LSL-tagged samples for analysis of kinetic parameters of WT and mutants. For this purpose, IP₅-dependent conversion of ATP to ADP was determined by HPLC. Assays were performed in 20 mM Hepes,

1 mM MgCl₂, pH 7.3, containing 0.2–5 μM Ins(1,3,4,5,6)P₅ and 50 μM ATP in a volume of 50–100 μl at an enzyme concentration of 4 μg ml⁻¹. Reactions were stopped by the addition of 50 μl of 60 mM (NH₄)₂HPO₄, pH 3.8, with H₃PO₄. Aliquots of the reaction products were resolved by ion-pair reverse-phase chromatography (56) with the following modifications: separations were performed on a 100 × 2.1-mm Agilent X-Bridge C18 (3.5-μm particle size) column eluted at a flow rate of 0.25 ml min⁻¹. Nucleotide substrates and products (ADP and ATP) were detected at 260 nm, and the extent of conversion of one to the other was determined from the ratio of integrated peaks. The ADP content of the ATP used was less than 0.1% of the ATP peak area. Reaction velocities were calculated assuming 1:1 stoichiometry of consumption of nucleotide and inositide. Experimental data were fitted by non-linear least squares regression to the Michaelis-Menten equation in GraFit (Erithacus Software). Assays were performed in triplicate and the experiment repeated an additional three times with similar results.

Verification of reactions catalyzed by *mIP*₅ 2-K

We undertook a variety of enzyme assays to determine the identities of products formed by the enzyme. We tested the ability of ΔC -*mIP*₅ 2-K to catalyze forward “kinase” and reverse reactions.

For the forward kinase reaction, 2.4 μg of ΔC -*mIP*₅ 2-K was incubated with 200 μl of 100 μM Ins(1,3,4,5,6)P₅ (sodium salt, SiChem, Germany), 40 μM ATP in 20 mM Hepes, 1 mM MgCl₂, pH 7.3, at 37 °C. At intervals, aliquots were withdrawn and 20 μl injected onto a 250 × 3-mm CarboPac PA200 column (Dionex) eluted at a flow rate of 0.4 ml min⁻¹ with a gradient derived from buffer reservoirs containing the following: A, water; B 0.6 M methanesulfonic acid according to the following profile: time (min), % B; 0, 0; 25, 100; 38, 100. The eluate from the column was mixed in a mixing tee with color reagent (0.1% w/v Fe(NO₃)₃·9H₂O in 2% v/v HClO₄) (57) delivered at a flow rate of 0.2 ml min⁻¹ by a second HPLC pump. The combined flow was monitored at 290 nm after passage through a knitted reaction coil.

For the reverse reaction, 2.4 μg of ΔC -*mIP*₅ 2-K was incubated with 200 μl of 100 μM IP₆ (sodium salt, Merck, Germany), 50 μM ADP in 20 mM Hepes, 1 mM MgCl₂, pH 7.3, at 37 °C, with subsequent processing as above. The identity of IP₅ and IP₆ products was confirmed by chromatography of standards (SiChem or Merck) and by analysis of an IP₆ hydrolysate obtained by overnight refluxing of IP₆ in 1 M HCl, with subsequent rotary evaporation to remove HCl.

Our Ins(1,3,4,5,6)P₅ (IP₅ 2-OH) substrate contained D- and/or L-Ins(1,2,4,5,6)P₅ (IP₅ 1/3-OH), but our analysis which resolves the two *meso*-compounds (IP₅ 2-OH and IP₅ 5-OH) from the two pairs of enantiomers (IP₅ 1/3-OH and IP₅ 4/6-OH) confirmed that *mIP*₅ 2-K is an inositol 1,3,4,5,6-pentakisphosphate 2-kinase that does not accept D- and/or L-Ins(1,2,4,5,6)P₅ (mixture unknown) as substrate (Fig. 3B). HPLC traces were exported from Jasco (Great Dunmow, UK) ChromNav software as ascii files and redrawn in GraFit (Erithacus Software).

