JBC Papers in Press. Published on February 24, 2011 as Manuscript M110.206813
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M110.206813

A SPRING LOADED RELEASE MECHANISM REGULATES DOMAIN MOVEMENT AND CATALYSIS IN PHOSPHOGLYCERATE KINASE

Louiza Zerrad[1], Angelo Merli[2], Gunnar F. Schröder[3], Andrea Varga[4], Éva Gráczer[4], Petra Pernot[1], Adam Round[5], Mária Vas[4] and Matthew W. Bowler[1]

From [1]Structural Biology Group, European Synchrotron Radiation Facility, 6 rue Jules Horowitz, F-38043 Grenoble, France; [2]Department of Biochemistry and Molecular Biology, University of Parma, Parco Area delle Scienze, 23/A 43100. Parma, Italy; [3]Institute of Structural Biology and Biophysics, ISB-3: Structural Biochemistry, Forschungszentrum Jülich, D-52425 Jülich, Germany; [4]Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, P.O. Box 7, H-1518 Budapest, Hungary and [5]European Molecular Biology Laboratory, Grenoble Outstation, 6 rue Jules Horowitz, 38042 Grenoble, France.

Running head: The complete reaction cycle of phosphoglycerate kinase

Address correspondence to: Dr. Matthew Bowler, Structural Biology Group, European Synchrotron Radiation Facility, 6 rue Jules Horowitz, F-38043 Grenoble, France, Tel: +33(0)476882928; Fax: +33(0)476882904; e-mail: bowler@esrf.fr

Phosphoglycerate kinase (PGK) is the enzyme responsible for the first ATP generating step of glycolysis and has been extensively implicated in oncogenesis and its development. Solution Small Angle X-ray Scattering (SAXS) data, in combination with crystal structures of the enzyme in complex with substrate and product analogues, reveal a new conformation for the resting state of the enzyme and demonstrate the role of substrate binding in the preparation of the enzyme for domain closure. Comparison of the X-ray scattering curves of the enzyme in different states with crystal structures has allowed the complete reaction cycle to be resolved both structurally and temporally. The enzyme appears to spend most of its time in a fully open conformation with short periods of closure and catalysis, thereby allowing the rapid diffusion of substrates and products in and out of the binding sites. Analysis of the open *apo* structure, defined through deformable elastic network (DEN) refinement against the SAXS data, suggests that interactions in a mostly buried hydrophobic region may favour the open conformation. This patch is exposed on domain closure making the open conformation thermodynamically more stable. Ionic interactions act to maintain the closed conformation to allow catalysis. The short time PGK spends in the closed conformation and its strong tendency to rest in an open conformation imply a spring-loaded release mechanism to regulate domain movement, catalysis and efficient product release.

Phosphoglycerate kinase (PGK) catalyses the transfer of phosphate from 1,3-bisphosphoglycerate (1,3BPG) to ADP in the first ATP-generating step of the glycolytic pathway. As a major controller of flux through the pathway, PGK is a viable target for drugs against anaerobic pathogens, such as *Trypanosoma* and *Plasmodium* species, that depend solely on glycolysis for energy metabolism (1). In addition to its metabolic role, the phosphoryl-transfer activity of PGK is important in the processing of anti-retroviral pro-drugs that take the form of L-nucleoside analogues (2). The rate-limiting step in the *in vivo* activation of such compounds has been demonstrated to be the addition of a third phosphate by PGK (3). A third activity of PGK is as a signalling molecule in chordates. It is integral in the response to hypoxia, when it is secreted from the cell and inhibits angiogenesis through a disulphide reductase activity that activates plasminogen autoproteolytic activity, producing angiostatin (4). This activity apparently uses the same mechanism as the normal metabolic reaction, and can be competitively inhibited by 3-phosphoglycerate (3PG) or ADP (5). Consequently, PGK has a crucial role in oncogenesis and its development.

