Published in final edited form as:

Biochemistry. 2010 March 9; 49(9): 2031–2041. doi:10.1021/bi9020988.

Structural basis for the mechanism and substrate specificity of glycoamine kinase, a phosphagen kinase family member†,‡

Kap Lim^{§,1}, Sadhana Pullalarevu^{§,1}, Karen Talin Surabian[§], Andrew Howard^{||}, Tomohiko Suzuki[‡], John Moul[§], and Osnat Herzberg^{§,*}

[§]Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland, USA

^{||}Illinois Institute of Technology, Chicago, Illinois, USA

[‡]Laboratory of Biochemistry, Faculty of Science, Kochi University, Kochi, Japan

Abstract

Glycoamine kinase (GK), a member of the phosphagen kinase family, catalyzes the Mg^{2+} -dependent reversible phosphoryl group transfer of the N-phosphoryl group of phospho glycoamine to ADP to yield glycoamine and ATP. This reaction helps to maintain the energy homeostasis of the cell in some multicellular organisms that encounter high and variable energy turnover. GK from the marine worm *Namalycastis* sp. is heterodimeric, with two homologous polypeptide chains, α and β , derived from a common pre mRNA by mutually exclusive N-terminal alternative exons. The N-terminal exon of GK β encodes a peptide that is different in sequence and is sixteen amino acids longer than that encoded by the N-terminal exon of GK α . The crystal structures of recombinant GK $\alpha\beta$ and GK $\beta\beta$ from *Namalycastis* sp. were determined at 2.6 Å and 2.4 Å resolution, respectively. In addition, the structure of the GK $\beta\beta$ was determined at 2.3 Å resolution in complex with a transition state analog, Mg^{2+} -ADP- NO_3^- -glycoamine. Consistent with the sequence homology, the GK subunits adopt the same overall fold as that of other phosphagen kinases of known structure (the homodimeric creatine kinase (CK) and the monomeric arginine kinase (AK)). As with CK, the GK N-termini mediate the dimer interface. In both heterodimeric and homodimeric GK forms, the conformations of the two N-termini are asymmetric and the asymmetry is different than that reported previously for the homodimeric CKs from several organisms. The entire polypeptide chains of GK $\alpha\beta$ are structurally defined and the longer N-terminus of the β subunit is anchored at the dimer interface. In GK $\beta\beta$ the 24 N-terminal residues of one subunit and 11 N-terminal residues of the second subunit are disordered. This observation is consistent with a proposal that the GK $\alpha\beta$ amino acids involved in the interface formation were optimized once a heterodimer emerged as the physiological form of the enzyme. As a consequence, the homodimer interface (either solely α or solely β chains) has been corrupted. In the unbound state, GK exhibits an open conformation analogous to that observed with ligand-free CK or AK. Upon binding the transition state analog, both subunits of GK undergo the same closure motion that clasps the transition state analog, in contrast to the transition state analog complexes of CK, where the corresponding transition state analog occupies only one subunit, which undergoes domain closure. The active site environments of the GK, CK and AK at

[†]Grant sponsors: National Institute of Health PO1-GM057890. The Advanced Photon Source is supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under contract W-31-109-Eng-38.

[‡]Protein Data Bank entry codes 3L2E for the ligand-free GK $\alpha\beta$; 3L2D for the ligand-free GK $\beta\beta$; 3L2F and 3L2G for the GK $\beta\beta$ /transition state analog complex (split into two entries).

*Corresponding author: Osnat Herzberg, Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville MD 20850, Tel: 240-314-6245; Fax: 240-314-6255; herzberg@umbi.umd.edu.

¹These authors contributed equally to the studies.

the bound states reveal the structural determinants of substrate specificity. Despite the equivalent binding in both active sites of the GK heterodimer, the conformational asymmetry of the N-termini is retained. Thus, the coupling between the structural asymmetry and negative cooperativity previously proposed for CK is not supported in the case of GK.

Glycocyamine kinase (ATP:guanidinoacetate N-phosphotransferase, guanidinoacetate kinase, GK², EC 2.7.3.1) belongs to the family of phosphagen kinases, enzymes that catalyze the Mg²⁺-dependent reversible conversion of guanidinium-containing substrates and ATP into the respective N^ω-phosphorylated substrate and ADP (1-3). GK acts on glycocyamine (guanidinoacetate) and phosphoglycocyamine (phosphoguanidinoacetate, N^ω-phosphonoglycocyamine) (Fig. 1). Phosphoglycocyamine and other phosphagens serve as energy reservoirs for rapid regeneration of ATP and for maintaining the energy balance of the cell. Whereas vertebrates produce only one phosphagen kinase, creatine kinase (CK), lower invertebrates produce multiple phosphagen kinases including GK (4). Eight phosphagen kinases have been identified in the phyla annelida, CK, arginine kinase (AK), GK, lombricine kinase, thalassemine kinase, taurocyamine kinase, hypotaurocyamine kinase and opheline kinase (1, 4, 5). The need for so many phosphagen systems in the same organism is poorly understood. All these enzymes are thought to have evolved from a common ancestor (6). It was proposed that because creatine contains a methyl-substituted internal nitrogen atom (guanidino N), the compound is thermodynamically less stable than other phosphagens containing unsubstituted internal N such as glycocyamine or arginine. Thus the latter phosphagens may be preferred over phosphocreatine for ATP production under stress conditions such as acidosis (7). The importance of phosphoglycocyamine for immediate energy provision under extreme environmental stress is underscored by the unprecedented high content of this compound in the warm *Nephtys hombergii*, a burrowing predator living in marine sediments (8).

Sequence analyses led to the proposal that AK is the oldest phosphagen kinase and CK diverged from it at an early stage of evolution (9, 10). Monomeric AK is the most prevalent phosphagen kinase in invertebrates. Dimeric AK exists in sea cucumber and may have back evolved from CK (11). There are three types of CK, cytosolic (muscle-type and brain-type), mitochondrial, and flagellar (12). The cytosolic CK assembles into dimers. Mitochondrial CK is in equilibrium between dimeric and octameric forms (13, 14). Flagellar CK comprises a fused trimer, which is found in many protostome and deuterostome invertebrates and in protochordate.

The GK from marine worms is a dimeric protein that is alternatively spliced at the N-termini of the subunits with the shorter exon encoding the α chain and the sixteen amino acid residues longer exon encoding the β chain (6). Ellington and colleagues proposed that the ancestral GK was composed of the $\alpha\alpha$ homodimer but alternative splicing occurred possibly to remove the deleterious effect of mutation in the original gene, thus resulting in the GK $\alpha\beta$ heterodimeric enzyme currently present in the marine worms. The molecular mass of the α subunit is ~42 kDa and that of the β subunit is ~44 kDa (6). Because of the high level of sequence identity between GK and CK (>50%), the three-dimensional structures of the two enzymes are expected to be similar.

