The Substrate-free and -bound Crystal Structures of the Duplicated Taurocyamine Kinase from the Human Parasite Schistosoma mansoni

Romain Merceron, Ayman M Awama, Roland Montserret, Olivier Marcillat, Patrice Gouet

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 290, NO. 20, pp. 12951–12963, May 15, 2015

The Substrate-free and -bound Crystal Structures of the Duplicated Taurocyamine Kinase from the Human Parasite *Schistosoma mansoni*[*][S]

Received for publication, December 3, 2014, and in revised form, April 1, 2015 Published, JBC Papers in Press, April 2, 2015, DOI 10.1074/jbc.M114.628909

Romain Merceron[‡], **Ayman M. Awama**[§], **Roland Montserret**[‡], **Olivier Marcillat**[§1], **and Patrice Gouet**[‡2]

From the [‡]*Institut de Biologie et Chimie des Protéines, BMSSI-IBCP, UMR 5086 CNRS Université Lyon 1, 7, Passage du Vercors, 69367 Lyon Cedex 07, France and the* [§]*Institut de Chimie et Biochimie Moléculaire et Supramoléculaire, UMR 5246 CNRS Université Lyon 1, 69622 Villeurbanne, France*

Background: Trematode taurocyamine kinases are contiguous dimers of unknown structure.
Results: The first reported crystal structure of taurocyamine kinase displays an original bilobal arrangement compared with true dimeric phosphagen kinases.
Conclusion: Each lobe is capable of enzymatic activity and substrate binding in a non-mutually exclusive manner.
Significance: This structure can serve as a tool for the rational design of anti-schistosomiasis drugs.

The taurocyamine kinase from the blood fluke *Schistosoma mansoni* (SmTK) belongs to the phosphagen kinase (PK) family and catalyzes the reversible Mg^{2+}-dependent transfer of a phosphoryl group between ATP and taurocyamine. SmTK is derived from gene duplication, as are all known trematode TKs. Our crystallographic study of SmTK reveals the first atomic structure of both a TK and a PK with a bilobal structure. The two unliganded lobes present a canonical open conformation and interact via their respective C- and N-terminal domains at a helix-mediated interface. This spatial arrangement differs from that observed in true dimeric PKs, in which both N-terminal domains make contact. Our structures of SmTK complexed with taurocyamine or L-arginine compounds explain the mechanism by which an arginine residue of the phosphagen specificity loop is crucial for substrate specificity. An SmTK crystal was soaked with the dead end transition state analog (TSA) components taurocyamine-NO_3^{2-}-MgADP. One SmTK monomer was observed with two bound TSAs and an asymmetric conformation, with the first lobe semiclosed and the second closed. However, isothermal titration calorimetry and enzyme kinetics experiments showed that the two lobes function independently. A small angle x-ray scattering model of SmTK-TSA in solution with two closed active sites was generated.

Guanidino kinases (EC 2.7.3), also called phosphagen kinases (PKs),[3] constitute a large family of enzymes that catalyze the Mg^{2+}-dependent phosphoryl transfer from ATP to various guanidino substrates to yield ADP and an *N*-phosphorylated guanidino product termed phosphagen. Because PKs can also efficiently catalyze the reverse reaction, they allow the establishment of a phosphagen reservoir that is readily available for ATP regeneration and maintain the ATP/ADP balance in cells with high and fluctuating energy requirements (1). PKs are widely distributed throughout the animal kingdom and are found in unicellular organisms, protists, and bacteria, although their presence in these organisms has seldom been described (2).

Depending on the organism studied, different guanidylic acceptors can be found to be associated with specific PKs. The best known of these is creatine, the substrate of the creatine kinase isoenzymes found in many multicellular organisms and the sole PK identified in vertebrates. Creatine kinases can be produced from isoform gene groups classified as cytoplasmic (CK), mitochondrial, or flagellar. Arginine kinases (AKs) are widely distributed in invertebrates and have been proposed to be the most closely related to a PK-like ancestor (3). Other PKs, such as taurocyamine kinase (TK), glycocyamine kinase (GK), or lombricine kinase, have been observed in annelids and allied groups (1, 4). PKs are generally 300–400 amino acids long and share extensive sequence similarity with several characteristic sequence signatures and a typical molecular mass of ~45 kDa (1). In addition to the use of different guanidino substrates, evolution within this family of enzymes also results in quaternary structures with increasing complexity. Although AKs are largely monomeric proteins, dimeric AKs can be found; they most likely evolved from a CK-like ancestor (3, 5). CKs are obligate dimers (6), and mitochondrial creatine kinases are found mainly as octamers bound to the outer face of the mitochondrial inner membrane, but these proteins can reversibly dissociate into active dimers (7–10). In addition to these noncovalently associated multisubunit proteins, duplicated AKs

* The crystallization work in the study benefited from the IBCP/UMS 3444/US8 crystallogenesis facilities of SFR Biosciences. This work was supported by the European Community's Seventh Framework Program (FP7/2007–2013) under Grant Agreement 283570 (BioStruct-X).
The atomic coordinates and structure factors (codes 4WO8, 4WOD, and 4WOE) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[S] This article contains supplemental Video 1.
[1] To whom correspondence for enzymatic studies may be addressed. E-mail: olivier.marcillat@univ-lyon1.fr.
[2] To whom correspondence for structural studies may be addressed. E-mail: p.gouet@ibcp.fr.
[3] The abbreviations used are: PK, phosphagen kinase; CK, cytoplasmic creatine kinase; AK, arginine kinase; TK, taurocyamine kinase; TSA, transition state analog; AMPPNP, 5′-adenylyl-β,γ-imidodiphosphate; SmTK, *S. mansoni* TK; LvAK, *L. vannamei* AK; SAXS, small angle x-ray scattering; ITC, isothermal titration calorimetry; PDB, Protein Data Bank.

(11, 12), TKs (13), and triplicated flagellar creatine kinases (6) have been described. These contiguous PKs have most likely originated through the duplication or triplication of a primordial gene and, to date, have been described without structural insight.

All known PK structures display a conserved topology, with an N-terminal α-helix bundle domain connected to a large $\alpha+\beta$ C-terminal domain, as observed in the crystal structures of CKs and mitochondrial creatine kinases (9, 10, 14–19), AKs (5, 20–26), GK (27), and lombricine kinase (28). In the substrate-free enzyme, the active site is open and stretches as a groove on the concave face of the structure. The N-terminal domain contains a variable 7–13-amino acid phosphagen specificity loop, the length of which is inversely correlated to the size of the guanidino substrate (29). The C-terminal domain is mainly involved in nucleotide binding and has a consequent inward bending motion when the PK is complexed with an abortive transition state analog (TSA) consisting of the guanidino substrate, the MgADP product, and a trigonal planar nitrate anion that mimics the transferred phosphoryl group (20). The crystal structures of CK-, AK-, and GK-TSA as well as NMR and mass spectrometry studies show that this structural rearrangement is accompanied by the stabilization of a C-terminal flexible loop that shields the active site from the solvent (14, 19, 20, 23, 24, 27, 30, 31). The PK-TSA structure is closed, and the bound compounds are locked in proper alignment for the in-line phosphoryl transfer (23, 26, 32–34). An invariant cysteine ($Cys^{268/631}$ in SmTK) and an invariant glutamate ($Glu^{222/585}$ in SmTK) contribute to the stabilization of the guanidylic acceptor and maximize the catalytic reaction (20, 35).

Although the advantage conferred by quaternary structures of increasing complexity remains elusive, it has been suggested that the evolution of PKs from a monomeric to an oligomeric or a contiguous form is linked to cooperative regulation (5, 6). Hence, for dimeric CKs, a negative intersubunit cooperativity mechanism has been inferred from asymmetric structures of both unliganded and liganded crystal forms (14, 19). Similar structural features were reported for a dimeric AK co-crystallized with the pretransition state analog arginine-AMPPNP, which achieved a closed liganded form in a single monomer (5). However, a dimeric GK with an asymmetric structure in the substrate-free form and a doubly closed structure in the GK-TSA complex have also recently been described (27), ruling out the hypothesis of a systematic coupling between asymmetrical unliganded dimers and catalytically interdependent monomers. In addition, several experiments performed on muscle CK did not provide any evidence for a functional asymmetry of the two identical subunits of the enzyme (30, 31, 36, 37).

Schistosoma mansoni is a trematode flatworm and the causative agent of schistosomiasis, which affects over 240 million people worldwide. A protein produced by this organism was classified as a TK (EC 2.7.3.4) based on sequence similarity and preliminary enzymatic assays (38). The *S. mansoni* TK (SmTK) is developmentally regulated and abundantly produced in the motile tail during the cercarial stage of the parasite's life cycle (39, 40). Similar to other known trematode TKs, SmTK is a contiguous dimer composed of 716 amino acids (Fig. 1*A*) with a structure that is thought to have evolved through duplication of a molluscan primordial AK gene (41). Each of the two fused sequence elements is catalytically active, as shown for *Schistosoma japonicum*, *Paragonimus westermani*, and *Clonorchis sinensis* (41–44). By contrast, the enzymes are nearly inactive against arginine and creatine substrates, as previously shown for SmTK (40).

We report herein a comprehensive structural and biochemical study of the SmTK contiguous dimer. We solved the crystal structure of the substrate-free form at 2.2 Å resolution. The structural bases of the catalytic mechanism were investigated by determining the crystal structures of SmTK in complex with TSA components at 2.3 Å resolution (SmTK-TSA) and in complex with L-arginine at 1.9 Å resolution (SmTK-Arg). Small angle x-ray scattering (SAXS) data were collected for the free and substrate-bound enzymes to characterize structural rearrangements in solution. Enzymatic and isothermal titration calorimetry (ITC) experiments were performed on the wild-type enzyme and invariant cysteine mutants to investigate a cooperative mechanism in the duplicated structure.