PGK is composed of two similarly-sized domains, both with Rossmann fold topology, termed the N-domain, which binds the phosphoglycerate species 3PG and 1,3BPG, and the C-domain, which binds the nucleotides ADP and ATP. In early crystal structures of PGK (6-9), it was apparent that this state of the enzyme was incapable of catalysis. The relative orientation of the two domains meant that the two substrates would be too distant from each other to allow phosphoryl-transfer. Therefore, significant "hinge-bending" was suggested to be required for catalysis to occur (6). A major step forward in the understanding of catalysis by PGK was the determination of partially (10,11) and fully closed (12) crystal structures of the enzyme. The latter, in complex with transition state analogues (TSAs),

defined for the first time the state of PGK responsible for its catalytic activity, in particular demonstrating that a catalytic triad of positively charged residues surrounds the transferring phosphoryl-group. On adopting a fully closed conformation the separation of donor and acceptor atoms is between 3.9-4.3 Å, as predicted to be required for catalysis (13,14). In the closed conformation of PGK the two domains reorient by 33°, compared to the open form, about central hinge regions in order to bring the substrates together. Much effort has been put into elucidating the mechanism of this hinge-bending and the role that substrate binding plays in inducing the conformational changes needed for catalysis to occur (15-19).

Here, we present Small Angle X-ray Scattering (SAXS) data on complexes of human PGK (HsPGK) in solution. In combination with Deformable Elastic Network (DEN) refinement (20,21) a new "fully open" conformation of the *apo* enzyme is defined and details are revealed of the relative time spent by the enzyme in the open and closed conformations during catalytic turnover. Together with the crystal structures of the binary (3PG) and ternary (3PG-ADP and 3PG-AMP-PCP) complexes in a half-open conformation, it is demonstrated that, in addition to the role of substrate binding in domain closure, the enzyme strongly favours the open conformation. The exposure of a hydrophobic patch may be important in promoting domain opening. A complete reaction pathway for the enzyme can thus be presented.

Experimental Procedures

SAXS experiments – Expression, purification and activity measurements of recombinant wild type HsPGK were performed as previously described (18,22). Small Angle X-ray Scattering data were collected at the bioSAXS beamline ID14-3 (23) at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) with a Pilatus 1M detector (Dectris Ltd, Baden, Switzerland) at a wavelength of 0.931 Å and a camera length of 2.42 m. Scattering curves were measured from solutions of HsPGK without substrates; the binary complexes with either 50 mM 3PG or 10 mM ADP; a quaternary complex inhibited with the TSA 3PG-AlF_4^-ADP [50 mM 3PG, 10 mM ADP, 25 mM $MgCl_2$, 10 mM NH_4F and 2 mM $AlCl_3$ added 3 hours before measurements]; the 3PG-ADP ternary complex [50 mM 3PG, 10 mM ADP and 25 mM $MgCl_2$] and the 3PG-ATP ternary complex [50 mM 3PG, 10 mM ATP and 25 mM $MgCl_2$, nucleotide added just before measurements] in SAXS buffer (50 mM Tris pH7.5, 20 mM DTT). Measurements were performed at protein concentrations between 5 and 15 mg/ml to verify that any inter-particle effects that may have been present could be accounted for and rule out their influence on the analysis. In order to exclude the possibility of radiation damage, 10 frames, each of 10 seconds duration, were collected while continuously exposing fresh sample to the beam, the resulting frames were then compared to ensure no differences in the SAXS profiles were induced by exposure to X-rays. All data were processed using the ATSAS program package (24). Radii of gyration (R_g) were evaluated from Guinier plots using PRIMUS and pair distance distribution functions, *P(r)*, were computed with GNOM (25), see Figures 1 and S1. The solution shapes of hsPGK were reconstructed from the experimental data (GNOM functions) using the *ab initio* method. For each sample twelve independent DAMMIN reconstructions were aligned, averaged and filtered using the program package DAMAVER. A homology model of the *apo* form of HsPGK was created from the mouse crystal structure ((26) PDB accession number 2P9Q). The N- and C-terminal domains were then treated as rigid bodies and their positions refined against the scattering curve using SASREF (24). The model produced was then used to calculate an electron density map to 8 Å resolution as an alternative representation of the scattering data. The *apo* homology model was then fitted to this density map using the real-space refinement program DireX (21), which uses deformable elastic network (DEN) restraints (20) to balance the model deformations and the fit to the density. The fitted model was then energy minimized with CNS. The DireX refinement removed numerous atom clashes and improved local geometry in the rigid body fitted model with almost no change to the chi-square value of the fit to the scattering curve (0.97 to 1.08). As the structure was refined against the 8 Å density map (i.e. indirectly against the scattering curve) the limited amount of data did not allow for a cross-validation procedure, therefore no "free-chi-square" value could be computed. The final structure had 90% of backbone torsion angles in the allowed region (and 2% outliers) of the Ramachandran plot. The relative contributions of models to scattering curves was assessed with OLIGOMER (27) (Table 1 and Figure S2).