The crystal structures of various homodimeric vertebrate CKs, with and without bound ligands, show that the N-terminal region is involved in dimer formation (15-20). With the exception of one example of dimer subunits that are related by perfect crystallographic symmetry and contain N-termini that adopts two alternative conformations (20), each of the

²The abbreviations used are: GK – glycocyamine kinase; CK – creatine kinase; AK – arginine kinase; TSA – transition state analog.

two CK N-termini adopts a different conformation, which introduces structural asymmetry irrespective of the presence of bound ligand (15-17, 21). Structures of rabbit muscle, and *Torpedo californica* CKs in complex with the transition state analog (TSA) Mg^{2+} -ADP- NO_3 -creatine revealed an asymmetry; a ligand-free subunit (or bound to only Mg^{2+} -ADP) that adopted an open conformation, and a TSA-bound subunit that adopted the closed conformation (16, 17). For the human brain-type CK, both the CK/TSA or CK/ADP complexes exhibited the closed conformation in the ligand-bound subunit and the open conformation in the ligand-free subunit (21). The relationship between the N-termini conformational asymmetry to the observed subunit binding asymmetry are unknown, although it was recently proposed that the N-terminal regions play a role in active site communication and catalytic regulation (17). The structural asymmetry supports the view that the CK subunits exhibit negative cooperativity (22-25), whereas other chemical modification experiments and data from deuterium exchange experiments contradict the notion of negative cooperativity in CK (26, 27). The two distinct alternative spliced GK N-termini imply an inherently asymmetric structure, which is confirmed by the crystal structures of both the homodimeric GK $\beta\beta$ and the heterodimeric GK $\alpha\beta$ reported here. In contrast to the CK-TSA complexes, GK binds its TSA, Mg^{2+} -ADP- NO_3 -glycocyamine, in both active sites and both subunits exhibit the closed conformation while maintaining the asymmetric dimer interface.

Materials and Methods

Cloning, Expression and purification

The GK α and β genes were cloned into the pDEST-His-MBP vector, kindly provided by David Waugh (28). The genes were amplified from the pMAL-c2/GK α and pMAL-c2/GK β plasmids (29) by PCR using the following primers.

GK α :

5'GAGAACCTGTACTTCCAGGGTATGTTTAAGGACTACACTAGGGAGAA
ATTTGC3' (forward),

5'
GGGGACCACTTTGTACAAGAAAGCTGGGTTATTATTACTTGGGAATCAT
GTCATCGATG3' (reverse).

GK β :

5'GAGAACCTGTACTTCCAGGGTATGGGTTTCAGCCATTCAGGATTACT3'
(forward),

5'GGGGACCACTTTGTACAAGAAAGCTGGGTTATTATTACTTGGGAATCA
TGTCATCGATG3' (reverse).

Each of these PCR fragments contained a N-terminal tobacco etch virus (TEV) protease cleavage site sequence designed to be inserted between a maltose binding protein (MBP) gene and the respective GK gene. The resulting PCR products were used as templates for a second PCR with the forward primer

5'GGGGACAAGTTTGTACAAAAAAGCAGGCTCGGAGAACCTGTACTTCCAGGGT
3' and the GK α or GK β reverse primer to anneal an attB1 recognition site. Each of the two resulting PCR fragments were inserted into pDNOR221 (Invitrogen) by recombinational cloning with the BP clonase system using the standard Gateway protocol (Invitrogen). The nucleotide sequence of the entry clones pGK α 221 and pGK β 221 were verified by sequencing. The His₆-MBP fusion protein expression vectors were constructed by recombining the GK genes from pGK α 221 and pGK β 221 entry vectors to the destination vector pDEST-His-MBP using the LR clonase enzyme (Invitrogen).

E. coli BL21star(DE3) cells transformed with GK α or GK β were grown in Luria broth (LB) media containing 100 μ g/mL ampicillin to OD_{600nm} ~ 0.6 at 37 °C, chilled on ice, and induced with 1 mM isopropyl-D-thiogalactopyranoside (IPTG) in a shaker maintained at 22° C. The cells were harvested by centrifugation after 23 h and the pellet was frozen at -80° C. The frozen cells were thawed and resuspended in lysis buffer containing 20 mM Tris-HCl (pH 8.5), 100 mM NaCl, 1 mM DTT and 0.5 mM EDTA, in 1x of Bug buster protein extraction reagent (Novagen), benzonase and protease cocktail for His-tagged proteins (Sigma). The cells were then lysed by sonication, centrifuged, and the supernatant was applied onto Ni-NTA column (Qiagen). The bound protein was eluted in a step gradient with 250 mM imidazole-containing buffer. TEV protease was added to the MBP-GK fusion protein solution at 1:200 molar ratio. After 16 hours, the reaction mixture was applied onto a second Ni-NTA column to bind the His-tagged cleaved MBP and the His-tagged TEV protease and elute the cleaved GK. The protein was further purified by size exclusion chromatography on Sephacryl S-100 column using 20 mM Tris-HCl (pH 8.5) and 100 mM NaCl. Protein expression level and sample purity were assessed by SDS-PAGE analysis and homogeneity of the purified protein was assessed by native PAGE analysis. The molecular weight was confirmed by MALDI-TOF mass spectrometry. The oligomeric state of the protein was determined by analytical size exclusion chromatography using a Superdex S200 analytical size exclusion column (GE healthcare) and also by dynamic light scattering using a DynaPro instrument (Protein solutions).

Crystallization and data collection

GK crystals were obtained at room temperature by vapor diffusion in hanging drops. Drops containing equal volumes of protein (at concentration of 10 mg/mL) and reservoir solution were equilibrated against the reservoir solution containing 22% w/v polyethylene glycol 3350, 200 mM NaNO₃, and 100 mM Tris-HCl (pH 8.0). To obtain GK α β crystals, GK α and GK β solutions were mixed in 1:1 molar ratio just prior to setting up crystallization experiments. Crystals of 0.2 x 0.2 x 0.1 mm³ size appeared within a week. Both GK β β and GK α β yielded crystals useful for X-ray work whereas the GK α α yielded only stacks of thin needles that could not be improved further.

Crystals of GK β β bound with the TSA comprising glycocyamine, Mg²⁺, ADP, and NO₃⁻, were obtained by the vapor diffusion method from reservoir solution containing 22.5% polymethylmetacrylate 2000, 100 mM calcium acetate, and 100 mM Tris-HCl (pH 8). The TSA was added to the protein solution to achieve final concentration of 5 mM each MgCl₂, ADP, and NaNO₃, and 1.25 mM glycocyamine. Rod-shaped crystals of 0.1 x 0.1 x 0.2 mm³ in size appeared within 3 days. Efforts to obtain GK α β /TSA crystals were unsuccessful.

The crystals were flash-cooled in liquid nitrogen or in liquid propane cooled with liquid nitrogen. The 2.4 Å diffraction data from a ligand-free GK β β crystal were collected at ~100 K on the home facility consisting of Osmic Max-Flux monochromated X-rays generated by a Rigaku MicroMax 007 rotating anode equipped with a Rigaku AXIS IV⁺⁺ image plate detector. These GK β β diffraction data were processed with the CrystalClear program (Rigaku/MSO, the Woodlands, TX). For the ligand-free GK α β , 2.6 Å data were collected at the SER-CAT 22BM beamline of the Advanced Photon Source (Argonne National Laboratory, Argonne, Illinois). The SER-CAT beamline was equipped with a MAR CCD detector. The HKL suite of programs was used to process the GK α β diffraction data (30). The 2.3 Å resolution diffraction data from a crystal of the GK β β -TSA complex were collected on the SER-CAT 22ID beamline and processed with the XGEN program (31). Data collection statistics are provided in Table 1.