EXPERIMENTAL PROCEDURES

Construction, Overproduction, and Purification of Enzymes—Wild-type SmTK was overproduced and purified from *Escherichia coli* BL21(DE3) cells (Novagen) as previously reported (38). The invariant cysteine mutants C268S, C631S, and C268S/C631S (doubly mutated) were obtained by site-directed mutagenesis following the QuikChange® site-directed mutagenesis protocol (Stratagene). The mutations were generated using two couples of synthetic oligonucleotides (C268S, 5′-CCCAAATTAGATGGA**G**AGAATGTGATAAATCCTAAACG-3′ (reverse) and 5′-CGTTTAGGATTTATCACATTCT**C**TCCATCTAATTTGGG-3′ (forward); C631S, 5′-CTAAGTTTGATGGA**G**AGCAAGTTATATAACCATATTTATC-3′ (reverse) and 5′-GATAAATATGGTTATATAACTTGCT**C**TCCATCAAACTTAG-3′ (forward)) (the mutations are shown in boldface type). The linear amplified mutant plasmids were transformed into chemically competent *E. coli* DH5α cells (Novagen), and the sequences were confirmed by DNA sequencing. For expression of the genes, the plasmids were transformed into *E. coli* BL21(DE3) cells. Expression and purification of mutant enzymes followed the wild-type SmTK protocol. In short, after disruption of the *E. coli* cells by sonication, the supernatants were loaded onto a Blue Sepharose 6 Fast Flow column (GE Healthcare), and fractions containing the protein of interest were pooled and loaded onto a MonoQ-HR 5/5 column (GE Healthcare). The purity was verified by the presence of a single band on Coomassie Blue-stained SDS-polyacrylamide gels. Concentrations of the purified proteins were determined spectrophotometrically using the theoretical molar extinction coefficient at 280 nm of $6.745 \times 10^4\ M^{-1}\cdot cm^{-1}$ as calculated for the wild-type enzyme using ProtParam (ExPASy server), which gave concentration values similar to those determined using the Pierce 660-nm assay (Thermo Scientific). The purified proteins were concentrated to 1 mg/ml in the final buffer (20 mM Tris-Cl, pH 8.5) by ultrafiltration using Vivaspin devices (30 kDa molecular mass cut-off; Sartorius) and then flash-frozen in liquid nitrogen for storage at 193 K.

Circular Dichroism Spectroscopy (CD)—Far UV CD spectra were recorded for wild-type SmTK and mutant enzymes using

a Chirascan spectrometer (Applied Photophysics). For all samples, experiments were performed at room temperature with 150 μl of protein solution at a final concentration of 0.1 mg/ml in 20 mM sodium phosphate buffer (pH 8.0). The measurements were made in a 0.1-cm quartz cuvette between 190 and 260 nm with an increment of 0.2 nm and 1-s integration time. Spectra were then baseline-corrected, smoothed, and converted to mean residue molar ellipticity (θ). Deconvolution of spectra was carried out on the DICHROWEB server (45).

ITC—Titrations of wild-type SmTK and mutant enzymes with ADP in the presence of magnesium were carried out with an ITC_{200} microcalorimeter (MicroCal). Measurements were performed at 303 K with a first injection of 0.5 μl followed by 19 successive injections of 2 μl of ligand during 8- and 120-s intervals at a stirring speed of 1,000 rpm in the calorimetric cell. Typically, the cell contained the enzyme solution at a concentration of 0.02 mM in 20 mM Tris-Cl buffer, pH 8.5. To study the SmTK-TSA complex, the cell was supplemented with 5 mM $Mg(NO_3)_2$ and 5 mM taurocyamine in addition to the enzyme. The binding of ADP in presence of magnesium was also examined in absence of some of the TSA components by omitting taurocyamine and replacing $Mg(NO_3)_2$ by 5 mM $Mg(Ac)_2$ in the experiment. The syringe contained the ADP usually at a concentration of 0.3–0.4 mM as determined by UV light absorbance using the 260-nm ADP molar extinction coefficient of $1.54 \times 10^4\ M^{-1}{\cdot}cm^{-1}$. The injection syringe was supplemented with taurocyamine and ion cofactors according to the type of experiment performed. The enzyme solution used for ITC measurements was concentrated with an ultrafiltration device (cutoff 30 kDa). The ultrafiltrate solution was used to prepare the ligand solution in order to minimize heats of dilution. Blank titrations were done in the absence of enzyme to record the heat of dilution. The data corrected for the heat of dilution were best fitted to a single set of identical independent binding sites model with the Origin version 7.0 software to obtain the binding parameters K_a (association constant) from which the K_D (dissociation constant, $1/K_a$) was obtained, ΔH (enthalpy), and n (stoichiometry).

Crystallization—We have previously published the crystallogenesis of SmTK and solved the phase problem by molecular replacement (38). However, the crystallographic refinement to 2.8 Å resolution led to the observation of a bound MES buffer molecule in the active site. In an effort to obtain an unliganded structure, crystallization conditions were rescreened at 293 K by the sitting drop vapor diffusion method using new commercial kits. An initial hit was obtained in the condition A3 of the JCSG+ screen (200 mM diammonium citrate, pH 5.0, 20% (w/v) polyethylene glycol 3350) and optimized. New SmTK crystals were obtained by equilibrating drops of protein (concentrated at 5 mg/ml in 20 mM Tris-Cl, pH 8.5) with the reservoir solution (200 mM diammonium tartrate, pH 5.4, 20% (w/v) polyethylene glycol 3350, 20% (v/v) ethylene glycol) with a 2:1 or 1:1 protein/solution volume ratio. Crystals grew within 2 weeks at 293 K to reach maximum dimensions of $200 \times 50 \times 50\ \mu m^3$. They were isomorphous to the previous ones with one monomer in the asymmetric unit and diffract x-rays to about 2 Å resolution. The new crystallization solution contained a cryoprotectant (20% (v/v) ethylene glycol), and mounted crystals were directly flash-frozen in liquid nitrogen before data collection. The L-arginine complex (SmTK-Arg) was obtained by the co-crystallization method by equilibrating a protein drop (same conditions as above) mixed with a 2:1 protein/solution volume ratio with the new crystallization solution supplemented with 5 mM L-arginine. Different methods (crystal soaking or co-crystallization) and substrate concentrations (5–100 mM) were assayed to obtain either the SmTK-TSA complex or the pre-transition state analog complex (pre-TSA; SmTK-taurocyamine-MgAMPPNP). Finally, the structure of the SmTK-TSA complex with a closed active site was obtained from a crystal soaked for 5 min in the optimized crystallization solution supplemented with 100 mM taurocyamine, 20 mM ADP, 50 mM $Mg(NO_3)_2$, and 1% (v/v) DMSO to improve ligand solubility. This soaking resulted in the doubling of the *a* unit cell dimension (see Table 1).

X-ray Diffraction Experiments and Structure Determination—Synchrotron data were collected at 100 K on the MX beamlines at the European Synchrotron Radiation Facility in Grenoble, France, and on the PXIII beamline at the Swiss Light Source in Villigen, Switzerland. The XDS and XSCALE programs were used for data reduction and scaling (46). Five percent of the randomly selected reflections were kept apart for cross-validation, and the phasing problem was solved by molecular replacement as described previously (38). The substrate-free structure was further refined by several cycles of rigid body in PHENIX (47) followed by positional and *B*-factor refinement. The model was manually improved during the course of the crystallographic refinement using the graphic program COOT (48). The SmTK-Arg structure was refined using a similar protocol. The SmTK-TSA structure was solved by molecular replacement using the program PHASER (49) and substrate-free SmTK as a search model. Because the taurocyamine structure is absent from the Protein Data Bank (PDB), a model was generated with the program eLBOW (50). All bound ligands were positioned in $F_o - F_c$ difference maps with COOT. The flexible loop at 671–680 was rarely observed in electron density maps and was not included in the final models of SmTK chain A, SmTK-Arg chain A, and SmTK-TSA chain B. The equivalent loop, 308–315, located in the first of the duplicated sequence elements, was not included in SmTK-TSA chain B. The quality of the final crystal structures was assessed with MOLPROBITY (51) prior to deposition at the PDB under the codes 4WO8 (SmTK), 4WOD (SmTK-Arg), and 4WOE (SmTK-TSA). Molecular images were generated with PyMOL (Schrödinger, LLC, New York). Secondary structure elements were assigned with DSSP (52). Figures of sequence alignments were prepared using ESPript/ENDscript (53). Data collection and final refinement statistics are presented in Table 1.

SAXS—All scattering data were collected at the BM29 European Synchrotron Radiation Facility beamline. For every experimental condition (including buffers), the temperature was maintained at 277 K, and 10 frames (with 2-s time exposure) were collected to monitor radiation damages before being merged. The buffer scattering was subtracted for each condition using PRIMUS (54), and the resulting curve was analyzed. The radii of gyration (R_g) and the forward scattering intensity $I(0)$ of all samples were extracted by the Guinier approximation

using AutoRg (55) and considering angles where s is $<1.3/R_g$ with the same program. For cross-validation, R_g and $I(0)$ were also obtained from the pairwise distribution functions $P(r)$ using GNOM (56). The maximum dimension of the particle (D_{max}) was calculated by restraining the $P(r)$ function to 0 at r_{max} and varying r_{max} over a wide range to find the value ($r_{max} = D_{max}$) that gave a plausible $P(r)$ function. The molecular weight of SmTK was estimated using the forward scattering $I(0)$ of SmTK sample and that of the standard protein BSA as described previously (57). The predicted scattering profiles of crystal structures were calculated using the program CRYSOL for comparison with SAXS data (58). Low resolution shape envelopes were generated with the *ab initio* bead-modeling program DAMMIF (59) by performing 10 independent reconstructions, aligning all with the most probable model with DAMSEL and DAMSUP, averaging models with DAMAVER, and adjusting to correspond with the experimentally determined excluded volume with DAMFILT (60). The final bead model was superimposed upon the x-ray crystal structure coordinates with SUPCOMB (61). For SAXS experiments, SmTK in the concentration range of 0.5–4.2 mg/ml was analyzed in the presence or absence of substrates (taurocyamine, 10 mM taurocyamine; $MgCl_2$, 10 mM $MgCl_2$; $Mg(NO_3)_2$, 10 mM $Mg(NO_3)_2$; MgADP, 10 mM $MgCl_2$, 10 mM ADP; MgATP, 10 mM $MgCl_2$, 10 mM ATP; ADP, 10 mM ADP; TSA, 10 mM taurocyamine, 10 mM $Mg(NO_3)_2$, 10 mM ADP; pre-TSA, 10 mM taurocyamine, 10 mM $MgCl_2$, 5 mM AMPPNP). The protein (in 20 mM Tris-Cl, pH 8.5) is typically previously filtered through a 0.02-μm pore size filter (Nalgene) before being concentrated with an ultrafiltration device (30 kDa molecular mass cut-off).

