Biochemical and structural characterization of a novel arginine kinase from the spider *Polybetes pythagoricus*

Aldana Laino¹ *, Alonso A. Lopez-Zavala² *, Karina D. Garcia-Orozco³, Jesus S. Carrasco-Miranda³, Marianela Santana¹, Vivian Stojanoff⁴, Rogerio R. Sotelo-Mundo³ and Carlos Fernando Garcia¹

¹Instituto de Investigaciones Bioquímicas de La Plata “Dr. Prof. Rodolfo R. Brenner”, Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina
²Departamento de Ciencias Químico-Biológicas, Universidad de Sonora, Hermosillo, Sonora, Mexico
³Laboratorio de Estructura Biomolecular, Centro de Investigación en Alimentación y Desarrollo, A.C., Hermosillo, Sonora, Mexico
⁴Photon Science Directorate, National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY, United States of America
*These authors contributed equally to this work.

**ABSTRACT**

Energy buffering systems are key for homeostasis during variations in energy supply. Spiders are the most important predators for insects and therefore key in terrestrial ecosystems. From biomedical interest, spiders are important for their venoms and as a source of potent allergens, such as arginine kinase (AK, EC 2.7.3.3). AK is an enzyme crucial for energy metabolism, keeping the pool of phosphagens in invertebrates, and also an allergen for humans. In this work, we studied AK from the Argenittenan spider *Polybetes pythagoricus* (*PpAK*), from its complementary DNA to the crystal structure.

The *PpAK* cDNA from muscle was cloned, and it is comprised of 1068 nucleotides that encode a 384-amino acids protein, similar to other invertebrate AKs. The apparent Michaelis-Menten kinetic constant (*Km*) was 1.7 mM with a *kcat* of 75 s⁻¹. Two crystal structures are presented, the apo*Pv*AK and *Pp*AK bound to arginine, both in the open conformation with the active site lid (residues 310–320) completely disordered. The guanidino group binding site in the apo structure appears to be organized to accept the arginine substrate. Finally, these results contribute to knowledge of mechanistic details of the function of arginine kinase.

**INTRODUCTION**

Arginine kinase (AK) (ATP: L-arginine phosphotransferase, EC 2.7.3.3) is an enzyme which catalyzes the reversible transfer of the γ-phosphoryl group (PO₄⁻) of ATP to L-arginine (*Arjunwadkar & Raghupathi Rami Reddy, 1985; Blethen, 1970*) and is widely distributed in invertebrates. AK is found mainly in muscle and heart tissue, although it is...
also expressed in gills, intestine, digestive glands (Blethen & Kaplan, 1968) and eggs (Ratto & Christen, 1988). Phosphoarginine is accumulated in tissues when ATP is present, and ADP is back-phosphorylated by AK.

AK has been described in insects such as beetles (Tanaka et al., 2007), lepidopterans (Binder et al., 2001; Chamberlin, 1997), hymenopterans (Wang et al., 2009a), cockroaches (Brown & Grossman, 2004) and locusts (Wu et al., 2007). It is present in crustaceans (Garcia-Orozco et al., 2007), mollusks and echinoderms among other invertebrates (Morrison, 1973). Considering that the sub-phylum Chelicerate is comprised of more than 75,000 species, only the AK from horseshoe crab (Strong & Ellington, 1995), scorpion (Arjunwadkar & Raghupathi Rami Reddy, 1985) and spider (Blethen & Kaplan, 1968; Bobolea et al., 2011) have been studied.

In spiders, fast movements are limited by depletion of phosphagens, accumulation of lactate and limited aerobic metabolism (Prestwich, 1988), suggesting that AK is key for this process. AK has a similar function to creatinine kinase (CK) from vertebrates (ATP: N-creatine phosphotransferase, EC 2.7.3.2,) providing fast energy during muscular contraction by the back-phosphorylation of ADP by phospho-arginine (Bessman & Carpenter, 1985). In the fire ant (Solenopsis invicta), AK is highly expressed in the tissues (thorax and head) and castes (workers) with a larger energetic demand, confirming the connection between the biochemical function and the physiology (Wang et al., 2009a).

Besides the role of AK in energy maintenance, it is also important as an allergen. Crustacean AK is a well-characterized crustacean allergen (Garcia-Orozco et al., 2007; Yu & Li, 2003) and also described in insects (Binder et al., 2001; Sookrung et al., 2006) and a spider (Bobolea et al., 2011). Recently it has been shown that there is cross-reactivity between crustacean and insect AKs, possibly due to structural epitopes (Srinroch et al., 2015). On the other hand, because the enzyme plays a significant role in energy metabolism and it is absent in vertebrates, AK was proposed as a target for insecticides (Wang et al., 2009a; Wu et al., 2009) like the insect Ctenocephalides sp. (Werr, Cramer & Ilg, 2009). To contribute to the biochemical and structural knowledge of phosphagen kinases, we studied the AK from Polybetes pytagoricus (PpAK). This represents the first structural study of a chelicerate phosphagen kinase and allergen.