Structure determination and refinement

The GK $\beta\beta$ structure was determined by the Molecular Replacement method with the computer program Phaser (32, 33) using a monomer of the chicken brain-type CK crystal structure as the search model ((15); PDB code 1QH4, 56 % sequence identity with GK). The N-termini were built de-novo because they adopt two different conformations, none of which is similar to that of the CK N-termini. Later, the refined GK $\beta\beta$ structure was used as the search model to obtain a Molecular Replacement solution for the GK $\alpha\beta$, again with the program Phaser.

The structure of the GK $\beta\beta$ -TSA complex was determined by Molecular Replacement using Phaser. A search model based on the structure of GK $\beta\beta$ in the unbound form did not yield a useful Molecular Replacement solution. Instead, the rabbit muscle CK-TSA structure (creatine-containing TSA, (17)) was modified to generate a search model as follows. The rabbit muscle CK-TSA structure forms an asymmetric dimer with one subunit containing the TSA and exhibiting closed conformation and the second subunit containing only Mg²⁺-ADP and exhibiting an open conformation. This asymmetric dimer did not yield a useful Molecular Replacement solution. The search model that produced the Molecular Replacement solution was obtained by generating a CK dimer in which both subunits adopted the closed conformation and the ligands omitted to avoid phase bias. The entire asymmetric unit, containing 18 dimers, was built gradually and the progress of the Molecular Replacement was assessed by the reduced R_{free} value (<0.40 when all 18 dimers were placed in the asymmetric unit) and the improved quality of the electron density map.

Structure refinements were performed initially with the CNS program (34) followed by either REFMAC (35) or PHENIX (36). For the ligand-bound GK $\beta\beta$ structure, the R_{free} and R_{work} values were reduced substantially when twinning refinement was applied in PHENIX with the twinning operator $(-h, -k, l)$. The estimated twinning fraction was 0.31. In contrast, the diffraction data for the substrate-free GK $\alpha\beta$ and GK $\beta\beta$ crystals did not indicate significant twinning and in both cases the structure refinement was completed with REFMAC. Water molecules were added to account for peaks in the $F_o - F_c$ electron density map higher than 3σ and observed within hydrogen bond distance from protein atoms. Model building and modifications were performed with the interactive graphics programs O (37) and COOT (38). Refinement statistics are provided in Table 1.

It worth noting that the diffraction data for the crystals of the GK $\beta\beta$ -TSA complex were processed in space group $P2_1$, which yields 18 dimers in the asymmetric unit, however the data could also be processed using the higher symmetry space group, $P2_12_12_1$. Although a Molecular Replacement solution was obtained also in the orthorhombic cell, the reduced symmetry (coupled with twinning) was required for the structure refinement to progress.

Figures were generated with MOLSCRIPT (39) followed by RASTER3D (40, 41). Conformational changes were analyzed with the computer program DynDom (42).

Results and Discussion

Note on sequence numbering

The GK α and β polypeptide chains are identical except in the N-terminal alternatively spliced segments (29):

```

•           MFKDYTREKFAKE.....YE...
•  MGSATQDYFVKNRVGHSGKFWESGKFKAAD.....KE...
      5      10      15      20      25      48

```


Exon 1 of the GK gene encodes the 48 N-terminal amino acids of GK β and exon 2 encodes the 32 N-terminal amino acids of GK α . The first residue of the longer β chain is numbered 1 and the first residue of the shorter α chain is numbered 17. As depicted above, the N-terminal sequences of the α and β GK variants differ over the first 29 amino acid residues, the ensuing 17 residues are identical, position 47 differs, and position 48 is identical. Beyond exons 1 and 2, the amino acid sequences of the GK α and β isoforms are encoded by the same exons and are thus identical.

GK purification

The expression constructs developed for crystallographic studies consist of a TEV protease cleavable His₆-MBP-GK fusion protein, which enables the application of an affinity purification step followed by cleavage of the fusion protein, separation of the His-tagged MBP, the uncleaved fusion protein, and the His-tagged TEV protease from the cleaved GK, and further purification of the GK by size exclusion chromatography. During the size exclusion step, most of the GK $\alpha\alpha$ sample formed high aggregates that were eluted in the void volume while the yield of the dimeric protein was low. In contrast, most of the GK $\beta\beta$ sample was eluted as dimers with high aggregates comprising the minor form. These results are consistent with previous reports that the MBP-GK $\alpha\alpha$ fusion protein was unstable (29). The GK $\alpha\alpha$ tendency to aggregate may also account for failure to form diffraction quality crystals. Nevertheless, the analytical size exclusion analysis and the dynamic light scattering of the non-aggregated fractions of both GK variants were consistent with dimeric association. The analytical size exclusion also shows that GK $\alpha\beta$ migrates differently than either of the two homodimers, consistent with spontaneous formation of the heterodimer from a 1:1 mixture of GK $\alpha\alpha$ and GK $\beta\beta$.

Overall structure

The crystals of apo GK $\alpha\beta$ and GK $\beta\beta$ are isomorphous and contain two dimers in the asymmetric unit. Except for the N-termini, which are fully defined in GK $\alpha\beta$ but exhibit some disorder in GK $\beta\beta$ (no interpretable electron density is associated with the 24 N-terminal residues of one subunit and 11 N-terminal residues of the second subunit), the structures are the same (Fig. 2A & B). The root-means-squared deviation (RMSD) of the common α -carbon residues 27-331 and 343-390 is 0.6 Å. As with other phosphagen kinase family members, the molecule consists of an N-terminal α -helical domain (residues 12-114) and a larger C-terminal α/β saddle domain (residues 115-390). The subunits' N-termini form much of the dimer interface, although the C-terminal domain loop residues 156-160 and 220-221 are also involved in subunit contacts. Different in length and sequence, the N-termini of the α - and β -subunits in the GK $\alpha\beta$ dimer adopt different conformations, which result in an asymmetric interface. Moreover, the two N-termini in the GK $\beta\beta$ homodimer also form an asymmetric interface despite their identical N-termini sequence. As discussed later in more detail, the interfaces of both structures are similar except for the partial disorder in the GK $\beta\beta$ N-termini.

Binding of the TSA, Mg²⁺-ADP-NO₃-glycocyamine to GK $\beta\beta$ (NO₃ mimics a planar phosphoryl group during the phosphotransfer reaction) induces a rigid body like conformational transition that changes the relationship between the α -helical and α/β domains so that the active site crevice is less accessible to bulk solvent and the two lobes of each subunit are closer to one another (Fig. 2C). The dimer interface is not affected, thus the rigid body movement is best envisaged as shifting the C-terminal α/β domains with respect to the fixed N-terminal α -helical domains. In the following discussion, we refer to the ligand-free GK conformation as the open conformation, and to the enzyme conformation induced by binding of the TSA as the closed conformation.