Enzyme Kinetics—Kinetic analysis of the forward reaction (*N*-phosphotaurocyamine synthesis) catalyzed by wild-type and mutant enzymes was achieved using a coupled pyruvate kinase/L-lactic acid dehydrogenase assay, where production of ADP is coupled to the oxidation of NADH. The 300-μl reaction volume contained 2.5 units of pyruvate kinase, 2.5 units of L-lactic acid dehydrogenase, 0.6 mM phosphoenolpyruvate, 0.25 mM NADH, and 5 mM KCl in 100 mM Tris-acetate, pH 8.2, 10 mM magnesium acetate, and 0.5 mM β-mercaptoethanol. Preliminary experiments ensured that the concentration of SmTK was the rate-limiting factor. Enzymatic reactions were followed by the decrease in absorbance at 340 nm of NADH on a Tecan Infinite M1000 spectrophotometer thermostated at 303 K. Samples where preincubated at this temperature for 30 min before adding the enzyme. Apparent kinetic constants were determined under pseudo-single substrate conditions by varying one substrate (taurocyamine or ATP, maximal tested concentration 5 mM) while holding the remaining substrate at constant and saturating concentration (5 mM). Apparent kinetic parameters ($k_{cat(app)}$ and $K_{m(app)}$) for wild-type enzyme and single reactive-cysteine mutants were determined from a nonlinear least-squares best fit of the data to the Michaelis-Menten equation using the program GraphPad Prism version 6.05 (GraphPad Software, San Diego, CA). The apparent catalytic constant for double mutant was determined at single and saturating concentrations of ATP and taurocyamine (5 mM).

Taurocyamine Synthesis—Taurocyamine was synthesized from *S*-methylisothiourea and taurine as described (62). 28.8 g of taurine (Sigma) solubilized in 100 ml of ammonium hydroxide (Sigma) were allowed to react with 48 g of *S*-methylisothiourea (Fluka) for 20 h at 333 K. Crystallized taurocyamine was then filtered and purified twice by crystallization with ice-cold water.

RESULTS

Solution Characterization and Crystal Structure of Substrate-free SmTK—The duplicated structure of SmTK was investigated by combining SAXS and x-ray crystallography approaches (see "Experimental Procedures"). This enzyme is monomeric in solution, as assessed by the estimation of its molecular mass by SAXS, which indicated values close to the expected 80 kDa at the maximal protein concentration tested (4.2 mg/ml). Similarly, a PISA analysis (63) detected no biological interface in the crystal packing of the 2.2 Å resolution crystal structure. The calculated SAXS envelope has a slightly twisted U shape into which the atomic structure of SmTK fits perfectly, with a normalized spatial discrepancy value of 0.94 (Fig. 1*B*). Moreover, deconvolution into secondary structural components of the circular dichroism spectra recorded for the wild-type enzyme agrees well with the percentages obtained from a DSSP analysis of the crystal structure (data not shown). Each of the fused sequence elements corresponds to a complete PK domain. The first (D1, residues 1–363) fills one half of the U, and the second (D2, residues 364–716) fills the other half. The active sites of D1 and D2 are located in the cradle of the U and are very distinct, with a separation of ~40 Å. The two PK domains are very similar to each other, with a sequence identity of 53% and a root mean square deviation of 1 Å for all Cα pairs after superimposition. These domains will be hereafter called "lobes" in reference to the terminology used for other proteins with duplicated domains, such as transferrins (64).

Each SmTK lobe displays a classical PK topology, with an all α-helix, 98-residue-long N-terminal domain connected to a large C-terminal domain composed of an eight-stranded antiparallel β-sheet flanked by seven α-helices and a β-hairpin (Fig. 1). According to a search for structural homologs performed with DALI (65), these lobes are most similar to substrate-free AKs. The highest similarity score ($Z = 46$) was obtained between lobe D1 and the monomeric AK of *Litopenaeus vannamei* (LvAK) (25), with 47% sequence identity and a root mean square deviation of 1.5 Å for 339 Cα pairs.

The main structural singularity of SmTK resides in the spatial arrangement between the two fused lobes. The C-terminal domain of D1 interacts with the N-terminal domain of D2 via an interface mediated by helix α-12 of D1, which buries an 840-Å^2 surface per lobe as calculated by PISA (Fig. 1*B*). Residues 289–292 and 295–296 of the C-terminal domain of D1 (α-10′, α-11 region) are in contact with a few residues of the N- and C-terminal domains of D2 at the interface (Fig. 1*A*). By contrast, the intermolecular interface of true dimeric PKs involves the N- and C-terminal domains of two monomers related by a pseudo-2-fold rotation axis and buries a larger surface of at least 1,000 Å^2 per monomer.

The Key Loops in SmTK—The phosphagen specificity loop (residues 56–64 in D1 and 419–427 in D2) shows similar conformations in the two lobes, despite the low sequence identity

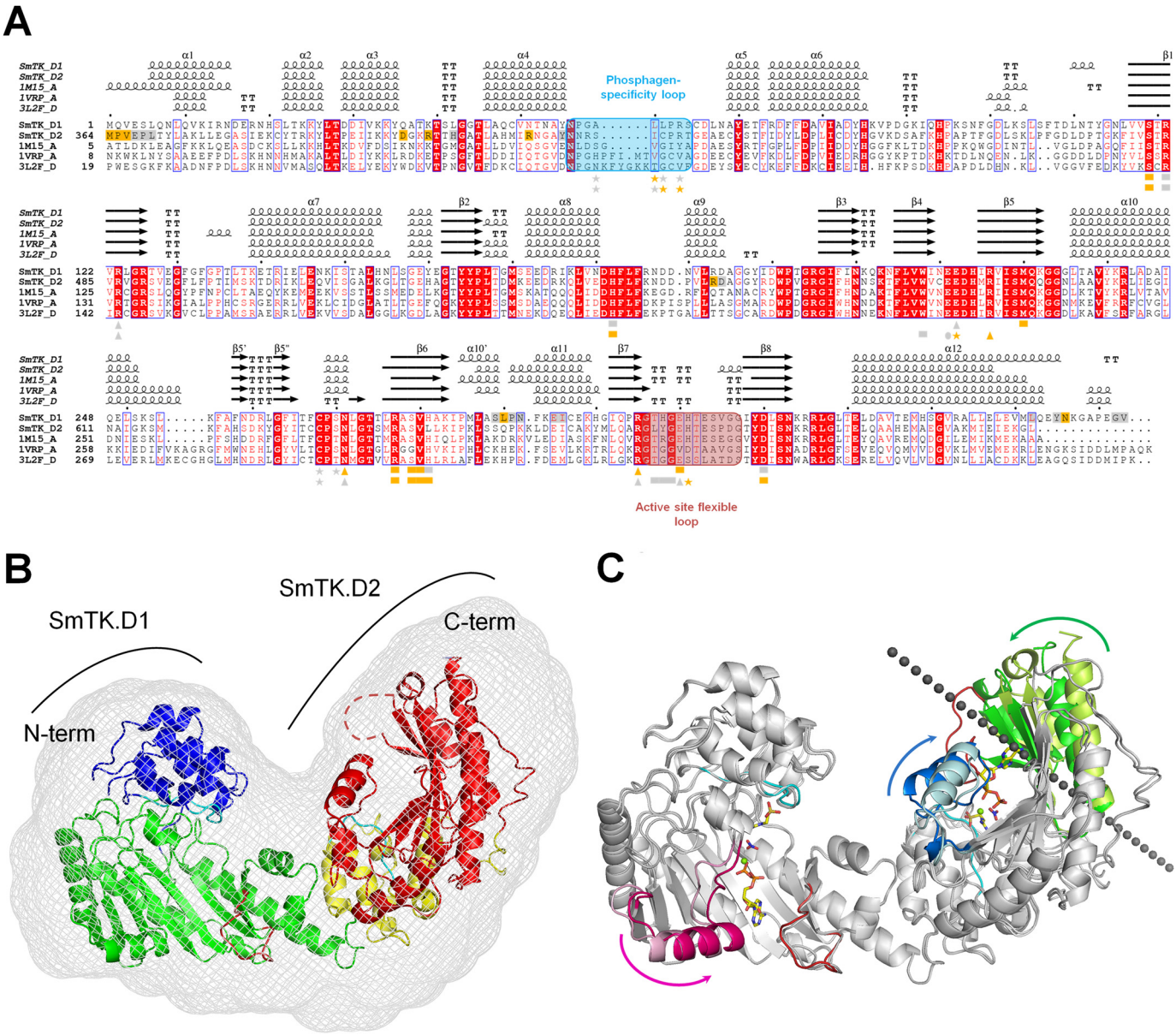


FIGURE 1. **Sequence alignment of SmTK and overall structure.** *A*, multiple sequence alignment and secondary structure elements of the contiguous lobes D1 and D2 of SmTK with the AK from *Limulus polyphemus* (PDB code 1M15), the CK from *Torpedo californica* (PDB code 1VRP), and the lombricine kinase from *Namalycastis* sp. ST01 (PDB code 3L2F). The phosphagen specificity and active site flexible loops are highlighted by *cyan* and *red transparent frames*, respectively. *Stars*, *squares*, *triangles*, and *circles* indicate residues of D1 and D2 that are in contact with taurocyamine, ADP, nitrate, and magnesium, respectively. Contact distances of <3.2 Å are shown in *orange*; contacts in the range of 3.2–4 Å are shown in *gray*. *B*, *ribbon diagram* of SmTK fitted in the SAXS envelope (*gray mesh*). The N- and C-terminal domains of D1 are *colored* in *blue* and *green*, respectively, and the N- and C-terminal domains of D2 are *yellow* and *red*, respectively. The phosphagen specificity loop is shown in *cyan* in both D1 and D2. The flexible loop is in *red* in D1 and shown with a *dotted line* in D2 because it is not observed in the electronic density maps. *C*, *schematic* superimposition between unliganded SmTK and liganded monomer A of SmTK-TSA. The conformational variant regions identified with the program DynDom are depicted in *pale color* for SmTK and *darkened color* for SmTK-TSA (*i.e.* the C-terminal part of D2 (residues 642–716) is shown in *green* with its flexible loop in *red*, and the α-8 helix of D1 and D2 is shown in *red* and *blue*, respectively). The hinge axis for the C-terminal part of D2 movement from SmTK to SmTK-TSA and detected by DynDom is represented by non-bonded *black spheres*. The bound ligands are shown as *sticks*, and the Mg^{2+} ion is depicted as a *green sphere*.