**MATERIALS & METHODS**

**RNA isolation and cDNA cloning of arginine kinase**

Adult spider *P. pythagoricus* were collected from eucalyptus trees in La Plata City, (Province of Buenos Aires, Argentina; 34°52’56”S, 57°56’07”W). Spiders were anesthetized by cooling, and leg muscle was dissected. Total RNA was isolated from 100 mg of tissue using Trizol reagent (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer’s instruction. The RNA concentration was measured using a nano spectrophotometer (Nanodrop ND-1000) and the integrity was determined by agarose gel electrophoresis. A double stranded cDNA library was constructed following the instructions of Smart cDNA library Construction Kit (Clontech Laboratories Inc. Palo Alto, CA, USA). The synthesized cDNAs were electrophoresed and excised from the agarose gel separating three groups.
according to the size (0.2–1 kb, >1–3 kb, >3 kb) to be ligated in the TA cloning vector TOPO XL PCR (Invitrogen, Carlsbad, CA, USA). The ligation and transformation were carried out according to the manufacturer’ manual. White colonies containing inserts bigger than 300 bp were detected by PCR and selected to be cultured overnight in 2 ml of LB broth with ampicillin. From these, 384 colonies were selected, plasmid was obtained using an alkaline lysis protocol (Ausubel et al., 1995), and sent for Sanger DNA sequencing to the University of Arizona Genetics Core Lab (Tucson AZ, USA). ESTs were assembled into contiguous sequences (contigs) using the CLC Main Workbench program (CLC Bio, a QIAGEN Company; Aarhus, Denmark). All contigs were searched against public databases in GenBank and Swiss-Prot by BLASTX and BLASTN and one 632 bp contigs with high identity to invertebrate AK was obtained. The Rapid Amplification of cDNA Ends (RACE) kit (Thermo Fisher Scientific, Waltham, MA, USA) was used to complete the partial sequence of cDNA that codify for arginine kinase (PpAK). The PpAK specific primers used for RACE were AKAF1 5’ ACCATGGTTGACCAAGCTAC, AKAF2 5’ GACAAACATCCTCCAACCGAC and AKAF2nes 5’ CACCATTGTCAACGTCGATCC. The nucleotide and deduced amino acid sequence were deposited at GenBank with the accession number MF001441.

**Purification of recombinant PpAK protein**

A synthetic gene optimized for *E. coli* overexpression including a His-tag at the N-terminus was synthesized using the *P. pythagoricus* AK amino acid sequence by DNA 2.0. The recombinant construct was transformed into the *E. coli* BL21(DE3) Rosetta strain. Cultures were grown at 37 °C by inoculating 1 ml of an overnight culture into 100 ml of LB medium containing 50 mg/ml ampicillin. Incubation in an orbital shaker at 200 rpm and 37 °C was continued until the culture reached an A600 of 0.6. and isopropyl thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM. After 4 h of induction at 28 °C, bacteria were harvested by centrifugation at 4,000× g for 5 min. The pellet was resuspended in cold lysis buffer (50 mM Tris, 0.5 mM PMSF, 500 mM NaCl, 5 mM EDTA, 1 mM DTT), sonicated on ice (10 W, 10 s, 10 times with 10 s interval). The supernatant was obtained by centrifuging the lysate at 10,000× g for 20 min at 4 °C.

The His-tagged recombinant protein was purified using immobilized metal ion affinity chromatography (IMAC) (IMAC Sepharose 6 Fast Flow; GE Healthcare, Little Chalfont, UK). The resin was equilibrated with buffer A (0.5 M NaCl, 20 mM Tris-HCl pH 7.9). The sample was loaded onto the IMAC column. The column was washed with 5 column volumes of buffer A and the protein was eluted with 5 volumes of buffer B (0.5 M NaCl, 20 mM Tris-HCl pH 7.9, 500 mM imidazole (Sigma-Aldrich, St. Louis, MO, USA)) at a flow rate of 0.5 ml/min. Fractions of 1 ml were collected during the elution and monitored at 280 nm to detect the presence of protein. The fraction containing protein was isolated and dialyzed against 100 mM glycine and 10 mM beta-mercapto-ethanol. The purity of the AK was determined by 12% SDS-PAGE using stain-free gels (Bio-Rad Laboratories, Hercules, CA, USA), and the protein concentration was estimated by the bicinchoninic acid method (Pierce Biotechnology, Waltham, MA, USA) using bovine serum albumin as standard.
AK activity
AK Activity was measured by determining the rate of formation of ADP according to the protocol described by Blethen (1970) with modifications (García-Orozco et al., 2007). In this system three coupled reactions are used: (1) the AK enzyme phosphorylates L-arginine from ATP to form ADP, (2) ADP is again phosphorylated by the action of pyruvate kinase forming pyruvate, (3) the latter is reduced by oxidizing NADH lactate dehydrogenase. The rate of NADH oxidation was measured spectrophotometrically at 340 nm. The reaction was done at 30 °C in a quartz cuvette containing 1 ml of buffer containing 178 mM glycine, 0.33 mM 2-mercaptoethanol, 133 mM potassium chloride, 13 mM magnesium sulfate, 20 mM phospho (enol) pyruvate, 6.7 mM ATP, 0.13 mM nicotinamide adenine dinucleotide (NADH), 2 U of pyruvate kinase, 3 U of lactic dehydrogenase, 17 mM L-arginine and 30 µg of PpAK at pH 8.6. This amount of enzyme is equivalent to 0.75 µmol using the amino acid deduced molecular mass or a final concentration of 0.76 mM in the reaction cuvette.