Crystallisation - Lyophilized HsPGK was resuspended in 50 mM Tris (pH 7.0), 20 mM DTT, 20 mM $MgCl_2$ and 20 mM 3PG and the protein concentration adjusted to 20 mg/ml.

Crystals with approximate dimensions 0.9 × 0.3 × 0.1 mm^3 were obtained by vapour diffusion in 4 µl hanging drops consisting of a 50:50 mix of protein and precipitant (26 – 31% PEG 4000 and 0.1 M Tris pH 7) solutions. Crystals of this complex diffracted poorly after cryoprotection (d_{min} = 3.0 Å) and were therefore conditioned by controlled dehydration using a HC1b dehydration device mounted on beamline ID14-2 at the ESRF (28). Crystals were harvested on micromeshes at a relative humidity (RH) of 98.5 % (the equilibrium RH with the mother liquor) and excess mother liquor removed. A reduction in the RH to 97.5% resulted in an increase in the diffraction limit of crystals and a small reduction in the *a* axis of the unit cell by *ca* 3 Å. After conditioning, crystals were cryo-cooled by plunging directly into liquid nitrogen, without the use of a cryo-protectant, and stored at 100K. To obtain the ADP and AMP-PCP open ternary complexes, crystals were transferred to a buffer containing 50 mM Tris pH 7.0, 20 mM DTT, 20 mM $MgCl_2$ and 20 mM 3PG, 31% PEG 4000 and 10mM ADP or AMP-PCP for 10 minutes prior to conditioning with the HC1b device as described above.

Data collection and structure solution - Diffraction data were collected from cryo-cooled crystals to resolutions of between d_{min} = 2.2 Å and 1.74 Å on an ADSC Q210 CCD detector at beamline ID14-2 (λ=0.933 Å) at the ESRF, Grenoble, France (see Table 2). The half open conformation crystallised in the orthorhombic space group $P2_12_12$ with unit cell dimensions of *a* = 61.6 Å, *b* = 73.0 Å and *c* = 93.6 Å with one molecule in the asymmetric unit (Table 2). Data were processed with MOSFLM (29) and programs from the Collaborative Computational Project Number 4 (CCP4) suite (30). The structures were solved by molecular replacement with MolRep (31) using the HsPGK structure (PDB accession code 2zgv (32)), stripped of ligands and water molecules, as a search model. In all subsequent refinement steps, 5% of the data, chosen at random, were excluded for calculating the free R-factor. Refinement was carried out alternately with REFMAC5 (31) and by manual rebuilding with the program COOT (33). Some solvent molecules were included using the ARP/waters function of ARP/wARP (34). The models contain the residues 4-416 for the HsPGK.3PG.open structure and 5-416 for the HsPGK.3PG.ADP.open and HsPGK.3PG.AMP-PCP.open structures. The final electron density maps for all the structures were of good quality, omit maps for all ligands are shown in Figures S3-S5. Stereochemistry was assessed with COOT (33) with all residues in preferred or allowed regions. Figures 2, 3 and 4 and S2 to S5 were produced with PyMol (35). Hinge bending analysis and interpolation of structures for movies was performed using the MORPH server (36).