(Fig. 1, *A* and *B*). This short loop is characterized by a 6-amino acid deletion in both SmTK lobes compared with CKs (Fig. 1*A*). This represents the largest gap reported thus far in the PK family and is observed in all known TKs from trematodes (44). Segment 671–680 of the C-terminal region of D2 is not visible in electron density maps (Fig. 1*B*). This loop is often described as highly flexible in substrate-free PKs (23, 30), whereas it is implicated in the stabilization and optimal alignment of the substrates in all known PK-TSA complexes (14, 19, 20, 23, 24, 27, 30, 31). The loop is delineated by two glycine residues (Gly^{670}–Gly^{681}) that account for its high flexibility. The equivalent segment in D1 (Gly^{306}–Gly^{317}) is visible in electron density maps due to crystallographic contacts with a neighboring monomer and adopts an outward conformation with respect to the active site (Figs. 1*B* and 5*A*).

TSA Binding—Crystals of SmTK were soaked with TSA components, resulting in the transition from one to two monomers in the asymmetric unit of the SmTK-TSA crystal (Table 1). Our

TABLE 1

Data collection and refinement statistics

Values in parentheses are for the high-resolution shell. ESRF, European Synchrotron Radiation Facility; SLS, Swiss Light Source; RMSD, root mean square deviation.

	SmTK	SmTK-Arg	SmTK-TSA
Data collection			
Synchrotron beamline	PXIII, SLS	ID29, ESRF	ID23–2, ESRF
Wavelength (Å)	1.00	0.97	0.87
Space group	$P2_1$	$P2_1$	$P2_1$
Unit cell parameters (Å)	$a = 52.8$, $b = 122.0$, $c = 63.0$	$a = 51.9$, $b = 123.1$, $c = 62.7$	$a = 107.7$, $b = 122.2$, $c = 63.1$
Unit cell parameter (degrees)	$\beta = 108.1$	$\beta = 108.1$	$\beta = 106.2$
Resolution limit (Å)	2.2 (2.26–2.2)	1.9 (1.95–1.9)	2.3 (2.36–2.3)
No. of measurements	118,958 (6,118)	246,668 (15,545)	215,036 (14,434)
Unique reflections	38,030 (2,743)	58,275 (4,190)	67,921 (4,914)
Completeness (%)	98.7 (97.0)	98.9 (96.2)	98.6 (96.4)
R_{meas} $(I)^a$ (%)	9.2 (63.6)	12.2 (53.2)	15.4 (54.2)
$CC(1/2)^b$ (%)	99.6 (65.5)	99.3 (89.2)	99.0 (74.6)
Mean $I/\sigma(I)$	11.71 (2.11)	12.11 (3.41)	7.18 (2.26)
Crystallographic refinement			
Asymmetric unit content	1 monomer	1 monomer	2 monomers
No. of non-hydrogen protein atoms	5,568	5,568	11,154
No. of water molecules	362	340	739
Other heteroatoms	0	12	159
Mean *B*-factor (Å^2)	28.0	22.6	21.8
$R_{factor}/R_{free}{}^c$ (%)	17.6/22.4	20.6/25.1	18.6/25.6
Stereochemical quality of the model			
RMSD bond lengths (Å)	0.008	0.008	0.009
RMSD bond angle (degrees)	1.2	1.1	1.2
Ramachandran plot favored (%)	96.4	97.1	96.3
Ramachandran outliers (%)	0.3	0.6	0.4

[a] $R_{meas} = \sum_{hkl}[N/(N-1)]^{1/2}\sum_i|I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where N is the multiplicity of a given reflection.
[b] CC(1/2) values were calculated using the program XSCALE.
[c] R_{factor} and R_{free} are given by $\sum|F_o - F_c|/\sum F_o$ with $R_{free} = R_{work}$ calculated using 5% random data excluded from the refinement.

SAXS experiments performed on SmTK in the presence of TSA or individual substrates are still consistent with an 80-kDa monomeric protein (Fig. 2), which rules out the possibility of a biological dimerization induced by the substrates. This result is supported by the PISA analysis of the packing of the new crystals. Interestingly, the two monomers, termed A and B, provide different structural information. The 2.3 Å resolution electron density maps clearly show that all TSA components are bound in each active site of A, whereas nitrate is absent from B.

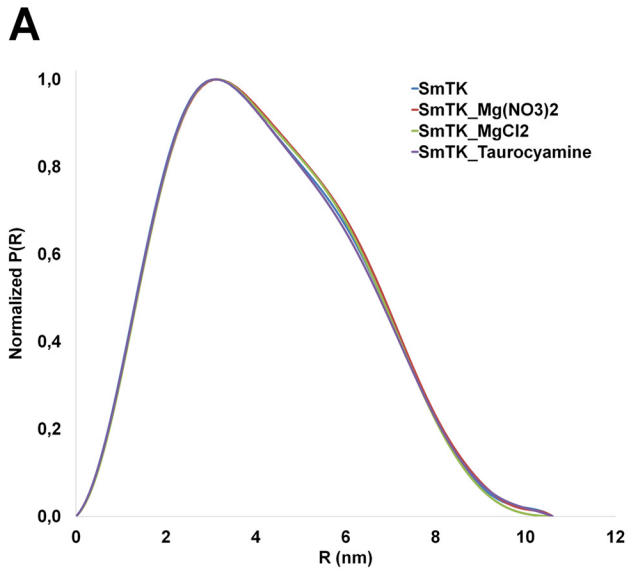
The TSA is firmly attached in lobe D2 of A and induces a global closure of the active site. A hinged rotation of 14° of the terminal region (residues 642–716) around an axis nearly parallel to the nucleotide-binding site cleft was calculated using the program DynDom (66) (Fig. 1*C*). This rigid body motion facilitates the stabilization of the flexible loop Gly^{670}–Gly^{681} that folds over the active site toward the phosphagen specificity loop (Fig. 3 and supplemental Video 1). A significant displacement, which affects helix α-8 and adjacent loop 544–552 on the other side of the active site crevice of D2, further constrains the bound nucleotide (Figs. 1*C* and 5*C*). Lobe D1 of A does not show such dramatic rearrangements upon TSA binding, and only helix α-8 and adjacent loop 181–189 move to restrict access to the nucleotide-binding site (Figs. 1*C* and 5*B*). The central interlobe zone of the protein mediated by helix α-12 of D1 appears to be unaffected by TSA binding (supplemental Video 1), as confirmed by a "Difference Distance Matrix Plot" produced by DDMP (available on the Yale University Web site) between the SmTK and SmTK-TSA structures (data not shown).

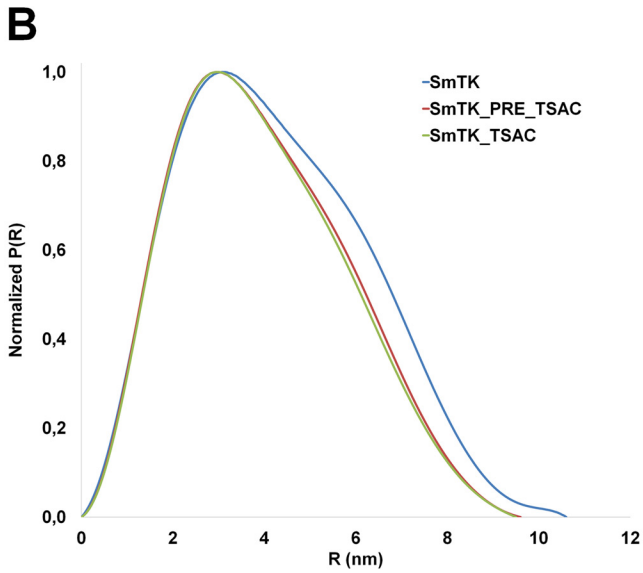
Superimposition of SmTK-TSA of B with unliganded SmTK did not reveal any significant rearrangements, despite the bound taurocyamine and ADP. A Mg^{2+} ion is only observed in the active site of D1, and the flexible loop is disordered in both lobes.

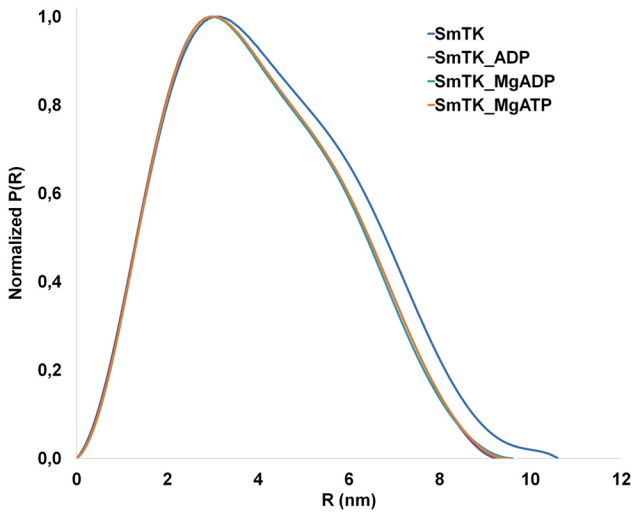
The Protein-TSA Complex—A dense network of interactions holds the TSA components, as described previously for other PK-TSA crystal structures (Fig. 3). The residues lining the ADP-binding site are almost identical between semiclosed D1 and closed D2 in monomer A. For D1, the invariant His^{182} (equivalent of His^{545} in D2) adjacent to the moving helix α-8 establishes a hydrogen bond with the 2′-hydroxyl group of the ribose, whereas the well conserved His^{281} (equivalent to Leu^{644} in D2) of strand β-6 makes a stacking interaction with the adenosine ring of the nucleotide. The distance between the Cα carbons of His^{182} and His^{281} diminishes by 5 Å between the open and semiclosed forms of D1. The same displacement is observed between the open and closed forms of D2.

The flexible loop in semiclosed D1 is in contact with the nucleotide ribose via a hydrogen bond between Glu^{310} and the 2′-hydroxyl group of the ribose. The equivalent Glu^{674} in D2 is directly implicated in taurocyamine binding because of the folding of the flexible loop over the TSA components (Fig. 3). Thus, the peculiar outward conformation of this loop in D1, which allows its interaction with the ADP-ribose, could represent an intermediary state between the open and closed conformations of the enzyme. The middle portion of the active site with the nucleotide diphosphate-, magnesium-, and nitrate-binding sites (Fig. 3) displays typical structural features of PKs in both D1 and D2 (14, 19, 20).