The computer program DynDom was used to analyze more accurately the quasi rigid body domain motion (42). The DynDom analysis yields an effective hinge axis, which is defined as an inter-domain screw axis that passes within 3 Å of the C α atom of any of the inter-domain residues involved in the inter-domain motion. Using the GK β subunit structure in the open (ligand free) and closed (ligand bound) states, the DynDom analysis defines a fixed domain that encompasses residues 27-115, 141-249, and 264-302, corresponding to the lobe close to the dimer interface. The domain that moves relative to the fixed domain encompasses residues 116-140, 250-263, and 303-388, and corresponds to the lobe remote from the dimer interface. This analysis yielded a 22° rotation around an axis approximately parallel to the dimer interface and traversing the β -sheet underlying the active site crevice (Fig. 3). The effective hinge axis has essentially no translation component (0.04 Å). Using DynDom analysis, effective hinge axes also describe the CK and AK conformational transition with 13° and 20° rotation, respectively and no significant translation component.

Two active site loops, an N-terminal domain loop (residues 75-84) involved in glycoamine binding (referred to as the phosphagen-specificity loop, Fig. 4) and a C-terminal domain loop (residues 333-342) involved primarily in ATP/ADP binding (referred to as the nucleotide-binding loop, Fig. 4), are positioned remotely from one another in the open conformation, whereas they interact with one another in the closed conformation (colored brown and magenta, respectively, in Fig. 2). In the ligand-free state, the nucleotide-binding loop exhibits multiple conformations, which may arise from the different crystal environments among the four molecules in the asymmetric unit.

TSA binding

Previous studies demonstrated that the combination of Mg²⁺, ADP, nitrate, and phosphagen comprise a transition state analog that inhibits the catalytic activities of phosphagen kinases. In such a TSA, the planar nitrate serves as a mimic of the planar phosphoryl group during phosphotransfer (43-45). Indeed, this complex was used in crystallographic studies of CKs and AK (16, 17, 21, 46-48), which confirmed that the nitrate is stationed between the ADP and the phosphagen. We followed the same strategy in studying the structure of the GK-TSA complex, using Mg²⁺-ADP, nitrate, and glycoamine (Fig. 5). The electron density map showed that both GK $\beta\beta$ subunits contained the bound TSA in all 18 dimers of the crystal asymmetric unit. Superposition of C α atoms of all dimer combinations resulted in a RMSD value of 0.5 $\pm$ 0.2 Å, with 0.9 Å for the largest deviating dimer pair. The structural variability occurs primarily in the C-terminal α -helix that tends to adopt slightly different orientations with respect to the rest of the structure.

In contrast to the GK-TSA complex, where both subunits bind the glycoamine TSA, the three reported CK-TSA complexes (from rabbit muscle, *T. californica*, and human brain-type) bind the creatine TSA only in one subunit, which adopts the closed conformation. The second subunit of the rabbit muscle and *T. californica* binds Mg²⁺-ADP (16, 17) whereas that of the human brain-type enzyme is ligand-free (21). In all these CK complexes the second subunit adopts the open conformation. It is also important to note that in contrast to the *T. californica* CK-TSA complex, the closed conformation has been recently observed also in the Mg²⁺-ADP bound subunit of human brain-type CK (21). The asymmetric binding of the TSA to CK has been viewed as supporting the notion of negative cooperativity (16), although the mechanism by which this occurs is not clear. The GK structures retain the same asymmetric interface in both the open conformation and when both subunits adopt the closed conformations. Thus there is no structural support for negative cooperativity in this enzyme.

The CK and GK subunits that bind the TSA undergo similar global conformational changes. This is easily discernible when N-terminal domain segments of the TSA-free and TSA-

bound molecules are superposed. The N-terminal domain segments encompassing residues 26-76 in GK and 14-64 in CK serve as the reference framework in this comparison because they mediate the dimer interface and remain unaltered in the two states. The analysis of the AK conformational change includes a 10 residues shorter N-terminus (residues 24-61) as the reference unit because this monomeric protein contains an N-terminal α -helix absent in GK and CK. The analyses show that the shifts are most pronounced in the phosphagen specificity loop, α -helices 7 and 8 on the C-terminal domain, and the ensuing β -strand except that the shortest phosphagen specificity loop of AK does not change upon TSA binding (Fig 6). AK exhibits the greatest displacement of α 7 perhaps because it is a monomeric enzyme and not restricted by interaction with a partner subunit (in the dimeric enzymes the N-terminus of α 7 is placed in close proximity to the N-terminal region of the partner subunit). Concomitant with the conformational transition of secondary structure units, the side chains of residues that must accommodate these global changes adjust as well, for example Thr188, Lys207 and Pro208 in GK and the respective Ser178, Lys196 and Pro197 in CK (Fig 6).

In a phosphotransfer that follows the associative mechanism (49), the transition state places the transferred phosphorous such that it forms pentagonal bipyramidal geometry with the donor and acceptor groups located in the apical positions ~ 2 Å from the phosphorous atom. The postulated transition state in the glycocyamine kinase reaction is depicted schematically in Figure 1. The nitrate in the GK-TCA complex mimics the planar phosphoryl group of the transition state (Figs. 5 & 6) as its nitrogen atom is positioned approximately midway between the guanidino group of the glycocyamine and the β -phosphoryl of ADP. Modeling shows that modest rotations around the ADP α and β phosphates bonds bring the β -phosphoryl group in line with the nitrate and with a guanidino N atom phosphoryl acceptor of glycocyamine. A similar arrangement occurs in the CK and AK with their respective TSA complexes (Fig. 6).

The Mg^{2+} cofactor forms octahedral coordination with the nitrate ion, the P_{α} and P_{β} phosphates of ADP and with three water molecules (Fig. 5). The coordinating waters molecule, in turn, form hydrogen bonds with the carboxylate groups of Glu242, Glu243 and Glu336. Similar Mg^{2+} coordination is present in the CK-TSA and AK-TSA complexes. In addition to the Mg^{2+} , the active site is enriched with arginine residues that interact with the phosphates of ADP and with the NO_3 , Arg141, Arg143, Arg247, Arg303, and Arg331 (Fig. 5). In turn (not shown in Fig. 5), Asp244 interacts with Arg247, Asp346 interacts with Arg303, Asn297 interacts with Arg143, and all these residues are conserved in the CK and AK amino acid sequences. Other conserved residues that interact with the nucleotide are His202 that is hydrogen bonded to the ADP ribose C(2)OH and Ser139 that is hydrogen bonded to the adenine's N1. The interaction of Ser305 with the adenine's NH2 and the N6 occurs also in AK, whereas a glycine residue occupies the same position in CK and instead a water molecule replaces the Ser305 hydroxyl group. The nucleotide-binding loop residues 333-336 flank the ribose and phosphate of ADP and contribute backbone interactions; Thr533 and Gly534 CO groups form hydrogen bonds with the ribose C(2)OH and C(3)OH, respectively. These two residues are conserved to maintain the structural integrity of the nucleotide-binding loop; Thr533 hydroxyl group is hydrogen bonded to the backbone NH and the hydroxyl group of Thr341, and the backbone of Gly335 (conserved in all phosphagen kinase sequences) adopts dihedral angles that lead to steric strain in C_{β} -containing residues. Notably, the nucleotide-binding loop is also involved in phosphagen binding specificity through Glu336 as discussed in the following section. Finally, the π -electron stacking interactions of His307 imidazole ring and Arg141 guanidinium group occur also in CK and AK.