Control reactions were performed with all reagents except the AK enzyme and were monitored by absorbance at 340 nm. Absorbance data was collected for 5 min. All data processing was done in GraphPad Prism version 6.0 using the non-linear regression kinetic model to calculate the Michaelis–Menten constants. One unit or arginine kinase is defined as the amount of enzyme that generates 1.0 µmol of phospho-L-arginine per minute using ATP and L-arginine as substrates, at pH 8.6 and 30 °C (Blethen & Kaplan, 1968).

Crystallization of apo form and arginine complex of PpAK
Purified PpAK was exhaustively dialyzed against buffer Tris-HCl 20 mM pH 7.9 to eliminate salts and imidazole and concentrated to 24 mg/ml using a 10-kDa membrane cartridge (Amicon; Millipore, Billerica, MA, USA). Crystallization experiments were done by the vapor-diffusion method using the hanging-drop technique in 24-well Limbro plates. Hampton Research Crystal Screen I kit was used to set up each drop by adding 0.4 mL of each precipitating solution to the well and mixing in the siliconized coverglass 2 µl PpAK and 2 µl of each crystallization solution. Each well was carefully sealed with vacuum grease and stored at 16 °C. Plates were monitored each week with a stereoscope and under polarized light. Apo-PpAK crystal grew after two weeks as a bunch of large square and clear plates (0.25 × 0.15 × 0.05 mm) in solution # 9 (0.2 M ammonium acetate, 0.1 M sodium citrate tribasic dihydrate pH 5.6 and 30% w/v polyethyleneglycol 4000). An acupuncture needle was used to carefully separate single crystals, transferred to a cryoprotectant solution for 5 min (solution # 9 supplemented with 25% glycerol) and then loop-mounted and super-cooled under direct liquid nitrogen at 100 K. PpAK in arginine (Arg-PpAK) binary complex crystal was prepared by soaking apo-PpAK single crystal in the same crystallization solution containing 20 mM arginine by 10 min at 16 °C. The soaked crystal was then super-cooled by using the cryoprotectant solution with arginine, to keep the substrate concentration constant.

X-ray data collection and crystal structure determination
Apo-PpAK X-ray diffraction experiments were performed in a Bruker D8 QUEST diffractometer equipped with a CuKα Incoatec microfocus sealed tube at 1.54 Å wavelength.
and a Photon 100 CMOS detector (100 cm²). X-ray data were collected at 100 K by omega-phi scan with 30 s exposure time with and oscillation range of 0.5° and the crystal to detector distance was kept at 60 mm. Dataset was integrated, scaled and space group determination by using Proteum II® software suite provided by Bruker. Apo-PpAK crystal belongs to P2₁ space group with unit cell dimensions of \( a = 48.13, b = 61.11, c = 60.99; \alpha = \gamma = 90^\circ \) and \( \beta = 95.86^\circ \). Cell content analysis indicates that was one PpAK molecules per asymmetric unit with a \( V_m = 2.23\text{Å}^3 \text{D}^{-1} \) and solvent content = 44.74% (Matthews, 1968).

Whereas, Arg-PpAK crystals were diffracted on beam line 14-1 of the Stanford Synchrotron Light Source at SLAC National Accelerator Laboratory (Menlo Park, CA, USA) using a MARmosaic CCD detector. X-ray data set (386 frames) was collected at \( \lambda = 0.98 \) Å with an oscillation range of 0.7° at 100 K. The crystal was exposed to 6 s to the beam and the crystal-to-detector distance was 230 mm. Arg-PpAK data set was analyzed using the XDS package (Kabsch, 2010). Also, Arg-PpAK crystal has the same space group as apo-PpAK crystal and slight differences in cell dimensions and cell content.