This structure of SmTK-TSA provides us with the first view of a taurocyamine-binding pocket in PKs. In the closed D2 form, the guanidinium group of taurocyamine is stabilized by bidentate interactions with both Glu^{674} of the flexible loop and the invariant Glu^{585} of the C-terminal domain, which enable







D

SmTK	APO	Taurocyamine	MgADP	MgATP	ADP	TSAC	PRE_TSAC	MgCl2	$Mg(NO_3)_2$
Rg (Guinier, nm)	3.35 ± 0.01	3.3 ± 0.2	3.08 ± 0.03	3.1 ± 0.1	3.08 ± 0.04	3.04 ± 0.01	3.08 ± 0.01	3.25 ± 0.01	3.35 ± 0.01
Rg (Gnom, nm)	3.30 ± 0.01	3.30 ± 0.01	3.11 ± 0.01	3.12 ± 0.01	3.11 ± 0.01	3.06 ± 0.01	3.08 ± 0.01	3.29 ± 0.01	3.32 ± 0.01
Dmax (nm)	10.6	10.6	9.6	9.5	9.4	9.5	9.6	10.6	10.6

FIGURE 2. **SAXS analysis and pair distance distribution functions *P*(*r*) obtained for SmTK in different conditions.** *P*(*r*) graphs for each condition (described under "Experimental Procedures") calculated with GNOM. The graphs were normalized to unity at their maximum value for comparison purposes. *A*, SmTK was analyzed in its ligand-free form or incubated with the guanidino substrate taurocyamine or with $Mg(Cl)_2$ or $Mg(NO_3)_2$ as a control for the effect of ions alone. *B*, SmTK was analyzed in its ligand-free form or incubated with the TSA components (MgADP-NO_3^{2-}-taurocyamine) or pre-TSA components (MgAMPPNP-taurocyamine). *C*, SmTK was analyzed in its ligand-free form or incubated with the nucleotide MgADP, ADP, or MgATP. *D*, *table* with the R_g values obtained either from the Guinier approximation or from GNOM for cross-validations, as well as Dmax, for SmTK in the different conditions tested.

the proper alignment of the phosphoryl group for the in-line transfer (Fig. 3). The side chain of the invariant Cys^{631} participates in the stabilization of the guanidinium group. Superimpositions of SmTK-TSA with known PKs complexed with TSA show a perfect match in the position and conformation of the bound guanidinium group, nitrate, and β-phosphate, further confirming the hypothesis of a highly conserved catalytic mechanism (67) (see Fig. 4*A* for superimposition with LvAK-TSA).

The sulfonate moiety of the taurocyamine bound in D2 is maintained by an extensive hydrogen-bonding network that involves His^{675} of the flexible loop (adjacent to Glu^{674}), two main-chain amine groups of the phosphagen-specificity loop (Ile^{423} and Cys^{424}), and the lateral chain hydroxyl group of Ser^{422}.

In addition to the shorter phosphagen specificity loop, the main difference between SmTK and AKs is the presence in this loop of Arg^{426}, which interacts through its guanidinium group with the sulfonate moiety of the taurocyamine substrate (Fig. 3). This arginine residue is conserved in D1 (Arg^{63}) as well as in all known TKs of trematodes but is substituted by a tyrosine in AKs (13) (Figs. 1*A* and 4*A*). Hydrogen bonds, similar to those observed for Arg^{426}, stabilize the interaction of Arg^{63} with the taurocyamine sulfonate group in the semiclosed lobe D1 of A. The flexible loop is not folded over the active site, and the bound guanidinium group is characterized by higher thermal B factors. However, a well ordered bound taurocyamine is observed in the open lobe D2 of B. The role of the incoming flexible loop is thus probably more to optimize the position of the guanidinium group than to stabilize it.

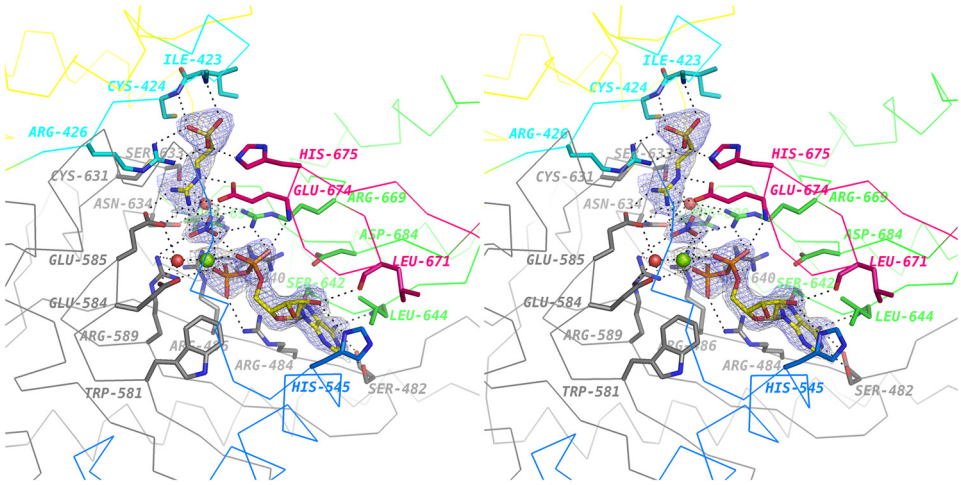


FIGURE 3. ***Walleye stereoview* of the closed active site of the SmTK-TSA D2 subunit.** The $2F_o - F_c$ maps (*black*) are contoured at 1σ, and the $F_o - F_c$ omit maps (*marine blue*) calculated without ligand coordinates are contoured at 3σ. Bound TSA components are represented in *sticks* as well as selected binding residues. Polar contacts < 3.2 Å are shown with *dashed lines*. Residues of the C-terminal part of SmTK (residues 642–716) are *colored* in *green*, and those of the region containing the helix α-8 are in *marine blue*. Residues of the phosphagen specificity and flexible loops are in *cyan* and *dark pink*, respectively. The N-terminal domain of D2 is in *yellow*, and the conformationally invariant region of D2 is in *gray*.

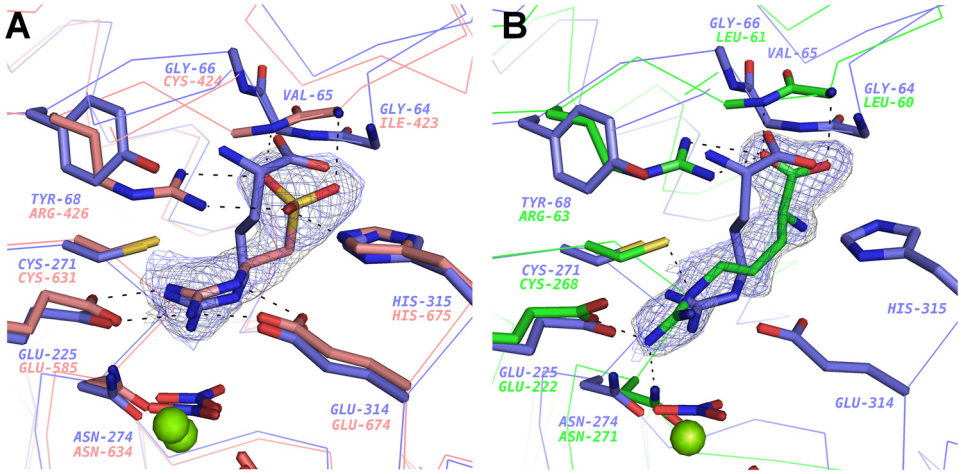


FIGURE 4. ***Close-up view* of the guanidino substrate binding site in SmTK and LvAK.** *A*, a *close-up view* centered on the guanidino substrate-binding site of LvAK-TSA (PDB code 4BG4) in *blue* and subunit D2 of SmTK-TSA monomer A in *pale red*. The Mg^{2+} ion is represented by a *green sphere*. *B*, a *close-up view* centered on the guanidino substrate-binding site of LvAK-TSA in *blue* and subunit D1 of SmTK-Arg in *green*. In both cases, the $2F_o - F_c$ maps (*black*) are contoured at 1σ, and the $F_o - F_c$ omit maps (*marine blue*) calculated without the guanidino substrate coordinates are contoured at 3σ.

More Insights into Phosphagen Specificity—Although SmTK is nearly inactive against L-arginine, we were able to determine the crystal structure of a binary SmTK-Arg complex. A bound L-arginine is clearly visible in the 1.9 Å resolution electron density maps of the D1 active site (Fig. 4*B*), and superimposition of SmTK-Arg and unliganded SmTK indicates no significant conformational change. This crystallographic result is consistent with our SAXS experiments performed with SmTK in the presence of taurocyamine, which indicated no significant change in R_g and D_{max} values (Fig. 2). Hence, in line with what has been described for CK and AK (25, 33, 34), binding of the guanidino substrate is not the main determinant of the conformational change cascade observed for the TSA complex.

Superimposition of SmTK-Arg and of the representative open LvAK-Arg or closed LvAK-TSA structures shows that the bound L-arginine adopts an elongated conformation in our binary complex that differs from the bent conformation observed in liganded AKs, in a manner that is independent of the folding of the flexible loop. A distinctive $\chi 3$ dihedral angle is observed between the two conformations, with a value of $-179°$ in the elongated L-arginine and $\sim 40°$ in the bent form. As a consequence, in SmTK-Arg, the bound guanidinium group is shifted toward the strictly conserved Glu^{222} and thus not properly positioned for in-line phosphoryl transfer. Its internal nitrogen (instead of the external nitrogen in liganded AKs) contacts the invariant cysteine Cys^{268} (Fig. 4*B*), whereas its main-chain nitrogen is displaced to avoid a steric clash with the long side chain of Arg^{63} of the phosphagen specificity loop (Arg^{426} in D2). In LvAK and in other AKs, Arg^{63} is substituted by the shorter Tyr^{68}, the hydroxyl group of which establishes a hydrogen bond with the bound L-arginine main-chain amine group, whereas this group protrudes into the solvent in SmTK-Arg.

Microcalorimetric and Enzymatic Investigations of a Putative Cooperative Mechanism—Because a significant structural asymmetry was observed between the liganded lobes of SmTK-TSA, the enzyme was examined by ITC for evidence of inter-

lobe cooperativity in TSA complex formation by direct titration of SmTK with ADP with or without other TSA components. In the presence of TSA components, ADP binding to wild-type SmTK was characterized by a monophasic sigmoidal titration curve, which enabled us to determine the stoichiometry of binding as 1.53 sites per SmTK protein (Fig. 6*A*). In the absence of TSA components, a clearly distinct binding isotherm was recorded with a binding pattern typical of a weak binding event, and no dissociation constant could be determined (data not shown). The same experiments were performed using the active-site cysteine single mutants, for which structural integrity was controlled by CD spectroscopy (Table 2). These mutants produced similar binding isotherms with K_D values in the micromolar range and a stoichiometry value reduced to ~0.9 (Fig. 6, *C* and *D*). The same experiment performed on the double cysteine mutant generated a binding isotherm characteristic of a very weak affinity binding event. No reasonable fit with a classical model of binding was possible, even when changing the concentrations of protein or ligand (data not shown). Together, these results support the observation that the two active sites of SmTK are able to bind TSA in solution with very similar thermodynamic parameters.