Results

The solution structure of apo PGK and domain movements in solution - The X-ray crystal structures of PGK available define snap-shots in the catalytic cycle but lose detail on the extent and timing of the domain movements that occur in solution. In order to determine the events during the reaction cycle in solution, Small Angle X-ray Scattering (SAXS) was employed to determine the conformations of the protein at different stages of the reaction cycle. Scattering curves of the *apo* enzyme, the 3PG and ADP binary complexes, the ATP.3PG ternary (equivalent to the enzyme in catalytic turnover), the ADP.3PG ternary complex (a dead-end complex) and the HsPGK-3PG-AlF_4^--ADP TSA complex (the fully closed conformation) were measured (Figure 1). While similar experiments have been performed before (18,19) the experiments described here benefit from the ability to trap the fully closed conformation in solution by inhibition with aluminium fluoride (the ADP.AlF_4^- 3PG TSA complex) and compare the curves to the crystal structure of this complex and to the various open forms from the same species. In this respect, the study provides a full explanation of domain movement during catalysis. The radius of gyration (R_g) of the *apo* protein was found to be larger than that calculated from crystal structures of the open conformation (Table 1). *Ab initio* modelling for the *apo* protein produced a molecular envelope close to that obtained from crystal structures of the open conformation but with a greater angle between the two domains (Figure S1). In order to define the domain orientation in this new conformation a homology model of the HsPGK *apo* enzyme was constructed from the crystal structure of the mouse *apo* enzyme (26) and the N- and C- domains refined as rigid bodies against the scattering curves to define their relative orientation in solution. DEN refinement (20,21) was then used to further characterise the resting state of the enzyme (Figure 2). The model obtained gives an excellent fit to the data (Table 1, coordinates available in supplementary materials) and shows a conformation considerably more open than observed in crystal structures, with a 56° rotation required to form the catalytically active conformation (as opposed to a maximum of 33° from the most open crystal structure (37, Figure 2B). The validity of the model produced is

supported molecular dynamics studies (16) that detected a further opening of PGK, when compared to 'open' crystal structures. The domains also exhibit a 5° twist such that amino acids in the two domains that interact in the closed conformation are 'misaligned' relative to each other. Without the constraints of a crystal lattice, the protein thus adopts its fully open resting state. Scattering curves of the binary complexes, with either 3PG or ADP alone, show an R_g close to that of the *apo* protein implying a requirement for both substrates to bind before domain closure, in agreement with previous suggestions (11,19). Scattering curves obtained from the AlF_4^- TSA complex have a similar R_g to the crystal structure of the same complex and *ab initio* modelling confirms a more closed conformation of this complex (Figure S2). SAXS data therefore define both the fully open and fully closed states of PGK in solution. The curves obtained from of the ATP.3PG ternary complex do not match any of the crystal structures, even though the R_g is similar to the half-open crystal structure (Table 1). The *ab initio* model calculated from these data presents a shape reminiscent of both open and closed conformations (Figure S2 C) as SAXS scattering curves originate from an average of all the particles in solution and the enzyme was in full catalytic turnover throughout the measurements. It is possible, however, to calculate the contribution of conformational states from each stage of the catalytic cycle to the overall scattering curves observed. In this manner, the relative time the enzyme spends in each conformation during catalysis can be determined. The program OLIGOMER was used to determine the relative contributions of conformational states to the scattering curves of the protein. The degree of inhibition in the TSA complex was also determined (Table 1) by calculating the percentage of particles inhibited in the closed conformation. Both two component (fully open and fully closed) and three component (including the half-open crystal structure) models were applied to the data to see if there was significant time spent in the half-open state. The results of the fit suggest that when the enzyme is in full catalytic turnover it spends 92% of its time in the fully open conformation and only 8% of the time in the catalytically competent fully closed conformation. This figure is even lower for the binary and ADP/3PG dead end complexes with 97% of time spent in the open conformation, presumably because these complexes cannot form all of the interactions needed to stabilise the closed conformation. The TSA complex was found to be only 60% closed. This would account for the larger R_g compared to the crystal structure of the same complex and is probably due to incomplete inhibition by aluminium fluoride.