Phases were determined by molecular replacement in Phaser (McCoy et al., 2005), as part of the CCP4 suite (Winn et al., 2011) using Litopenaeus vannamei arginine kinase coordinates as the template (PDB 4AM1) for apo-PpAK. While apo-PpAK was used as the search model for solving phase of Arg-PpAK data set. Refinement was achieved by combining several automated and manual building-cycles using Phenix (Adams et al., 2010) and Coot (Emsley et al., 2010) programs, respectively. Final models were validated by Molprobity (Adams et al., 2010) and deposited in the Protein Data Bank (apo-PpAK = PDB 5U8E and Arg-PpAK = PDB 5U92). Figures were drawn using CCP4-molecular graphics software (McNicholas et al., 2011). Electron density of ligands was calculated using Polder maps (Liebschner et al., 2017) as implemented in Phenix, which removes solvent effects and provides a better representation for ligand into the active site. Quality of the structures was assessed by parameters like clash score, rotamer outliers, C-β outliers and residues with bad angles and bad bond distances as implemented in Molprobity (Chen et al., 2010).

RESULTS

During the sequencing of an P. pythagoricus cDNA library, a partial cDNA sequence that clearly matched for an arginine kinase was obtained. Using RACE, we obtained a 1068 cDNA sequence that encodes an open reading frame for AK, with a deduced amino acid sequence of 43 kDa protein, with a theoretical isoelectric point of 6.37.

AK is highly conserved as observed in a sequence alignment with invertebrates (Fig. 1). The sequence contains the arginine and ADP-binding site, as described for Limulus AK, and the highly-conserved residues Asp62 and Arg193 that interact with the substrates (Fig. 1).

Based on the amino acid sequence, an E. coli codon-optimized synthetic gene was designed for expression in E. coli. An N-terminal histidine tag was included in the design of the synthetic gene to allow amenable purification by nickel affinity chromatography (IMAC). Upon expression in E. coli BL21 (DE3) cells, induction of a protein with the expected size was obtained in the soluble fraction. Using IMAC, a single protein was purified with molecular weight of 43 kDa.

**Table 1**  Comparison of kinetic parameters of different arthropod AKs.

<table>
<thead>
<tr>
<th>Species</th>
<th>Reference</th>
<th>$K_{m}$ (mM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_{m}$ (s$^{-1}$ mM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Polybetes pythagoricus</em> (spider)</td>
<td>Present study</td>
<td>1.7</td>
<td>75</td>
<td>44.1</td>
</tr>
<tr>
<td><em>Dugesia hentzi</em> (spider)</td>
<td>Blethen (1972)</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Limulus polyphemus</em> (horseshoe crab)</td>
<td>Blethen (1972)</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Palamneus phipsoni</em> (scorpion)</td>
<td>Arjunwadkar &amp; Raghupathi Rami Reddy (1985)</td>
<td>0.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Locusta migratoria manilensis</em> (Insect)</td>
<td>Wu et al. (2007)</td>
<td>0.95</td>
<td>159.4</td>
<td>169</td>
</tr>
<tr>
<td><em>Cissites cephalotes</em> (insect)</td>
<td>Tanaka et al. (2007)</td>
<td>1.01</td>
<td>2.02</td>
<td>2.01</td>
</tr>
<tr>
<td><em>Periplaneta americana</em> (Insect)</td>
<td>Brown &amp; Grossman (2004),</td>
<td>0.49</td>
<td>1.30</td>
<td>2.65</td>
</tr>
<tr>
<td><em>Antarctic krill</em> (Crustacea)</td>
<td>Si et al. (2014)</td>
<td>0.30</td>
<td>84.8</td>
<td>282.7</td>
</tr>
<tr>
<td><em>Litopenaeus vannamei</em> (marine shrimp)</td>
<td>Lopez-Zavala et al. (2016)</td>
<td>0.32</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The steady-state kinetic analysis of PpAK using L-Arg as substrate led to an apparent Michaelis–Menten kinetic constant ($K_{m}$) was 1.7 mM; $V_{max}$ was and 27.8 μmol min$^{-1}$, and $k_{cat}$ was 75 s$^{-1}$. Table 1 shows in comparative form the kinetics of AK in relation to other arthropods.

Besides the biochemical characterization, we obtained the crystallographic structures of PpAK (Fig. 2) in substrate-free from (apo-Pp-AK) and in a binary complex with the guanidino group substrate-arginine (Arg-Pp-AK) at 2.1 and 2.0 Å resolution, respectively
Figure 2. **Structural alignment.** Superposition of crystallographic structures of free-ligand (grey) and arginine binary complex (yellow) PpAK. The arginine substrate is shown as spheres colored by atom type (green for carbon, blue for nitrogen and red for oxygen) in the hinge between N-terminal and C-terminal domains.

(Table 2). Both structures crystallized in the same spacegroup and were solved by the molecular replacement method using the *Litopenaeus vannamei* (marine Pacific shrimp) arginine kinase (PDB 4AM1) as the search model (Lopez-Zavala et al., 2012). Structures were well refined and good stereochemical parameters were obtained (Table 3).