TABLE 2

Deconvolution of circular dichroism spectra into secondary structural components

Data are average values of separate calculations using CONTIN, CDSSTR, and SELCON3 ± S.D.

	α-Helix[a]	β-Strand[b]	Turn	Unordered
	%	%	%	%
$SmTK_{WT}$	38.7 ± 1.2	15.5 ± 1.0	17.9 ± 0.3	27.5 ± 0.5
$SmTK_{C268S}$	38.9 ± 0.2	15.4 ± 0.6	18.2 ± 0.7	27.7 ± 0.7
$SmTK_{C631S}$	38.1 ± 1.2	15.7 ± 0.9	17.6 ± 0.4	27.8 ± 0.9
$SmTK_{C268S/C631S}$	38.4 ± 0.5	15.8 ± 0.9	18.1 ± 0.8	27.9 ± 0.8

[a] Calculated values for regular and distorted helix are summed.
[b] Calculated values for regular and distorted β-strand are summed.

Enzyme kinetics experiments were also performed on wild-type and the invariant cysteine mutants of SmTK without any significant evidence for cooperativity (data not shown, consistent with a classical Michaelis-Menten behavior and Table 3). Although the kinetic parameters of single cysteine mutants should be taken with caution because the activity for the double mutant was severely reduced but not abolished ($k_{cat(app)}$ = 14.1 ± 0.7 s^{-1}), the results presented herein show that both sites displayed equivalent enzymatic activity.

DISCUSSION

A Bilobal Structure—These structures of SmTK with different liganded states represent the first atomic results obtained for a duplicated PK. Although the D1 and D2 contiguous lobes of SmTK interact in a manner different from that observed for known PKs, it appears that both duplicated and true dimeric PKs share a conserved generally U-shaped aspect. Our results also represent the first structural insight into a TK family member and allow a better understanding of both the substrate specificity of these enzymes and their molecular mechanism.

We observed that the unliganded D1 and D2 lobes display the classical open state form described previously for other substrate-free PKs (9, 10, 15–18). The flexible loop of the C-terminal domain is defined in density in D1 because of crystal contacts. This region adopts an outward conformation similar to that described for substrate-free marine worm GK and human brain-type CK, whereby it is also stabilized by crystal packing (19, 27). The disordered conformation of this loop observed in D2 allows equal access to the active site and is consistent with solution NMR measurements performed with AK (68) and CK (30).

The soaking of SmTK crystals in TSA components permitted the trapping of a series of transient states for each lobe of the two monomers contained in the crystal asymmetric unit (*i.e.*

TABLE 3

SmTK wild-type and invariant cysteine mutant enzyme kinetic data determined for the forward reaction (phosphagen synthesis)

Values represent mean ± S.D. (n = 3).

Construct	$k_{cat(app)}$ ATP	$K_{m(app)}$ ATP	$k_{cat(app)}$ ATP/$K_{m(app)}$ ATP	$k_{cat(app)}$ taurocyamine	$K_{m(app)}$ taurocyamine	$k_{cat(app)}$ taurocyamine/K_m(app) taurocyamine
	s^{-1}	mM		s^{-1}	mM	
$SmTK_{WT}$	87 ± 2	0.26 ± 0.02	334.6	108 ± 3	0.46 ± 0.04	234.8
$SmTK_{C268S}$	61 ± 1	0.36 ± 0.02	172.2	70 ± 1	0.57 ± 0.03	122.8
$SmTK_{C631S}$	36 ± 1	0.21 ± 0.02	171.4	47 ± 1	0.30 ± 0.03	156.7

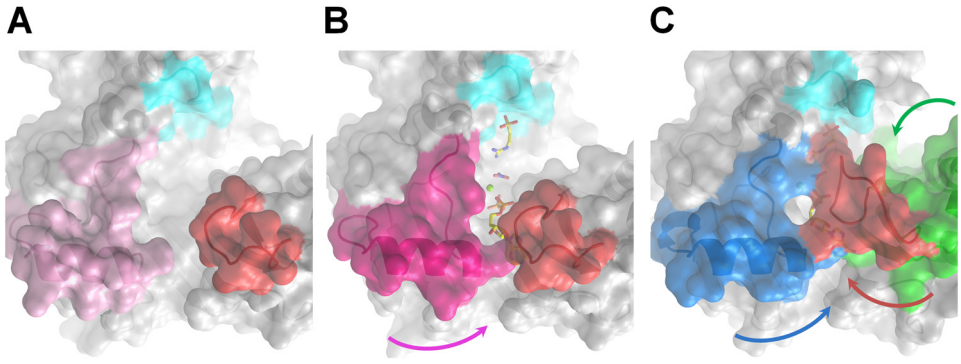
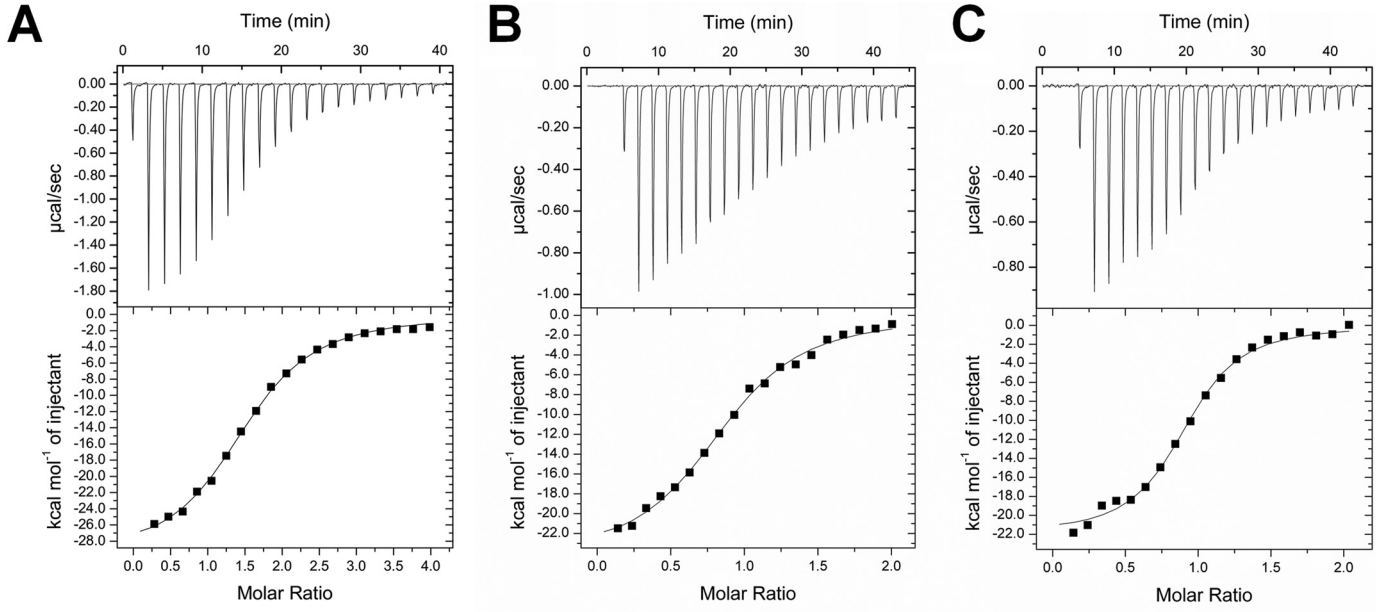


FIGURE 5. **Transparent surface representations of unliganded D1 in SmTK (*A*), liganded and semi-closed D1 in SmTK-TSA monomer A (*B*), and liganded and closed D2 in SmTK-TSA monomer A (*C*).** Protein is represented as a *ribbon*, and bound ligands are shown as *stick models* and *colored* in *yellow*. Conformational changes upon ligand binding are indicated by *colored arrows*. The *color code* of Fig. 3 is conserved.



Panel	Construct	K_D (µM)	n	ΔH (cal / mol)	ΔS (cal / mol / deg)
A	$SmTK_{WT}$	3.8 ± 0.3	1.53 ± 0.02	-3 x 10^4 ± 575	- 75
B	$SmTK_{C268S}$	2.7 ± 0.2	0.89 ± 0.01	-2.4 x 10^4 ± 558	- 55
C	$SmTK_{C631S}$	1.1 ± 0.1	0.90 ± 0.01	-2.2 x 10^4 ± 421	- 45

FIGURE 6. **ITC measurements of ADP binding to wild-type and mutated SmTK in the presence of the other TSA components.** The *top panels* show the raw data thermograms, and the *bottom panels* show the binding isotherms with the peak-integrated heats plotted *versus* the ADP/protein molar ratio. The *smooth line* corresponds to the fit of experimental data to a model consisting of one class of independent binding sites. *A*, titration of wild-type SmTK. *B*, titration of SmTK C268S mutant. *C*, titration of SmTK C631S mutant.

open, semiclosed, and closed forms, as shown in Fig. 5). Thus, one lobe exhibits a classical open form with partially bound TSA components and may represent the first transient intermediate. We postulate that the semiclosed form in which helix α-8 moves to accommodate ADP is produced next and precedes the closed form in which a rigid body motion of half of the C-terminal region stabilizes the flexible loop over the active site. The SmTK-TSA crystal also reveals that one SmTK monomer (chain A) can accommodate the TSA components in its two active sites, unlike what has been described for dimeric CK and AK complexed with ADP, pre-TSA, or TSA, which display a single-liganded active site (5, 14, 19). Additionally, this monomer displays a semiclosed D1 and closed D2, also differing from the fully bound and doubly closed dimer observed for the GK-TSA complex (27).

Asymmetrical features have been observed previously in crystal structures of dimeric PKs and are often viewed as support for the negative cooperativity behavior displayed by these enzymes (5, 14), although conflicting biochemical results have been published (30, 31, 37, 69, 70). Our ITC experiments on SmTK-TSA formation were best fitted to a monophasic binding event with a stoichiometry of 1.53 (Fig. 6). This value most likely indicates that the two active sites of the protein can adopt the TSA conformation because the stoichiometry of TSA formation is significantly reduced in each of the single cysteine mutants (0.89 and 0.9), but the close binding affinities remain similar. This result is consistent with our enzyme activity measurements that indicate that the lower activity of cysteine mutants is due mainly to reduced k_{cat} values. The observed K_D value for ADP within the context of the TSA complex (micromolar range in Fig. 6) is similar to that previously determined by fluorescence spectroscopy using rabbit muscle CK, whereas the binding affinity is weaker in the absence of TSA components (71). Formation of the TSA complex is severely affected in the double cysteine mutant and could not be characterized by ITC. Taken together, these data suggest that the two SmTK lobes are equivalent and bind to the ADP substrate in a non-mutually exclusive manner without signs of cooperativity. Moreover, enzyme kinetics experiments performed on SmTK are consistent with the Michaelis-Menten model, again without significant signs of cooperativity affecting catalytic properties.