Inductively coupled plasma-optical emission spectroscopy

Metal analysis of mIP₅ 2-K was performed by optical emission spectroscopy on a Varian Vista Pro ICP-OES. Protein or buffer in which protein was prepared was diluted 125–139-fold in 18.2 megohm·cm water containing 1 N HNO₃. Diluted protein or buffer was subjected to ICP-OES on a machine calibrated with 0–4.0 μM standards of cobalt, copper, nickel, and zinc in 1 N HNO₃. All metals except zinc were close to the limit of detection in the diluted protein sample; they were only slightly above the background in the buffer (supplemental Table S2).

Author contributions—E. F.-E and J. M. G. prepared the constructs for the experiments. E. F.-E expressed, purified, and crystallized all the protein samples. E. F.-E and B. G. solved the protein structures. E. F.-E., J. S.-A., and B. G. analyzed the structural data. C. A. B. performed all enzyme assays and zinc measurements. J. S.-A. participated in the work and results and discussion. B. G. designed the research and wrote the paper. All authors edited the manuscript.

Acknowledgments—Data collection was performed at the XALOC-BL13 beamline at ALBA Synchrotron with the collaboration of ALBA staff. We thank Graham Chilvers, UEA SCI Faculty Analytical Facility, for assistance with ICP-OES. We acknowledge the support and the use of resources of the French Infrastructure for Integrated Structural Biology FRISBI ANR-10-INSB-05 and of Instruct, a Landmark ESFRI project.