In AKs and in phosphagen kinases in general, the canonical fold is comprised of two domains; a small N-terminal globular domain and a large C-terminal α-β domain. In PpAK, the N-terminal domain (residues 1–100), is comprised of α-helices and the amino acids residues involved in arginine substrate binding (Fig. 2). This domain is classified as the ATP:guanido phosphotransferase N-terminal domain on the InterPro database (IPR022413) (Finn et al., 2017), or as the transferase creatine kinase superfamily 1.10.135.10 in the CATH (Class-Architecture-Topology-Homologous superfamily) protein structure classification (Dawson et al., 2017).

The large C-terminal domain (residues 101–357), is formed by an α-β fold and contains the ATP-binding site. It is classified as the ATP:guanido phosphotransferase, catalytic domain (IPR022414) or as the CATH superfamily 3.30.590.10.
Table 2  Data reduction statistics of PpAK structures. Values in parenthesis represent the statistics at the highest resolution bin.

<table>
<thead>
<tr>
<th>Data set</th>
<th>apo-PpAK</th>
<th>arg-PpAK</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray source</td>
<td>Bruker D8-Quest</td>
<td>SSRL Beam-line-14-1</td>
</tr>
<tr>
<td>Detector</td>
<td>Photon 100 CMOS</td>
<td>MAR-Mosaic 325 CCD</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.54 Å</td>
<td>0.98 Å</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
<td>P2₁</td>
</tr>
<tr>
<td>Unit-cell parameters (Å)</td>
<td>$a = 48.13$, $b = 61.11$, $c = 60.99$ $β = 95.86^\circ$</td>
<td>$a = 48.09.3$, $b = 61.45$, $c = 60.6$ $β = 95.75^\circ$</td>
</tr>
<tr>
<td>Number of residues</td>
<td>357</td>
<td>357</td>
</tr>
<tr>
<td>Monomers per asymmetric unit</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Mathews coefficient (Å Da$^{-1}$)</td>
<td>2.23</td>
<td>2.21</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>44.74</td>
<td>44.22</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>20.3–2.18 (2.25–2.18)</td>
<td>19.73–2.0 (2.07–2.0)</td>
</tr>
<tr>
<td>Total reflections</td>
<td>246,281 (18,398)</td>
<td>130,527 (12,408)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>35,766 (4,274)</td>
<td>23,798 (2,308)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>6.8 (4.3)</td>
<td>5.5 (5.4)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.43 (94.27)</td>
<td>99.79 (98.80)</td>
</tr>
<tr>
<td>Mean $I/\sigma(I)$</td>
<td>24.7 (3.34)</td>
<td>17.5 (5.64)</td>
</tr>
<tr>
<td>$CC_{1/2}$</td>
<td>0.98 (0.739)</td>
<td>0.99 (0.97)</td>
</tr>
<tr>
<td>$R_{\text{symm}}$</td>
<td>0.0847 (0.389)</td>
<td>0.0635 (0.293)</td>
</tr>
<tr>
<td>$R_{\text{meas}}$</td>
<td>0.0476 (0.2845)</td>
<td>0.0703 (0.328)</td>
</tr>
<tr>
<td>Wilson plot $B$ value</td>
<td>19.38</td>
<td>24.9</td>
</tr>
</tbody>
</table>

Notes.

$^aR_{\text{symm}} = \sum_h \sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_h \sum_i I_i(hkl)$, where $I_i(hkl)$ and $\langle I(hkl)\rangle$ represent the diffraction-intensity values of the individual measurements and the corresponding mean values. The summation is over all unique measurements.

$^bR_{\text{meas}}$ is a redundancy-independent version of $R_{\text{symm}}$, $R_{\text{meas}} = \sum_h \sqrt{n_h/n_T} - 1 \sum_h n_h I_h / \sum_h n_h I_h$, where $I_h = 1/m_\sum h_i$.

Interestingly, the active site is located between both domains and a large conformational change occurs during catalysis (Zhou, Ellington & Chapman, 2000). The crystallographic model includes 340 of the 357 amino acids deduced from the cDNA sequence. Two loops (292–297 and 310–320) show weak electron density and therefore those residues are not present in the final structural model. The guanidino group binding loop (Asp62, Ser63, Gly64 and Val65) and stabilizing residues (Try68, Cys271 and Glu225) also are well-ordered. 347 amino acid residues were fitted in the final model (see Tables 2 and 3 for statistical details), as in apo-PpAK structure residues 310–320 are disordered. Arg-PpAK binary complex was found is the open conformation similar to the free substrate PpAK model (Lopez-Zavala et al., 2013). Superposition of both structures do not show overall conformational changes upon arginine binding (RMSD = 0.36 Å for C-α atoms) (Fig. 2).

The structure of substrate-bound PpAK was obtained by soaking crystals of the apoAK in 20 mM arginine dissolved in the crystallization solution. The electron density was weak in a regular electron density map, although the substrate could be modelled in the arginine-binding site (Fig. 3A). Using a Polder map, which is an omit map where the bulk solvent is excluded during calculations, generated a much better electron density
Table 3  Refinement statistics. Crystallographic refinement statistics of PpAK structures.