The GK interactions with the glycoamine are fewer than the interactions with the nucleotide (Fig. 5). Two of these interactions are conserved in all phosphagen kinases of known structure; Glu243 and Cys294 interact with the glycoamine guanidinium group. The glycoamine carboxylate group forms a hydrogen bond with the backbone NH of Thr84 located on the phosphagen-specific loop, a conserved interaction in CK and AK. The remaining interactions are phosphagen specific. Glu336 interacts with the glycoamine guanidine group. Lys83 and Thr84 of the phosphagen-specific loop flank the glycoamine. The Thr84 hydroxyl group is hydrogen bonded with the glycoamine carboxylate group. The side chain of Lys84 contributes to the sequestering of the glycoamine from solvent but its amino group does not form an ion pair with the substrate's carboxyl group.

Determinants of phosphagen specificity

Both glycoamine and arginine lack the internal guanidinium N-methyl present in the creatine molecule. A glutamate residue conserved in GK and AK but missing in CK defines the specificity towards a non-substituted guanidinium group by engaging it in a salt bridge (Glu336 and Glu314 in GK and AK, respectively). CK displays in the equivalent position a smaller and hydrophobic residue (Val325), which complements the creatine's N-methyl substituent (Fig. 7) (16). The other end of the phosphagen molecule interacts with the phosphagen-specificity loop, which is longest for the smallest phosphagen, glycoamine, and the shortest for the largest phosphagen, arginine (Fig. 4 & 7). Here, the interaction of GK and CK with the phosphagen's carboxylate group is the same, employing the Thr84 or Val72 backbone NH, respectively). A shorter AK loop is required to accommodate the additional 4 atoms of the arginine substrate. Nevertheless, the backbone interaction with the carboxylate group is conserved in AK (*via* Val65 NH), and in addition, the Gly64 NH also engages the arginine carboxylate group in a hydrogen bond interaction. In AK, the amino group of the arginine substrate interacts with Tyr68 hydroxyl group, located on the phosphagen-specific loop, and with an internal water molecule that in turn is hydrogen bonded to Ser63, also an AK phosphagen-specific loop residue. Thus, Tyr68 and Ser63 are determinants of AK activity with the exception of the two dimeric enzymes from the phylum Echinodermata (sea cucumber and sea urchin) where the sequences contain two hydrophobic residues (Phe and Leu) in the analogous positions (Fig. 4). Presumably, in these cases the enzymes employ a different strategy to complement the sequestered arginine amino group.

Figure 7 shows that the nucleotide-binding and phosphagen-specific loops in GK and CK interact with one another in the closed conformation. One of the CK conserved phosphagen-specificity loop residues, His66, interacts with the nucleotide-binding loop, Asp326, conserved in all CKs. Site directed mutagenesis studies of the human CK support the importance of this interaction in catalysis (50). In GK, the phosphagen-specificity loop is longer than that of CK or AK. Closure of the GK loops brings Phe79 and Tyr80 on the N-terminal domain in contact with Leu339 on the C-terminal domain, which "locks" the GK closed conformation. These residues are conserved in the few known GK sequences, and distinguish the GK from CK.

The short phosphagen-specificity loop of AK does not reach the nucleotide-binding loop (Fig. 7C). Instead, both loops exploit interactions with α helix 9 to establish the closed conformation; Asp62 forms an ion pair with Arg193, and the His315 imidazole ring stacks against the benzene ring of Phe194 (His in the Echinodermata AKs; Fig. 4). The GK and CK loops lack equivalent interactions, therefore the AK Asp62, Arg193, Phe194, and His315 may be considered signature residues of AK activity. Currently, only a few AK sequences have been reported. As more sequences become available, other residue combinations that anchor the loops may emerge, and the validity of these four AK signature residues may need to be revised.

It is important to emphasize that although the structure-based signature residues discussed above help identify the substrate specificity of phosphagen kinases, by no means is it possible to switch enzyme specificity solely by exchanging the relevant residues. This has been demonstrated by the mutagenesis studies of Jourden *et al* (51). The authors replaced the CK valine signature residue discussed above by the GK glutamate signature residue and exchanged the phosphagen-specific loop alone or in combination with the Val to Glu replacement. All mutants impaired both CK and GK activities dramatically and the huge reduction in k_{cat} values rules out conclusions about specificity change. Presumably, as proteins evolve in nature, specificity changes are followed by further sequence changes. Some changes optimize the emerging new activity and some are random drift from the ancestral molecule, ultimately arriving at a structure that cannot be reversed to perform the ancestral function by simply changing the signature sequences. Indeed, although CK and GK exhibit over 50% sequence identity, there is ample opportunity to drift away from the ancestral enzyme in the remaining residues.

Mutagenesis studies of AK also demonstrate that structures evolve further after function changes occur (52). In this case, the entire AK loop containing the Glu314 (analogous to GK Glu336) was replaced with the analogous CK loop, so that position 314 was occupied by a valine, a proposed determinant of CK activity. This replacement and four additional replacements all resulted in smaller side chains compared with the AK amino acid residues. No CK kinase activity was observed whereas the AK activity was reduced only slightly. Apparently, other protein groups that interact with the arginine substrate contribute sufficiently to binding to retain the authentic AK activity. Lacking a crystal structure of the five-residue substituted AK mutant, it remains unknown whether the grafted loop adopts the same conformation as seen in the CK structure, which would be expected for an enzyme that switches substrate specificity.

The N-termini mediate dimer assembly

Mutually exclusive alternative splicing of the two N-terminal exons produces the α and β variants of GK (Fig. 4). The substrate-free GK $\alpha\beta$ heterodimer structure shows that the two N-termini complement one another through hydrophobic interactions that contribute to the dimer interface (Fig. 8A). The longer N-terminus of the β chain in the GK $\alpha\beta$ structure is fully visible in the electron density map and traverses nearly the entire dimer interface while the shorter N-terminus of the α chain is also fully ordered and located on one side of the dimer interface. The two termini interact only through their respective first α -helices (labeled $\alpha 1a$ and $\alpha 1b$ in Fig. 4) such that Ile5 and Phe9 of the β chain form hydrophobic interactions with Phe26 of the α chain (Fig. 8A).

The spontaneous formation of GK $\alpha\beta$ from a 1:1 mixture of GK $\alpha\alpha$ and GK $\beta\beta$ suggests that the heterodimer is the most stable form. Indeed, the GK $\alpha\beta$ interface is better packed than the GK $\beta\beta$. The N-termini of the GK $\beta\beta$ homodimer exhibit substantial disorder; one chain (the counterpart of the α chain in GK $\alpha\beta$) is visible only from residue 24 (instead of residue 18 as defined in the α chain of the heterodimeric structure) and the second chain (analogous to the β chain of the GK $\alpha\beta$) is visible only from residue 12. This structural disorder is attributed to the structural incompatibility arising from two identical sequences instead of the complementing α and β sequences.

The catalytic rates of GK are comparable with those of CK, and the activity of recombinant GK $\beta\beta$ is similar to that of native GK $\alpha\beta$ (29). Thus, the disorder of the GK $\beta\beta$ N-termini does not affect enzyme activity. In contrast, GK $\alpha\alpha$ tends to aggregate and, in our hands, did not yield single crystals useful for X-ray work. A model of a GK $\alpha\alpha$ dimer using GK $\alpha\beta$ or GK $\beta\beta$ as the template for dimer association shows much reduced contact between the two

subunits. Thus, if as proposed by Ellington and colleagues the ancestral enzyme consisted of GK $\alpha\alpha$ (6), the interface of that ancestral protein could not resemble the current interface.