References

- Stephens, L. R., Hawkins, P. T., Stanley, A. F., Moore, T., Poyner, D. R., Morris, P. J., Hanley, M. R., Kay, R. R., and Irvine, R. F. (1991) Myo-inositol pentakisphosphates. Structure, biological occurrence and phosphorylation to myo-inositol hexakisphosphate. *Biochem. J.* **275**, 485–499
- Shears, S. B. (2001) Assessing the omnipotence of inositol hexakisphosphate. *Cell. Signal.* **13**, 151–158
- Ouyang, Z., Zheng, G., Tomchick, D. R., Luo, X., and Yu, H. (2016) Structural basis and IP6 requirement for Pds5-dependent cohesin dynamics. *Mol. Cell* **62**, 248–259
- Alcázar-Roman, A. R., Tran, E. J., Guo, S., and Wenthe, S. R. (2006) Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. *Nat. Cell Biol.* **8**, 711–716
- Macbeth, M. R., Schubert, H. L., Vandemark, A. P., Lingam, A. T., Hill, C. P., and Bass, B. L. (2005) Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science* **309**, 1534–1539
- Boullaran, C., Scott, M. G., Bourougaa, K., Bellal, M., Esteve, E., Thuret, A., Benmerah, A., Tramier, M., Coppey-Moisand, M., Labbé-Jullié, C., Fähræus, R., and Marullo, S. (2007) β-Arrestin 2 oligomerization controls the Mdm2-dependent inhibition of p53. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18061–18066
- Scherer, P. C., Ding, Y., Liu, Z., Xu, J., Mao, H., Barrow, J. C., Wei, N., Zheng, N., Snyder, S. H., and Rao, F. (2016) Inositol hexakisphosphate (IP6) generated by IP5K mediates cullin-COP9 signalosome interactions and CRL function. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 3503–3508
- Byrum, J., Jordan, S., Safrany, S. T., and Rodgers, W. (2004) Visualization of inositol phosphate-dependent mobility of Ku: depletion of the DNA-PK cofactor InsP6 inhibits Ku mobility. *Nucleic Acids Res.* **32**, 2776–2784
- Barker, C. J., Illies, C., Gaboardi, G. C., and Berggren, P. O. (2009) Inositol pyrophosphates: structure, enzymology and function. *Cell. Mol. Life Sci.* **66**, 3851–3871
- Chakraborty, A., Koldobskiy, M. A., Sixt, K. M., Juluri, K. R., Mustafa, A. K., Snowman, A. M., van Rossum, D. B., Patterson, R. L., and Snyder, S. H. (2008) HSP90 regulates cell survival via inositol hexakisphosphate kinase-2. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1134–1139
- Verbsky, J., Lavine, K., and Majerus, P. W. (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. U.S.A.* **102**, 8448–8453
- Brehm, M. A., Wundenberg, T., Williams, J., Mayr, G. W., and Shears, S. B. (2013) A non-catalytic role for inositol 1,3,4,5,6-pentakisphosphate 2-kinase in the synthesis of ribosomal RNA. *J. Cell Sci.* **126**, 437–444
- Brehm, M. A., Schenk, T. M., Zhou, X., Fanick, W., Lin, H., Windhorst, S., Nalaskowski, M. M., Kobras, M., Shears, S. B., and Mayr, G. W. (2007) Intracellular localization of human Ins(1,3,4,5,6)P₅ 2-kinase. *Biochem. J.* **408**, 335–345
- Berridge, M. J., and Irvine, R. F. (1984) Inositol trisphosphate, a novel second messenger in cellular signal transduction. *Nature* **312**, 315–321
- Wang, H., Falck, J. R., Hall, T. M., and Shears, S. B. (2011) Structural basis for an inositol pyrophosphate kinase surmounting phosphate crowding. *Nat. Chem. Biol.* **8**, 111–116
- Qian, X., Mitchell, J., Wei, S. J., Williams, J., Petrovich, R. M., and Shears, S. B. (2005) The Ins(1,3,4)P₃ 5/6-kinase/Ins(3,4,5,6)P₄ 1-kinase is not a protein kinase. *Biochem. J.* **389**, 389–395
- Onnebo, S. M., and Saiardi, A. (2007) Inositol pyrophosphates get the vip1 treatment. *Cell* **129**, 647–649
- Saiardi, A., Caffrey, J. J., Snyder, S. H., and Shears, S. B. (2000) The inositol hexakisphosphate kinase family. Catalytic flexibility and function in yeast vacuole biogenesis. *J. Biol. Chem.* **275**, 24686–24692
- Shears, S. B. (2004) How versatile are inositol phosphate kinases? *Biochem. J.* **377**, 265–280
- Phillippy, B. Q., Ullah, A. H., and Ehrlich, K. C. (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
- Irvine, R. F., Letcher, A. J., Heslop, J. P., and Berridge, M. J. (1986) The inositol tris/tetrakisphosphate pathway—demonstration of Ins(1,4,5)P₃ 3-kinase activity in animal tissues. *Nature* **320**, 631–634
- Cheek, S., Zhang, H., and Grishin, N. V. (2002) Sequence and structure classification of kinases. *J. Mol. Biol.* **320**, 855–881
- González, B., Schell, M. J., Letcher, A. J., Vepintsev, D. B., Irvine, R. F., and Williams, R. L. (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
- Wang, H., DeRose, E. F., London, R. E., and Shears, S. B. (2014) IP6K structure and the molecular determinants of catalytic specificity in an inositol phosphate kinase family. *Nat. Commun.* **5**, 4178
- González, B., Baños-Sanz, J. I., Villate, M., Brearley, C. A., and Sanz-Aparicio, J. (2010) 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. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 9608–9613
- Holmes, W., and Jögl, G. (2006) Crystal structure of inositol phosphate multikinase 2 and implications for substrate specificity. *J. Biol. Chem.* **281**, 38109–38116
- Baños-Sanz, J. I., Sanz-Aparicio, J., Whitfield, H., Hamilton, C., Brearley, C. A., and González, B. (2012) Conformational changes in inositol 1,3,4,5,6-pentakisphosphate 2-kinase upon substrate binding: role of N-terminal lobe and enantiomeric substrate preference. *J. Biol. Chem.* **287**, 29237–29249
- Gosein, V., Leung, T. F., Kraiden, O., and Miller, G. J. (2012) Inositol phosphate-induced stabilization of inositol 1,3,4,5,6-pentakisphosphate 2-kinase and its role in substrate specificity. *Protein Sci.* **21**, 737–742
- Gosein, V., and Miller, G. J. (2013) Conformational stability of inositol 1,3,4,5,6-pentakisphosphate 2-kinase (IPK1) dictates its substrate selectivity. *J. Biol. Chem.* **288**, 36788–36795
- Gosein, V., and Miller, G. J. (2013) Roles of phosphate recognition in inositol 1,3,4,5,6-pentakisphosphate 2-kinase (IPK1) substrate binding and activation. *J. Biol. Chem.* **288**, 26908–26913
- Verbsky, J. W., Wilson, M. P., Kisseleva, M. V., Majerus, P. W., and Wenthe, S. R. (2002) The synthesis of inositol hexakisphosphate. Characterization of human inositol 1,3,4,5,6-pentakisphosphate 2-kinase. *J. Biol. Chem.* **277**, 31857–31862