The asymmetry observed in the structures is thus probably the result of the crystal environment. The SmTK-TSA structure was obtained by the soaking method with restrictions on the conformational freedom needed to obtain a doubly closed state for SmTK. We regularly observed that our soaked crystals lost diffraction to x-ray and dissolved after ~1 h, most likely because of structural rearrangements in D1 due to lobe closure. Our results are thus consistent with the idea of PK enzymes existing in solution in an equilibrium of multiple and intermediary states (fully open to fully closed) of nearly equal energies in the presence or absence of ligands, with environmental factors selecting the conformer observed for each crystal struc-

ture. Such conformational changes could be rate-limiting on turnover (28).

In addition, we performed SAXS experiments to gain information on the doubly closed form that can be attained in solution, and significant changes were observed in the R_g and D_{max} values between substrate-free SmTK (or SmTK in the presence of taurocyamine) and SmTK in the presence of ADP, pre-TSA, or TSA components. These observations agree with the idea of nucleotide binding mainly accounting for the shift between conformations (32–34). Interestingly, the presence of magnesium does not appear to be necessary for this conformational change to occur, at least for ADP binding (see Fig. 2*C*), unlike what was previously reported for SAXS experiments performed on CK (32–34). We did, however, observe by thermal shift assays that the binding affinities for ADP and ATP nucleotides are severely reduced in the absence of magnesium (data not shown). A SAXS envelope of SmTK-TSA was calculated, and the general U-shape of the enzyme was still visible. The resulting atomic model of SmTK-TSA fitted into the SAXS envelope suggested that D1 and D2 adopt a similar closed state that results in a more compact enzyme (data not shown).

Substrate Specificity—PKs have been identified in a number of important human and animal parasites and have been suggested to constitute novel therapeutic targets (72). Because TK is absent in mammals but is well conserved in the human pathogens *S. mansoni* and *S. japonicum*, our crystal structures of free and substrate-bound SmTK can provide a template for the rational development of new drugs against schistosomiasis.

Structural analyses of the liganded monomers contained in the SmTK-TSA crystal and comparisons with LvAK-TSA revealed that position 63/426 (SmTK numbering) of the phosphagen-binding loop is essential for substrate specificity in trematode TKs and AKs. In SmTK, this position bears an arginine, a residue with a basic group that tightly interacts with the specific sulfonate moiety of the taurocyamine substrate, regardless of whether the lobe is in an open state or a closed state. This structural result is supported by previous kinetics experiments and site-directed mutagenesis studies on this arginine, which is essential for *S. japonicum* TK activity (13). A tyrosine has been observed in an equivalent position in AKs, where it binds to the main-chain amine group of the L-arginine substrate. It is worth noting that the adjacent Glu^{674} and His^{675} of the flexible loop reinforce taurocyamine binding in closed SmTK-TSA and are conserved in AKs, suggesting a similar role. Hence, the glutamate residue contacts the bound guanidinium group in both closed LvAK-TSA and SmTK-TSA, whereas the histidine residue interacts with the L-arginine main chain in LvAK-TSA and the taurocyamine sulfonate moiety in SmTK-TSA. Such significant sequence conservation in the flexible loop is in clear agreement with the hypothesis that TKs are most similar to AKs and evolved from an AK-like ancestor (41).

The determination of the binary complex SmTK-Arg and the comparison with LvAK-TSA helped us to further explain the determinants of guanidino substrate specificity. Our results highlight the importance of the bent conformation of the guanidino substrate. This configuration is first ensured by proper binding with the phosphagen specificity loop, as shown by the inappropriate orientation of the L-arginine main chain in inactive SmTK-Arg, which results in an elongated conformation of the compound when bound at high occupancy in the crystal. A survey of liganded and active PK structures showed that the bent conformation of the guanidinium group with respect to the other ligand extremity is conserved in all bound guanidino substrates (creatine, L-arginine, glycocyamine, and now taurocyamine) once they are engaged in a bidentate interaction with the invariant glutamate of the C-terminal domain ($Glu^{222/585}$ in SmTK). Accordingly, it was previously shown that phosphagen specificity is not derived from a lock-and-key mechanism but potentially from subtle distortions in bound substrate configurations (24).

Our enzymatic study shows that both the D1 and D2 lobes are capable of TK activity. Gene duplications have been recognized as an important source of evolutionary innovation and adaptation. Indeed, it has been argued that the presence of duplicate genes could lead to the conservation of gene function in both copies and be beneficial simply because of the additional amount of active protein, with implications primarily for highly expressed genes, the products of which are in high demand (concerted evolution model). Another process involving gene duplication is neo-functionalization, in which one copy undergoes directional selection to perform a novel function after duplication (73). Our results suggest that if this is the case for SmTK, the advantage is not linked to a significant improvement in enzymatic properties. Additional investigations are required to better understand the nature of the selective pressure that leads to SmTK duplication.

Acknowledgments—We thank Dr. Xavier Robert for critical reading of the manuscript, and we thank the staff of the PXIII (Swiss Light Source) and MX (European Synchrotron Radiation Facility) beamlines for technical support. We thank Clément Danis, who did a 3-month internship on the project during work on a master's degree at Université Lyon 1.

REFERENCES

1. Ellington, W. R. (2001) Evolution and physiological roles of phosphagen systems. *Annu. Rev. Physiol.* **63,** 289–325
2. Suzuki, T., Soga, S., Inoue, M., and Uda, K. (2013) Characterization and origin of bacterial arginine kinases. *Int. J. Biol. Macromol.* **57,** 273–277
3. Compaan, D. M., and Ellington, W. R. (2003) Functional consequences of a gene duplication and fusion event in an arginine kinase. *J. Exp. Biol.* **206,** 1545–1556
4. Wyss, M., and Kaddurah-Daouk, R. (2000) Creatine and creatinine metabolism. *Physiol. Rev.* **80,** 1107–1213
5. Wu, X., Ye, S., Guo, S., Yan, W., Bartlam, M., and Rao, Z. (2010) Structural basis for a reciprocating mechanism of negative cooperativity in dimeric phosphagen kinase activity. *FASEB J.* **24,** 242–252
6. Hoffman, G. G., Davulcu, O., Sona, S., and Ellington, W. R. (2008) Contributions to catalysis and potential interactions of the three catalytic domains in a contiguous trimeric creatine kinase. *FEBS J.* **275,** 646–654
7. Marcillat, O., Goldschmidt, D., Eichenberger, D., and Vial, C. (1987) Only one of the two interconvertible forms of mitochondrial creatine kinase binds to heart mitoplasts. *Biochim. Biophys. Acta* **890,** 233–241
8. Vial, C., Marcillat, O., Goldschmidt, D., Font, B., and Eichenberger, D. (1986) Interaction of creatine kinase with phosphorylating rabbit heart mitochondria and mitoplasts. *Arch. Biochem. Biophys.* **251,** 558–566
9. Fritz-Wolf, K., Schnyder, T., Wallimann, T., and Kabsch, W. (1996) Structure of mitochondrial creatine kinase. *Nature* **381,** 341–345
10. Eder, M., Fritz-Wolf, K., Kabsch, W., Wallimann, T., and Schlattner, U.