Interestingly, several but not all dimers in the crystal asymmetric unit of the GK $\beta\beta$ -TSA complex (containing 18 dimers) exhibit more order in the interface-traversing β chain compared with the corresponding β chain of apo GK $\beta\beta$. For two of these dimers, the electron density map accounts the complete α -helix 1b except for the first couple of residues (Fig 8B). This α -helix is misaligned by 25° with the α 1b of GK $\alpha\beta$, presumably because of incompatibility in packing against the N-terminus of the partner β chain. As a result of the α -helix shift, Phe9 packs against Trp54 of the partner protein in contrast to GK $\alpha\beta$, where it is Tyr8 that interacts with Trp54.

The interface-traversing β chain in all versions of GK $\beta\beta$ and GK $\beta\beta$ -TSA complex is well defined in the electron density map beginning with Arg13. Arg13 forms a salt bridge with Glu161 of the same β subunit, an interaction that is independent of the second subunit (Fig. 8A&B). Interactions of residues beyond Arg13 involve the partner subunit beyond residue 47, which are identical in both α and β isoforms. The Arg13/Glu161 salt bridge occurs only in one subunit whereas Glu161 of the partner subunit interacts with Ser17, i.e. one Glu161 is involved directly in dimer formation and the second is not (Fig. 8A&B).

Both hydrophobic and hydrophilic interactions mediate the contact between the two subunits. There are also water molecules in the interface that bridge the subunits. An inter subunit salt bridge between Asp65 and Arg159 (occurring twice) is a contact conserved in both GK and CK (the corresponding CK residues are Asp54 and Arg148). The Asp/Arg pair is also conserved in the sequences of the dimeric AK from sea cucumber and sea urchin but not in the sequences of the monomeric AK enzymes (Fig. 4), which suggests that the presence of these residues may serve to distinguish between dimeric and monomeric phosphagen kinases.

The N-termini of the CK homodimers are also located at the subunit interface and they are fully or partially ordered in the various structures deposited in the PDB. Two examples of interfaces are depicted in Figure 8C&D. When visible, the arrangement is asymmetric despite the identical sequence of each subunit. Unlike the GK $\alpha\beta$ arrangement, where the N-terminus of the β chain mediates most of the inter subunit contacts, the shorter two N-termini of CK mediate inter subunit contacts to the same extent. Nevertheless, the CK residue equivalent to GK Glu161, Glu150, also interacts asymmetrically with the N-terminus except that Glu150 in each subunit interacts with a residue on the partner molecule. For the rabbit muscle CK one Glu150 interacts with Tyr10 (this interaction is not seen in *T. californica* CK because of a larger extent of N-terminal disorder) and the second Glu150 interacts with Asn8 (and with Trp10 in *T. californica* CK) (Fig. 8C&D). Taken together, the structural data show that CK and GK evolved different dimerization properties despite the evolutionary relationship they share in terms of sequence, function, overall structure, and the employment of the same dimerization strategy, which exploits the N-termini to generate an asymmetric interface.

Acknowledgments

We thank Buvaneshwari Narayanan for help with data collection. We also thank the staff at SER-CAT in Argonne National Laboratory for their assistance during the data collection.

References

1. Ellington WR. Evolution and physiological roles of phosphagen systems. *Ann Rev Physiol.* 2001; 63:289–325. [PubMed: 11181958]

2. McLeish MJ, Kenyon GL. Relating structure to mechanism in creatine kinase. *Crit Rev Biochem Mol Biol.* 2005; 40:1–20. [PubMed: 15804623]
3. Suzuki T, Furukohri T. Evolution of phosphagen kinase: primary structure of glycocyamine kinase and arginine kinase from invertebrates. *J Mol Biol.* 1994; 237:353–357. [PubMed: 8145248]
4. Morrison, JF. Arginine kinase and other invertebrate guanidino kinases. In: Boyer, PC., editor. *The Enzymes*. Academic Press; New York: 1973. p. 457–486.
5. Van Thoai, N. Homologous phosphagen phosphokinases. In: Van Thoai, N.; Roche, J., editors. *Homologous Enzymes and Biochemical Evolution*. Gordon and Breach; New York: 1968. p. 199–229.
6. Ellington WR, Yamashita D, Suzuki T. Alternative splicing produces transcripts coding for alpha and beta chains of a hetero-dimeric phosphagen kinase. *Gene.* 2004; 334:167–174. [PubMed: 15256266]
7. Ellington WR. Phosphocreatine represents a thermodynamic and functional improvement over other muscle phosphagens. *J Exp Biol.* 1989; 143:144–194.
8. Arndt C, Schiedek D. *Nephtys hombergii*, a free-living predator in marine sediments: energy production under environmental stress. *Marine Biol.* 1997; 129:643–650.
9. Sona S, Suzuki T, Ellington WR. Cloning and expression of mitochondrial and protoflagellar creatine kinases from a marine sponge: implications for the origin of intracellular energy transport systems. *Biochem Biophys Res Comm.* 2004; 317:1207–1214. [PubMed: 15094398]
10. Tanaka K, Uda K, Shimada M, Takahashi K, Gamou S, Ellington WR, Suzuki T. Evolution of the cytoplasmic and mitochondrial phosphagen kinases unique to annelid groups. *J Mol Evol.* 2007; 65:616–625. [PubMed: 17932618]
11. Suzuki T, Kamidochi M, Inoue N, Kawamichi H, Yazawa Y, Furukohri T, Ellington WR. Arginine kinase evolved twice: evidence that echinoderm arginine kinase originated from creatine kinase. *Biochem J.* 1999; 340:671–675. [PubMed: 10359650]
12. Suzuki T, Mizuta C, Uda K, Ishida K, Mizuta K, Sona S, Compaa DM, Ellington WR. Evolution and divergence of the genes for cytoplasmic, mitochondrial, and flagellar creatine kinases. *J Mol Evol.* 2004; 59:218–226. [PubMed: 15486695]
13. Pineda AO, Ellington WR. Structural and functional implications of the amino acid sequences of dimeric, cytoplasmic and octameric mitochondrial creatine kinases from a protostome invertebrate. *Eur J Biochem.* 1999; 264:67–73. [PubMed: 10447674]
14. Kalds P, Wallimann T. Functional differences between dimeric and octameric mitochondrial creatine kinase. *Biochem J.* 1995; 308:623–627. [PubMed: 7772050]
15. Eder M, Schlattner U, Becker A, Wallimann T, Kabsch W, Fritz-Wolf K. Crystal structure of brain-type creatine kinase at 1.41 Å resolution. *Protein Sci.* 1999; 8:2258–2269. [PubMed: 10595529]
16. Lahiri SD, Wang PF, Babbitt PC, McLeish MJ, Kenyon GL, Allen KN. The 2.1 Å structure of Torpedo californica creatine kinase complexed with the ADP-Mg(2+)-NO(3)(-)-creatine transition-state analogue complex. *Biochemistry.* 2001; 41:13861–13867. [PubMed: 12437342]
17. Ohren JF, Kundracik ML, Borders CL, Edmiston P, Viola RE. Structural asymmetry and intersubunit communication in muscle creatine kinase. *Acta Cryst.* 2007; D63:381–389.
18. Rao JK, Bujacz G, Wlodawer A. Crystal structure of rabbit muscle creatine kinase. *FEBS letters.* 1998; 439:133–137. [PubMed: 9849893]
19. Shen YQ, Tang L, Zhou HM, Lin ZJ. Structure of human muscle creatine kinase. *Acta crystallographica.* 2001; 57:1196–1200.
20. Tisi D, Bax B, Loew A. The three-dimensional structure of cytosolic bovine retinal creatine kinase. *Acta crystallographica.* 2001; 57:187–193.
21. Bong SM, Moon JH, Nam KH, Lee KS, Chi YM, Hwang KY. Structural studies of human brain-type creatine kinase complexed with the ADP-Mg²⁺-NO₃⁻-creatine transition-state analogue complex. *FEBS letters.* 2008; 582:3959–3965. [PubMed: 18977227]
22. Degani C, Degani Y. Further evidence for nonsymmetric subunit association and intersubunit cooperativity in creatine kinase. Subunit-selective modifications by 2,4-dinitrophenylthiocyanate. *The Journal of biological chemistry.* 1980; 255:8221–8228. [PubMed: 7190972]