32. Sweetman, D., Johnson, S., Caddick, S. E., Hanke, D. E., and Brearley, C. A. (2006) Characterization of an *Arabidopsis* inositol 1,3,4,5,6-pentakisphosphate 2-kinase (AtIPK1). *Biochem. J.* **394**, 95–103
33. Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549
34. Franco-Echevarria, E., Sanz-Aparicio, J., Troffer-Charlier, N., Poterszman, A., and Gonzalez, B. (2017) Crystallization and preliminary X-ray diffraction analysis of a mammal inositol 1,3,4,5,6-pentakisphosphate 2-kinase. *Protein J.* 10.1007/s10930–017-9717-y
35. Angulo, I., Acebrón, I., de las Rivas, B., Muñoz, R., Rodríguez-Crespo, I., Menéndez, M., García, P., Tatenó, H., Goldstein, I. J., Pérez-Agote, B., and Mancheño, J. M. (2011) High-resolution structural insights on the sugar-recognition and fusion tag properties of a versatile β -trefoil lectin domain from the mushroom *Laetiporus sulphureus*. *Glycobiology* **21**, 1349–1361
36. Juanhuix, J., Gil-Ortiz, F., Cuní, G., Colldelram, C., Nicolás, J., Lidón, J., Boter, E., Ruget, C., Ferrer, S., and Benach, J. (2014) Developments in optics and performance at BL13-XALOC, the macromolecular crystallography beamline at the ALBA synchrotron. *J. Synchrotron Radiat.* **21**, 679–689
37. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132
38. Evans, P. R. (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292
39. Collaborative Computational Project No. 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
40. 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., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
41. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 22–25
42. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix. refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367
43. Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S., and Tucker, P. A. (2009) On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1089–1097
44. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
45. Terwilliger, T. C. (1999) Reciprocal-space solvent flattening. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 1863–1871
46. Terwilliger, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965–972
47. Cowtan, K. (2000) General quadratic functions in real and reciprocal space and their application to likelihood phasing. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 1612–1621
48. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
49. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
50. Laskowski, R. A., M. M. W., Moss DS, Thornton J M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291
51. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System*. Version 1.6.0.0. DeLano Scientific LLC, San Carlos, CA
52. Finn, R. D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L., Potter, S. C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G. A., Tate, J., and Bateman, A. (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285
53. Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y. M., Buso, N., and Lopez, R. (2015) The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* **43**, W580–W584
54. Gouy, M., Guindon, S., and Gascuel, O. (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224
55. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acids Res.* **31**, 3320–3323
56. Caddick, S. E., Harrison, C. J., Stavridou, I., Mitchell, J. L., Hemmings, A. M., and Brearley, C. A. (2008) A *Solanum tuberosum* inositol phosphate kinase (StITPK1) displaying inositol phosphate-inositol phosphate and inositol phosphate-ADP phosphotransferase activities. *FEBS Lett.* **582**, 1731–1737
57. Phillippy, B. Q., and Bland, J. M. (1988) Gradient ion chromatography of inositol phosphates. *Anal. Biochem.* **175**, 162–166