(2000) Crystal structure of human ubiquitous mitochondrial creatine kinase. *Proteins* **39,** 216–225
11. Suzuki, T., Sugimura, N., Taniguchi, T., Unemi, Y., Murata, T., Hayashida, M., Yokouchi, K., Uda, K., and Furukohri, T. (2002) Two-domain arginine kinases from the clams *Solen strictus* and *Corbicula japonica*: exceptional amino acid replacement of the functionally important D(62) by G. *Int. J. Biochem. Cell Biol.* **34,** 1221–1229
12. Suzuki, T., Kawasaki, Y., Unemi, Y., Nishimura, Y., Soga, T., Kamidochi, M., Yazawa, Y., and Furukohri, T. (1998) Gene duplication and fusion have occurred frequently in the evolution of phosphagen kinases: a two-domain arginine kinase from the clam *Pseudocardium sachalinensis*. *Biochim. Biophys. Acta* **1388,** 253–259
13. Tokuhiro, S., Nagataki, M., Jarilla, B. R., Uda, K., Suzuki, T., Sugiura, T., and Agatsuma, T. (2014) Phosphagen kinase in *Schistosoma japonicum*: II. Determination of amino acid residues essential for substrate catalysis using site-directed mutagenesis. *Mol. Biochem. Parasitol.* **194,** 56–63
14. Lahiri, S. D., Wang, P. F., Babbitt, P. C., McLeish, M. J., Kenyon, G. L., and Allen, K. N. (2002) The 2.1 Å structure of *Torpedo californica* creatine kinase complexed with the ADP-Mg^{2+}-NO_3^--creatine transition-state analogue complex. *Biochemistry* **41,** 13861–13867
15. Rao, J. K., Bujacz, G., and Wlodawer, A. (1998) Crystal structure of rabbit muscle creatine kinase. *FEBS Lett.* **439,** 133–137
16. Shen, Y. Q., Tang, L., Zhou, H. M., and Lin, Z. J. (2001) Structure of human muscle creatine kinase. *Acta Crystallogr. D Biol. Crystallogr.* **57,** 1196–1200
17. Tisi, D., Bax, B., and Loew, A. (2001) The three-dimensional structure of cytosolic bovine retinal creatine kinase. *Acta Crystallogr. D Biol. Crystallogr.* **57,** 187–193
18. Eder, M., Schlattner, U., Becker, A., Wallimann, T., Kabsch, W., and Fritz-Wolf, K. (1999) Crystal structure of brain-type creatine kinase at 1.41 Å resolution. *Protein Sci.* **8,** 2258–2269
19. Bong, S. M., Moon, J. H., Nam, K. H., Lee, K. S., Chi, Y. M., and Hwang, K. Y. (2008) Structural studies of human brain-type creatine kinase complexed with the ADP-Mg^{2+}-NO_3^--creatine transition-state analogue complex. *FEBS Lett.* **582,** 3959–3965
20. Zhou, G., Somasundaram, T., Blanc, E., Parthasarathy, G., Ellington, W. R., and Chapman, M. S. (1998) Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 8449–8454
21. Azzi, A., Clark, S. A., Ellington, W. R., and Chapman, M. S. (2004) The role of phosphagen specificity loops in arginine kinase. *Protein Sci.* **13,** 575–585
22. Fernandez, P., Haouz, A., Pereira, C. A., Aguilar, C., and Alzari, P. M. (2007) The crystal structure of *Trypanosoma cruzi* arginine kinase. *Proteins* **69,** 209–212
23. Niu, X., Bruschweiler-Li, L., Davulcu, O., Skalicky, J. J., Brüschweiler, R., and Chapman, M. S. (2011) Arginine kinase: joint crystallographic and NMR RDC analyses link substrate-associated motions to intrinsic flexibility. *J. Mol. Biol.* **405,** 479–496
24. Clark, S. A., Davulcu, O., and Chapman, M. S. (2012) Crystal structures of arginine kinase in complex with ADP, nitrate, and various phosphagen analogs. *Biochem. Biophys. Res. Commun.* **427,** 212–217
25. López-Zavala, A. A., García-Orozco, K. D., Carrasco-Miranda, J. S., Sugich-Miranda, R., Velázquez-Contreras, E. F., Criscitiello, M. F., Brieba, L. G., Rudiño-Piñera, E., and Sotelo-Mundo, R. R. (2013) Crystal structure of shrimp arginine kinase in binary complex with arginine-a molecular view of the phosphagen precursor binding to the enzyme. *J. Bioenerg. Biomembr.* **45,** 511–518
26. Yousef, M. S., Clark, S. A., Pruett, P. K., Somasundaram, T., Ellington, W. R., and Chapman, M. S. (2003) Induced fit in guanidino kinases: comparison of substrate-free and transition state analog structures of arginine kinase. *Protein Sci.* **12,** 103–111
27. Lim, K., Pullalarevu, S., Surabian, K. T., Howard, A., Suzuki, T., Moult, J., and Herzberg, O. (2010) Structural basis for the mechanism and substrate specificity of glycocyamine kinase, a phosphagen kinase family member. *Biochemistry* **49,** 2031–2041
28. Bush, D. J., Kirillova, O., Clark, S. A., Davulcu, O., Fabiola, F., Xie, Q., Somasundaram, T., Ellington, W. R., and Chapman, M. S. (2011) The structure of lombricine kinase: implications for phosphagen kinase conformational changes. *J. Biol. Chem.* **286,** 9338–9350
29. Suzuki, T., Kawasaki, Y., Furukohri, T., and Ellington, W. R. (1997) Evolution of phosphagen kinase. VI. Isolation, characterization and cDNA-derived amino acid sequence of lombricine kinase from the earthworm *Eisenia foetida*, and identification of a possible candidate for the guanidine substrate recognition site. *Biochim. Biophys. Acta* **1343,** 152–159
30. Rivière, G., Hologne, M., Marcillat, O., and Lancelin, J. M. (2012) Dynamical properties of the loop 320s of substrate-free and substrate-bound muscle creatine kinase by NMR: evidence for independent subunits. *FEBS J.* **279,** 2863–2875
31. Mazon, H., Marcillat, O., Forest, E., and Vial, C. (2003) Changes in MM-CK conformational mobility upon formation of the ADP-Mg^{2+}-NO_3^--creatine transition state analogue complex as detected by hydrogen/deuterium exchange. *Biochemistry* **42,** 13596–13604
32. Dumas, C., and Janin, J. (1983) Conformational changes in arginine kinase upon ligand binding seen by small-angle x-ray scattering. *FEBS Lett.* **153,** 128–130
33. Forstner, M., Kriechbaum, M., Laggner, P., and Wallimann, T. (1996) Changes of creatine kinase structure upon ligand binding as seen by small-angle scattering. *J. Mol. Struct.* **383,** 217–222
34. Forstner, M., Kriechbaum, M., Laggner, P., and Wallimann, T. (1998) Structural changes of creatine kinase upon substrate binding. *Biophys. J.* **75,** 1016–1023
35. Londergan, C. H., Baskin, R., Bischak, C. G., Hoffman, K. W., Snead, D. M., and Reynoso, C. (2015) Dynamic asymmetry and the role of the conserved active-site thiol in rabbit muscle creatine kinase. *Biochemistry* **54,** 83–95
36. Awama, A. M., Mazon, H., Vial, C., and Marcillat, O. (2007) Despite its high similarity with monomeric arginine kinase, muscle creatine kinase is only enzymatically active as a dimer. *Arch. Biochem. Biophys.* **458,** 158–166
37. Forstner, M., Berger, C., and Wallimann, T. (1999) Nucleotide binding to creatine kinase: an isothermal titration microcalorimetry study. *FEBS Lett.* **461,** 111–114
38. Awama, A. M., Paracuellos, P., Laurent, S., Dissous, C., Marcillat, O., and Gouet, P. (2008) Crystallization and x-ray analysis of the *Schistosoma mansoni* guanidino kinase. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **64,** 854–857
39. Shoemaker, C. B. (1994) The *Schistosoma mansoni* phosphagen kinase gene contains two closely apposed transcription initiation sites and arose from a fused gene duplication. *Mol. Biochem. Parasitol.* **68,** 319–322
40. Stein, L. D., Harn, D. A., and David, J. R. (1990) A cloned ATP:guanidino kinase in the trematode *Schistosoma mansoni* has a novel duplicated structure. *J. Biol. Chem.* **265,** 6582–6588
41. Jarilla, B. R., Tokuhiro, S., Nagataki, M., Hong, S. J., Uda, K., Suzuki, T., and Agatsuma, T. (2009) Molecular characterization and kinetic properties of a novel two-domain taurocyamine kinase from the lung fluke *Paragonimus westermani*. *FEBS Lett.* **583,** 2218–2224
42. Xiao, J. Y., Lee, J. Y., Tokuhiro, S., Nagataki, M., Jarilla, B. R., Nomura, H., Kim, T. I., Hong, S. J., and Agatsuma, T. (2013) Molecular cloning and characterization of taurocyamine kinase from *Clonorchis sinensis*: a candidate chemotherapeutic target. *PLoS Negl. Trop. Dis.* **7,** e2548
43. Uda, K., Yamamoto, K., Iwasaki, N., Iwai, M., Fujikura, K., Ellington, W. R., and Suzuki, T. (2008) Two-domain arginine kinase from the deep-sea clam *Calyptogena kaikoi*: evidence of two active domains. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **151,** 176–182
44. Tokuhiro, S., Uda, K., Yano, H., Nagataki, M., Jarilla, B. R., Suzuki, T., and Agatsuma, T. (2013) Phosphagen kinase in *Schistosoma japonicum*: characterization of its enzymatic properties and determination of its gene structure. *Mol. Biochem. Parasitol.* **188,** 91–98
45. Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32,** W668–W673
46. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 133–144
47. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Echols, N., Headd, J. J., Hung, L. W., Jain, S., Kapral, G. J., Grosse Kunstleve, R. W., McCoy,

A. J., Moriarty, N. W., Oeffner, R. D., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2011) The Phenix software for automated determination of macromolecular structures. *Methods* **55,** 94–106
48. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
49. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C., and Adams, P. D. (2008) Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* **426,** 419–435
50. Moriarty, N. W., Grosse-Kunstleve, R. W., and Adams, P. D. (2009) electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* **65,** 1074–1080
51. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21
52. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22,** 2577–2637
53. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42,** W320–W324
54. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36,** 1277–1282
55. Petoukhov, M. V., Konarev, P. V., Kikhney, A. G., and Svergun, D. I. (2007) ATSAS 2.1: towards automated and web-supported small-angle scattering data analysis. *J. Appl. Crystallogr.* **40,** s223–s228
56. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25,** 495–503
57. Mylonas, E., and Svergun, D. I. (2007) Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J. Appl. Crystallogr.* **40,** s245–s249
58. Svergun, D. I., Barberato, C., and Koch, M. H. J. (1995) CRYSOL: a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28,** 768–773
59. Franke, D., and Svergun, D. I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42,** 342–346
60. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of *ab initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36,** 860–864
61. Kozin, M. B., and Svergun, D. I. (2001) Automated matching of high- and low-resolution structural models. *J. Appl. Crystallogr.* **34,** 33–41
62. Morrison, J. F., Ennor, A. H., and Griffiths, D. E. (1958) The preparation of barium monophosphotaurocyamine. *Biochem. J.* **68,** 447–452
63. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372,** 774–797
64. Gaffney, J. P., and Valentine, A. M. (2012) Beyond bilobal: transferrin homologs having unusual domain architectures. *Biochim. Biophys. Acta* **1820,** 212–217
65. Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38,** W545–W549
66. Hayward, S., and Lee, R. A. (2002) Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graph. Model.* **21,** 181–183
67. Pruett, P. S., Azzi, A., Clark, S. A., Yousef, M. S., Gattis, J. L., Somasundaram, T., Ellington, W. R., and Chapman, M. S. (2003) The putative catalytic bases have, at most, an accessory role in the mechanism of arginine kinase. *J. Biol. Chem.* **278,** 26952–26957
68. Davulcu, O., Flynn, P. F., Chapman, M. S., and Skalicky, J. J. (2009) Intrinsic domain and loop dynamics commensurate with catalytic turnover in an induced-fit enzyme. *Structure* **17,** 1356–1367
69. Degani, C., and Degani, Y. (1980) Further evidence for nonsymmetric subunit association and intersubunit cooperativity in creatine kinase. Subunit-selective modifications by 2,4-dinitrophenylthiocyanate. *J. Biol. Chem.* **255,** 8221–8228
70. Price, N. C., and Hunter, M. G. (1976) Non-identical behaviour of the subunits of rabbit muscule creatine kinase. *Biochim. Biophys. Acta* **445,** 364–376
71. Borders, C. L., Jr., Snider, M. J., Wolfenden, R., and Edmiston, P. L. (2002) Determination of the affinity of each component of a composite quaternary transition-state analogue complex of creatine kinase. *Biochemistry* **41,** 6995–7000
72. Jarilla, B. R., and Agatsuma, T. (2010) Phosphagen kinases of parasites: unexplored chemotherapeutic targets. *Korean J. Parasitol.* **48,** 281–284
73. Zhang, J. (2003) Evolution by gene duplication: an update. *Trends Ecol. Evol.* **18,** 292–298