23. Degani Y, Degani C. Subunit-selective chemical modifications of creatine kinase. Evidence for asymmetrical association of the subunits. *Biochemistry*. 1979; 18:5917–5923. [PubMed: 574773]
24. Hornemann T, Rutishauser D, Wallimann T. Why is creatine kinase a dimer? Evidence for cooperativity between the two subunits. *Biochimica et biophysica acta*. 2000; 1480:365–373. [PubMed: 10899637]
25. Nevinsky GA, Ankilova VN, Lavrik OI, Mkrtchyan ZS, Nersesova LS, Akopyan JI. Functional non-identity of creatine kinase subunits of rabbit skeletal muscle. *FEBS letters*. 1982; 149:36–40. [PubMed: 7152033]
26. Mazon H, Marcillat O, Forest E, Vial C. 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*. 2003; 42:13596–13604. [PubMed: 14622006]
27. Wang XC, Zhou HM, Wang ZX, Tsou CL. Is the subunit the minimal function unit of creatine kinase? *Biochimica et biophysica acta*. 1990; 1039:313–317. [PubMed: 2378889]
28. Nallamsetty S, Austin BP, Penrose KJ, Waugh DS. Gateway vectors for the production of combinatorially-tagged His6-MBP fusion proteins in the cytoplasm and periplasm of *Escherichia coli*. *Protein Sci*. 2005; 14:2964–2971. [PubMed: 16322578]
29. Mizuta C, Tanaka K, Suzuki T. Isolation, characterization, and cDNA-derived amino acid sequence of glycocyamine kinase from the tropical marine worm *Namalycastis* sp. *Comp Biochem Physiol B*. 2005; 140:387–393. [PubMed: 15694586]
30. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Meth Enzymol*. 1997; 276:307–326.
31. Howard, AJ. Data processing in macromolecular crystallography. In: Bourne, PE.; Watenpugh, KD., editors. *Crystallographic Computing 7: Proceedings from the Macromolecular Crystallographic Computing School*. Oxford University Press; Oxford: 2000.
32. McCoy AJ, Grosse-Kunstleve RW, Storoni LC, Read RJ. Likelihood-enhanced fast translation functions. *Acta Cryst*. 2005; D61:458–464.
33. Storoni LC, McCoy AJ, Read RJ. Likelihood-enhanced fast rotation functions. *Acta Cryst*. 2004; D60:432–438.
34. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst*. 1998; D54:905–921.
35. Winn MD, Isupov MN, Murshudov GN. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Cryst*. 2001; D57:122–133.
36. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. *Acta Cryst*. 2002; D58:1948–1954.
37. Jones TA. Interactive electron-density map interpretation: from INTER to O. *Acta Cryst*. 2004; D60:2115–2125.
38. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Cryst*. 2004; 60:2126–2132.
39. Kraulis PJ. A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst*. 1991; 24:946–950.
40. Bacon DJ. A fast algorithm for rendering space-filling molecule pictures. *J Mol Graph*. 1988; 6:219–220.
41. Merritt EA, Bacon DJ. Raster3D: Photorealistic molecular graphics. *Meth Enzymol*. 1997; 277:505–524. [PubMed: 18488322]
42. Hayward S, Berendsen HJ. Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins*. 1998; 30:144–154. [PubMed: 9489922]
43. Held BC, Wright-Weber B, Grossman SH. Kinetic analysis of two purified forms of arginine kinase: absence of cooperativity in substrate binding of dimeric phosphagen kinase. *Comp Biochem Physiol B*. 2007; 148:6–13. [PubMed: 17572125]

44. Milner-White EJ, Watts DC. Inhibition of adenosine 5'-triphosphate-creatine phosphotransferase by substrate-anion complexes. *Biochem J.* 1971; 122:727–740. [PubMed: 5129268]
45. Borders CL, Snider MJ, Wolfenden R, Edmiston PL. Determination of the affinity of each component of a composite quaternary transition-state analogue complex of creatine kinase. *Biochemistry.* 2002; 41:6995–7000. [PubMed: 12033932]
46. Yousef MS, Clark SA, Pruet PK, Somasundaram T, Ellington WR, Chapman MS. Induced fit in guanidino kinases - comparison of substrate-free and transition state analog structures of arginine kinase. *Protein Sci.* 2003; 12:103–111. [PubMed: 12493833]
47. Yousef MS, Fabiola F, Gattis JL, Somasundaram T, Chapman MS. Refinement of the arginine kinase transition-state analogue complex at 1.2 Å resolution: mechanistic insights. *Acta crystallographica.* 2002; 58:2009–2017.
48. Zhou G, Somasundaram T, Blanc E, Parthasarathy G, Ellington WR, Chapman MS. Transition state structure of arginine kinase: implications for catalysis of bimolecular reactions. *Proc Natl Acad Sci U S A.* 1998; 95:8449–8454. [PubMed: 9671698]
49. Knowles JR. enzyme-catalyzed phosphoryl transfer reactions. *Annu Rev Biochem.* 1980; 49:877–919. [PubMed: 6250450]
50. Wang PF, Flynn AJ, McLeish MJ, Kenyon GL. Loop movement and catalysis in creatine kinase. *IUBMB life.* 2005; 57:355–362. [PubMed: 16036620]
51. Jourden MJ, Clarke CN, Palmer AK, Barth EJ, Prada RC, Hale RN, Fraga D, Snider MJ, Edmiston PL. Changing the substrate specificity of creatine kinase from creatine to glycocyamine: evidence for a highly evolved active site. *Biochimica et biophysica acta.* 2007; 1774:1519–1527. [PubMed: 17976392]
52. Azzi A, Clark SA, Ellington WR, Chapman MS. The role of phosphagen specificity loops in arginine kinase. *Protein Sci.* 2004; 13:575–585. [PubMed: 14978299]
53. Sonnhammer EL, Hollich V. Scoredist: a simple and robust protein sequence distance estimator. *BMC Bioinformatics.* 2005; 6:108. [PubMed: 15857510]

Figure 1.
The reaction catalyzed by GK.

Figure 2.