<table>
<thead>
<tr>
<th>Data set</th>
<th>apo-PpAK</th>
<th>Arg-PpAK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{work}/R_{free}(5%)$</td>
<td>0.1747/0.2316</td>
<td>0.1679/0.2365</td>
</tr>
</tbody>
</table>

Content of asymmetric unit
- Protein residues: 340
- Ligands: Sodium
- Water molecules: 278

RMSD from ideal
- Bond length (Å): 0.008/0.007
- Bond angles (°): 1.07/1.034

Mean overall $B$ value (Å$^2$)
- Protein: 22.3/27.8
- Solvent: 30.40/38.7

Ramachandran plot, residues in
- Most favored regions: 333 (98%)/341 (98%)
- Additionally allowed regions: 7 (2.10%)/7 (2.06%)
- Outliers: 1 (0.3%)/0 (0%)

Clash score$^a$ (all atoms): 5.0/2.93

Rotamer outliers (%), goal < 1%: 0/0.34

C-β outliers deviations > 0.25 Å(goal 0): 0/0

Residues with bad bonds (%), goal < 1%: 0/0

Residues with bad angles (%), goal < 0.5%: 0/0.03

PDB code: 5U8E/5U92

Notes.
$^a$Clash score is the number of serious steric overlaps (>0.4 Å) per 1,000 atoms.
RMSD, root mean square deviation.

that supports modeling L-arginine into the active site and the interactions with PpAK side chains (Fig. 3B).

Arginine-substrate carboxylate is stabilized by hydrogen bonds to backbone Gly64, Val65 and Gly66 and the hydroxyl group of Tyr68 to arginine amino group. The guanidino group of the arginine substrate has specific contacts in the active site. A salt bridge with Glu225 and a hydrogen bond with Cys271 position the substrate in a region with poor density. The sidechains around the Cα carbons are better ordered and have contacts with Tyr68, Ser63 and Gly64 (Fig. 3B).

Also, a superposition of N-terminal domain of both PpAK structures we found that all residues involved in arginine binding-site are in the same position (Fig. 3). Comparing the apo vs. Arg-bound PpAK structures, it is observed that the Glu225 carboxylate lateral chain is slightly rotated to favour a hydrogen bond of arginine guanidino group end (Fig. 3). As mentioned before, in both apo- and PpAK-Arg structures, the loop 310–320 was disordered. A structural similarity search of Pp AK-Arg using the Dali server (Holm & Rosenstrom, 2010) against the PDB database, we found that the most similar structure was the shrimp apo-AK, (PDB entry 4AM1, RMSD 1.0 Å). The most similar fully occupied active site AK was the one from the horseshoe crab (Limulus polyphemus transition state complex, PDB entry 1M15, RMSD = 2.86 Å).
DISCUSSION

Although spiders are one of the most significant and numerous orders of the arachnid class, they are underrepresented in terms of molecular and structural information. At this time, the information regarding the AK from spiders is restricted to only a few species, such as *Holocnemus pluchei* (Bobolea et al., 2011), *Pholcus phalangioides* and *Dugesiella hentzi* (Blethen & Kaplan, 1968). In this work, we sequenced, cloned and overexpressed a spider arginine kinase (PpAK), an enzyme crucial for rapid energy fluxes in invertebrates. Based on the cDNA PpAK sequence, the recombinant enzyme was overexpressed showing an electrophoretical molecular mass of 40 kDa, similar to the AKs of arthropods, horseshoe crab, crustaceans and insects (Rosenthal, Dahlman & Robinson, 1977; Wang et al., 2009b). In relation to the calculated isoelectric point pH 6.4 of expressed PpAK which is different to native AK from lobster *Homarus vulgaris* at pH 5.4 (Virden, Watts & Baldwin, 1965), and expressed AK2 and AK3 from shrimp *Neocaridina denticulata* at pH 5.8, 6.0 (Iwanami et al., 2009), but similar to the AK from insect *Locusta migratoria manilensis* pH 6.3 (Wu et al., 2007) and AK from the crustacean *N. denticulata* (AK1) pH 6.3 (Iwanami et al., 2009).

A multiple alignment revealed that PpAK has the highest identity with the unique AK sequence from the spider *Stegodyphus mimosarum* (93%). Although *P. pythagoricus* and *S. mimosarum* share the araneomorphae suborder they belong to two phylogenetically distant families. The AK from *Limulus polyphemus* (the horseshoe crab) shares 77% identity which is expected for an arthropod in the class Chelicerata as they diverged early in the evolution of insect and crustaceans (Strong & Ellington, 1995), and is considered one of the oldest/primitive arthropods.

Since AK is related in function to creatine kinase of vertebrate, glycocyamine kinase of invertebrates, lombricine of annelida and taurocyamine kinase of protics and invertebrates (Dumas & Camonis, 1993; Suzuki et al., 2000), the relationship of these different enzymes
with AK possibly due to these different enzymes evolved from a common ancestor (Morrison, 1973; Perovic-Ottstadt et al., 2005; Suzuki & Furukohri, 1994).