These results imply that while the enzyme moves between the open and closed state it has an energetic preference for the open conformation. A model of cycling and catalysis may be proposed where unliganded PGK remains open (ready to bind substrates, or release products) most of the time. This assertion is supported by isothermal calorimetric titration experiments, where a very small energy difference could be calculated between the open and the closed conformations of the enzyme (38). The binding of substrates induces structural changes that allow the stabilisation of the closed state for sufficient time to enable catalysis.

Preparation of the closed conformation by the binding of ligands - If PGK has a preference for an open conformation, and only the binding of both substrates affects this conformation, what is the role of substrate binding in the formation of the closed conformation? Previous studies have pointed to an induction of hinge bending upon ligand binding (17-19); however, this model does not adequately describe how products are subsequently released. In order to investigate the effects of ligand binding on the open conformation, crystals of the 3PG-HsPGK binary complex in an open conformation (HsPGK.3PG.open) were grown and ADP and the non-hydrolysable ATP analogue, AMP-PCP introduced, by soaking, in order to investigate the effect of their binding on the open conformation. The diffraction properties of the crystals were significantly improved by controlled dehydration using the HC1b dehydration device (28). Crystal structures were then solved of the 3PG binary complex and of the 3PG.ADP and 3PG.AMP-PCP ternary complexes. The complexes obtained are in a half open conformation with an opening angle between the two domains of 28° relative to the closed conformation observed in the crystal structure of the MgF_3^--TSA complex (Figure 2C). This conformation is very similar to previously observed crystal structures of open forms of PGK from various species (8,9,39). Comparison of the structures of these half-open binary complexes with the open structure of *apo* PGK shows how binding of 3PG prepares the enzyme for catalysis. The conclusions drawn are entirely consistent with previous findings derived from the crystal structure of the 3PG binary complex of pig muscle PGK (9). Upon binding of 3PG, residues R65 and R170 (located in the N-domain) are aligned in order to be able to form essential salt bridges (present only in the closed conformation) with a

reoriented D218 (located in the C-domain, see below). These salt bridges stabilise the catalytically active closed conformation (Figure 3A). In the half-open ternary complex with 3PG and ADP (HsPGK.3PG.ADP.open) binding of the nucleotide effects changes that provide information on the catalytic cycle. On binding ADP, the loop region at the N-terminus of helix-8 (C-domain) composed of residues 214 to 219 adopts two conformations, one moving *ca* 2 Å towards the nucleotide. This motion may assist in domain movement and leads to the positioning of the catalytic residue K215 to interact with the 3-phosphate of 3PG (or both the 1- and 3-phosphate groups of 1,3BPG), bound in the N-domain, upon closure and stabilise this conformation (an arrangement not previously observed in other open conformation structures). A reorientation of D218 to a position ready to form a salt bridge with R65, another interaction important in stabilising the closed conformation, also occurs (Figure 3B and C). This movement leads to an extension of helix-8. Therefore, the binding of ADP moves essential catalytic residues on a 'catalytic loop' into position by extending helix-8 and prepares the C-domain for stabilising the closed conformation.

In contrast, in the half-open ternary complex of PGK with AMP-PCP and 3PG (HsPGK.3PG.AMP-PCP.open) the 214-219 loop adopts the same conformation seen in the binary complex (Figure 3C). The adenosine moiety is bound in a similar manner to the ADP ternary complex but the β- and γ-phosphate groups protrude into the solvent and are poorly ordered. A conformation that is similar to the same complex of pig muscle PGK (40). As the γ-phosphate has been shown to possess high extent of mobility it could temporarily interact with Lys215, even in the open conformation (22). Thus, ATP may be able to stabilise the conformation seen in the HsPGK.3PG.ADP.open structure which is similar to that observed in the closed conformation (3PG-MgF_3^--ADP TSA complex) with K215 interacting with the phosphate groups of 3PG and the transition state analogue (Figure 3D). Before the enzyme closes, binding of ADP has positioned this residue in order to form the interactions needed to stabilise the closed conformation, in its absence, K215 is not positioned correctly and the closed conformation cannot be stabilised. Therefore, both 3PG and the nucleotide, bound in the open conformation, act to prepare the interactions that will exist only in the catalytically competent closed form.