Ribbon representations of GK structures. (A) Ligand-free GK $\alpha\beta$ heterodimer. (B) Ligand-free GK $\beta\beta$ homodimer. (C) The GK $\beta\beta$ homodimer bound with glycocyamine, Mg^{2+} , ADP, and NO_3^- (transition state analog). For each dimer, the α subunit (or the α -equivalent β subunit in the GK $\beta\beta$) is colored yellow, and the β subunit is colored cyan. The N-terminus of the α subunit is colored red and N-terminus of the β subunit is colored blue. The phosphagen-specific loop (brown color) and nucleotide-binding loop (magenta color) are far apart in the ligand-free state (A & B) and are placed in close proximity in the ligand-bound structure. The nucleotide binding loop is partially disordered in the β subunit of the ligand-free GK $\alpha\beta$ (A).

Figure 3.

Stereoscopic representation of the GK overall conformational change displayed together with the effective hinge calculated by the program DynDom (42). The open conformation is depicted in gray color. The fixed domain of the closed conformation is colored cyan and the moving domain is colored yellow. The arrow corresponds to the effective hinge axis around which the moving domain rotates 22° upon TSA binding.

Figure 4.

Amino acid sequence alignment of GK, CK, and AK. GK secondary structure elements are indicated above the sequence blocks as cylinders for α helices and arrows for β strands. The phosphagen-specificity loop and the nucleotide binding loop are indicated. The GK α and β N-termini sequences are different. Beyond residue 47 (β isoform numbering) the α and β isoforms exhibit identical sequences. The following sequences are aligned below the GK sequence: Rabbit muscle CK (rmCK), *T. californica* CK (electric ray, erCK), chicken brain CK (bbCK), horseshoe crab AK (hsAK), *Trypanosoma cruzi* AK (tcAK), sea cucumber (scAK) and from sea urchin (suAK). The hsAK and tcAK are monomeric proteins whereas scAK and suAK are dimeric proteins. Residue numbers are provided for AK and CK above the respective block. Amino acid residues are shaded according to BLOSUM62 scores of 3.0 (cyan), 1.5 (pink), and 0.5 (yellow) (53).

Figure 5.

Stereoscopic representation of the binding of the TSA (Mg^{2+} -ADP- NO_3 -glycocyamine) to $\text{GK}\beta\beta$. The $F_o - F_c$ electron density map was calculated prior to the inclusion of the TSA and water molecules in the model and is contoured at 2.5σ level. The phosphagen-specific loop and the nucleotide-binding loop are indicated and colored brown and magenta, respectively. Key interactions with the ADP, NO_3^- , and glycocyamine are shown.

Figure 6.

Global conformational changes associated with TSA binding in phosphagen kinases. The N-terminal domains in the ligand-free and ligand-bound structures were superposed as described in the text and shifting structural units in the vicinity of the active site are shown. The arrows depict the direction of movement upon TSA binding.

Figure 7. Stereoscopic representations showing phosphagen-specific residues in the vicinity of the TSA-bound active sites of (A) GK, (B) CK, and (C) AK. Phosphagen abbreviation used: Glyc, glycocyamine; Cr, creatine; Arg, arginine. The phosphagen-specific loop is colored brown and the nucleotide-binding loop is colored magenta.

Figure 8.

N-termini interactions in the dimer interfaces of phosphagen kinases. (A) ligand-free GK $\alpha\beta$ (B) GK $\beta\beta$ -TSA complex (most ordered dimer). Hydrophobic amino acid residues (Ile5, Phe9, and Phe26) mediate the interaction between neighboring α -helices in GK $\alpha\beta$ (colored blue and red). In GK $\beta\beta$, the two N-termini are incompatible and exhibit more conformational disorder compared with those of GK $\alpha\beta$. The GK $\beta\beta$ α -helix of the chain analogous to the β chain in GK $\alpha\beta$ (colored blue) is seen only in 2 out of the 18 dimers in the asymmetric unit of the GK $\beta\beta$ -TSA complex, and the β -chain analogous to the α chain of GK $\alpha\beta$ (colored red) is almost entirely disordered. The interface interaction involving Glu161 is asymmetric in both GK structures (one Glu161 interacts with Arg13 of the same subunit and the second interacts with Ser17 of the partner subunit). The segments colored in cyan and yellow correspond to N-terminal domain segments of partner subunits that adopt the same conformation, and which have been used as reference for structure superposition. (C) rabbit muscle CK-TSA complex ((17), PDB 1u6r) and (D) *T. californica* CK CK-TSAC ((16)PDB 1vrp). The N-termini of rabbit muscle CK are fully ordered, whereas those of the electric ray CK are partially disordered. Glu150 (a residue equivalent to Glu161 of GK)

interacts with the N-termini (of the partner subunit) with the exception of one subunit of the electric ray CK which exhibit more N-terminal disorder.

\$watermark-text

\$watermark-text

\$watermark-text

Table 1

X-ray data collection and structure refinement statistics.

	GKαβ	GKββ	GKββ-TSA
Data collection			
Space group	$P2_1$	$P2_1$	$P2_1$
Cell dimension (Å, °)	$a = 84.1, b = 98.2, c = 93.1, \beta = 91.3$	$a = 84.6, b = 99.7, c = 93.3, \beta = 92.4$	$a = 243.1, b = 114.3, c = 259.9, \beta = 90.3$
# dimers in the asymmetric unit	4	4	18
Solvent content (%)	44	44	46
Wavelength (Å)	1.00	1.54	1.00
Resolution (Å)	2.6	2.4	2.3
# observed reflections	116,527	228,348	1,872,589
Completeness (%) ^a	90.7 (84.7)	99.6 (99.4)	96.9 (81.2)
No. unique reflections	42,293	60,423	610,182
R_{merge}^b	0.085 (0.317)	0.084 (0.344)	0.094 (0.403)
$\langle I/\sigma(I) \rangle$	10.7 (3.4)	11.6 (3.5)	13.8 (3.1)
Redundancy	2.8 (2.7)	3.8 (3.7)	3.1 (2.3)
Refinement			
# reflections used	42,274	60,171	586,126
# protein atoms	12,034	11,833	107,131
# ligand atoms	–	–	1,440
# water molecules	268	623	6,774
R_{cryst}^c	0.192 (0.293)	0.197 (0.276)	0.197 (0.273)
R_{free}^d	0.266 (0.371)	0.267 (0.341)	0.263 (0.344)
RMSd from ideal geometry			
Bond length (Å)	0.015	0.014	0.009
Bond angle (°)	1.5	1.6	1.3
Average B factor (Å ²)			
Protein	52	43	28
TSA	–	–	22
Water	49	41	23

^aThe values in parentheses are for the highest resolution shell (2.72-2.60 Å for ligand-free GKαβ, 2.49-2.40 Å for ligand-free GKββ, and 2.38-2.30 Å for GKββ-TSA).

^b $R_{\text{merge}} = \sum_{hkl} [(\sum_j |f_j| - \langle I \rangle) / \sum_j |f_j|]$.

^c $R_{\text{cryst}} = \sum_{hkl} ||F_O| - |F_C|| / \sum_{hkl} |F_O|$, where F_O and F_C are the observed and calculated structure factors, respectively.

^d R_{free} is computed from 2,118 randomly selected reflections omitted from the refinement for ligand-free GKαβ, 3,038 for ligand-free GKββ, and 31,024 for GKββ-TSA