PpAK expression in *E. coli* and determination of enzymatic activity in the purified recombinant protein showed that the gene encodes an active AK. The specific activity of the recombinant protein (14 U/mg) is similar to that obtained from AK in muscle of the spider *Pholcus phalangoides* (15.3 U/mg) (Blethen & Kaplan, 1968) and in the claw muscle of the scorpion *Palameus phipsoni* (14.1 U/mg) (Arjunwadkar & Raghupathi Rami Reddy, 1985). It also matches with other recombinant AKs: fly *Lucilia cuprina* 14 U/mg (Ilg & Werr, 2012), *Musca domestica* (10.3 U/mg) (Rockstein & Kumar, 1972). Furthermore, it is close to the activity values of the AK purified from muscle from *Litopenaeus vannamei* (8.8 U/mg) (García-Orozco et al., 2007) and *Homarus americanus* (15.6 U/mg) (Blethen & Kaplan, 1968). The specific activities of the tarantula *Dugesiella hentzi* (58.5 and 75 U/mg) (Blethen & Kaplan, 1968) and activity of the recombinant AK from the insect *Ctenocephalides felis* 75 U/mg (Werr, Cramer & Ilg, 2009) were much higher. The variations seen may be due to differences in experimental techniques.

The kinetic parameters determined for the substrate L-arginine were $K_m$ 1.7 mM and $V_{max}$ 27.8 µmol min$^{-1}$ mg protein$^{-1}$; within the range described for other arthropods. The horseshoe crab *Limulus polyphemus* (Blethen, 1972), the scorpion *Palameus phipsoni* (Arjunwadkar & Raghupathi Rami Reddy, 1985) and tarantula *Dugesiella hentzi* (Blethen & Kaplan, 1968) had a $K_m$ between 0.08 and 0.67 mM; in insects *Cissites cephalotes* (Tanaka et al., 2007), *Locusta migratoria* (Ilg & Werr, 2012; Wu et al., 2007), *Periplaneta americana* (Brown & Grossman, 2004), $K_m$ was between 0.49–1.0 mM; and in the crustacean *Litopenaeus vannamei* (Lopez-Zavala et al., 2016), *Euphausia superba* (Si et al., 2014), the $K_m$ was 0.3 (Table 1).

The value of $k_{cat}$ calculated for PpAK (as shown in Table 1) is within the range reported in arthropods, which seems to be highly variable from 1.3 s$^{-1}$ in *P. americana* (Brown & Grossman, 2004) to 159 s$^{-1}$ for recombinant AK and AK from *L. migratoria* leg muscle (Wu et al., 2007). The PpAK catalytic efficiency for arginine ($k_{cat}/K_m$ Arg) was 44.1 s$^{-1}$ mM$^{-1}$, well within the range reported in arthropods: 2.01 s$^{-1}$ mM$^{-1}$ for *C. cephalotes* (Tanaka et al., 2007) to 282.7 s$^{-1}$ mM$^{-1}$ for *Euphausia superba* (Si et al., 2014).

Structural characterization has been focused in both free-ligand and fully occupied active site of AK, that are called open and closed conformation, respectively (Fernandez et al., 2007; Pruett et al., 2003; Zhou et al., 1998). Only, a few reports show structural details of binding of arginine in a binary complex, which is a crucial step during catalysis (Lopez-Zavala et al., 2013; Wang et al., 2015). To date, horseshoe crab crystal structure has been reported for Chelicerate, which is comprised of more than 75,000 species (Strong & Ellington, 1995). Apo-PpAK structure was found in the open conformation as reported for other phosphagen kinases in the substrate-free form (Fig. 3). In this conformation, the loop comprised from amino acid residue 310–320 is disordered and therefore, was not present in final experimental model. Conversely, this loop is well ordered and serves as a lid to the active site when both substrates are the catalytic closed conformation.

The guanidino group binding loop (residues 59–64 ENLDSG for PpAK) in the phosphagen kinases are invariant and are located in the N-terminal domain. The length of
this loop is related to the size of the substrate (Azzi et al., 2004). PpAK has a well-conserved guanidino group binding loop (see Fig. 2) that is indicative of its preferred substrate, arginine. The binary complex of PpAK and L-Arg has a similar conformation compared to the unbound structure (Fig. 3). This suggest that the enzyme has a conformation ready to bind the guanidino group and the loop sequence is optimal for ligand binding. In more detail, there are precise hydrogen bonds with residues of the guanidino group binding loop (Fig. 3). Specifically, arginine carboxylate groups are in contact with the amide backbone of Gly64, Val65 and Gly66; also, the hydroxyl group of Tyr68 stabilizes the arginine amide. Another highly conserved residue is Cys271, which has been implicated in catalysis (Buechter et al., 1992; Marletta & Kenyon, 1979). In PpAK this residue interacts via hydrogen bonds with N(ε) guanidino group. Glu225 has been proposed as one of two bases that assist acid–base catalysis mechanisms in phosphagen kinases by substrate alignment during phosphate group transference (Pruett et al., 2003). Litopenaeus vannamei crystal structure in arginine binary complex (Lv AK-Arg; PDB entry 4BHL) shows that Glu225 stabilizes the arginine guanidino group end by two ionic bonds in a pre-stabilized catalytic intermediate (Lopez-Zavala et al., 2013). In the PpAK-Arg structure we found that Glu225 is positioned in a similar way, making only one salt-bridge with NH2(ω)-1 of the arginine guanidino group end. This small local conformational change was also observed in LvAK-Arg structure and it is postulated to lead to a productive active site in the closed conformation (López-Zavala et al., 2014; Zhou et al., 1998).

In spiders, the resting metabolic rate is much lower compared to other animals (Anderson, 1970), about half of the expected rate for other poikilotherm (Schmitz, 2016). Fast movements are limited by depletion of phosphagens, accumulation of lactate and a limited aerobic metabolism (Prestwich, 1988), besides a low number of mitochondrias in the muscle and the respiratory system (Figueroa et al., 2010). Biochemical studies realized in spiders indicate low activity of mitochondrial citrate synthase, glutamate dehydrogenase and isocitrate synthase, suggesting a low aerobic metabolism (Linzen & Gallowitz, 1975). Also in spiders, the structure of muscle fiber suggests that most of the fiber is comprised of the contractile apparatus, with limited energy production. This in contrast to the flight muscle of Locusta migratoria, where 40% of the volume in muscle fibers comes from mitochondria (Linzen & Gallowitz, 1975).

CONCLUSION

In conclusion, we successfully studied molecular, enzymatic and structural aspects of the arginine kinase from the spider P. pythagoricus. Our resolved structures are the first views of AK in arachnids. Both ligand-free and arginine binary complex were found in the open conformation. The PpAK guanidino group binding site appears to be pre-organized to accept the arginine substrate since it overlaps with the unbound crystal structure. Binding of the second substrate is needed for the conformational changes to occur in the AK structure of other invertebrates. Finally, these results contribute to increase the basic knowledge of spider bioenergetics.
ADDITIONAL INFORMATION AND DECLARATIONS

Funding
This work was supported by CONACYT (Mexico’s National Council for Science and Technology) Binational Mexico-Argentina grant C0005-2012-189333 to R Sotelo-Mundo and CONICET (Argentina’s National Council for Scientific and Technical Research) grant COOP-CTES-D680 to CF García. CF García also received grants from: Agencia Nacional de Promoción Científica y Tecnológica PICT 2014-2580, CONICET PIP 0515 and UNLP for 11/N667. R Sotelo-Mundo was supported by CONACYT grant INFR-2014-01-225455 for the X-ray diffractometer. This work was also supported by the NSLS-SSRL user transition program supported jointly by the LSBR, NSLSII and SMB, SSRL under NIH-NIGMS grants P41GM111244 and P41GM103393, and DOE BER contracts DE-SC0012704. SSRL is operated under DOE BES contract DE-AC02-76SF00515. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures
The following grant information was disclosed by the authors:
CONACYT (Mexico’s National Council for Science and Technology) Binational Mexico-Argentina: C0005-2012-189333.
Agencia Nacional de Promoción Científica y Tecnológica: PICT 2014-2580, CONICET PIP 0515.
UNLP: 11/N667.
CONACYT: INFR-2014-01-225455.
NSLS-SSRL, LSBR, NSLSII and SMB, SSRL under NIH-NIGMS and DOE-BES and -BER: P41GM111244, P41GM103393, DE-SC0012704, DE-AC02-76SF00515.

Competing Interests
The author declares there are no competing interests.

Author Contributions
• Aldana Laino, Jesus S. Carrasco-Miranda and Carlos Fernando García conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
• Alonso A. Lopez-Zavala conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
• Karina D. Garcia-Orozco conceived and designed the experiments, performed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.
• Marianela Santana performed the experiments, analyzed the data, wrote the paper, reviewed drafts of the paper.
• Vivian Stojanoff analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, reviewed drafts of the paper.
• Rogerio R. Sotelo-Mundo conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables.

DNA Deposition
The following information was supplied regarding the deposition of DNA sequences:

The nucleotide and deduced amino acid sequence were deposited at GenBank under the accession number MF001441.

Data Availability
The following information was supplied regarding data availability:

The coordinates and structure factors deposited at PDB are included as Supplemental Files. The assigned PDB codes are 5U8E and 5U92.

Supplemental Information
Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.3787#supplemental-information.

REFERENCES


kinase as a major cockroach allergen among Thai patients with major cockroach allergies. *Environmental Health Perspectives* **114**:875–880 DOI 10.1289/ehp.8650.