Discussion

A hydrophobic spring may favour the open conformation - The data presented here point to a new model of catalysis in PGK. SAXS data demonstrate that the protein remains mostly in an open conformation, with (even in the *apo* form) short bursts of closure, while the X-ray structures reveal that ligand binding prepares the interactions needed to stabilise the closed conformation for the 8% of time spent on catalysis, rather than actually inducing closure. The need for both ligands to be bound in order for the closed conformation to be stabilised and for catalysis to occur is also shown. Analysis of the solvent accessible area and electrostatic potential of the different conformations (Figure 4A) shows that the fully open SAXS-observed conformation is needed for efficient substrate exchange. Only in this conformation are the substrate binding sites fully accessible to solvent with large positive and negative patches available for binding. The transition to the crystallographically-observed half-open conformation occludes the binding sites implying that this conformation would not be able to efficiently exchange ligands. No significant time is spent in this conformation in solution and it is therefore part of the transition between open and closed, and not functionally significant itself. Movement to the catalytically active conformation completely excludes the ligands from the solvent. The fully open conformation is clearly optimal for ligand exchange, but how is this conformation maintained/readopted in the presence of the strong electrostatic interactions present when ligands are bound? A detailed comparison of the fully open and fully closed structures points to a possible "spring" that drives the preference for the open conformation. The N- and C-domains are linked by a single helix (helix-7, residues 187 to 198) that is highly conserved across all kingdoms of life. Many of these residues are hydrophobic. Analysis of the fully open conformation shows extensive interaction between these residues and other conserved hydrophobic residues in the core of the protein (for a full list of residues see Table S1). When the enzyme moves to the half-open and fully closed conformations some of these interactions are disrupted and many residues form a solvent-exposed hydrophobic patch (see supplementary movie 1 and Figure 4B). This hydrophobic patch has a solvent exposed area of 502 $Å^2$ in the fully closed conformation which reduces to 150 $Å^2$ in the half open conformation and has a minimum of 92 $Å^2$ in the open state (masking 82% of the area from solvent, Table S1). These residues only become buried when the

opening angle between the domains is greater than 50°. This masking of the hydrophobic region from solvent could account for the preference of PGK to adopt an open conformation, as this will be thermodynamically more stable.

An analysis of 7 amino acid sequences of PGKs show that the hydrophobic nature of almost all of the residues in this region is conserved across all kingdoms, implying an essential role (Figure 5). The exposure to solvent of this hydrophobic patch in the catalytically competent closed conformation, stabilised when both substrates are bound, suggests that this conformation will only exist long enough for catalysis to occur and that the overall thermodynamic preference for the open conformation will eventually destabilise the closed conformation, causing it spring open, leading to the release of products and rebinding of substrates. The hydrophobic patch could act as a 'spring' by constantly applying pressure to the closed conformation, leading to the release of products after catalysis.

The "spring-loaded" release mechanism suggested here has parallels with ATP synthase. In this enzyme, a domain movement of 19° is effected mechanically, using the energy of an electrochemical gradient, in order to produce the catalytically competent conformation (41). The majority of the energy used is in the subsequent mechanical transformations to the open conformation that are required to release the newly synthesised ATP molecule. In PGK, the production of ATP is a less efficient process as the transformation between conformations is not controlled mechanically. Instead, a balance between thermodynamically stable states could be used to ensure that a tightly bound state is opened in order to release ATP and 3PG and allow binding of 1,3BPG and ADP for the next cycle. However, if the enzyme has a stable state that is non-catalytic and spends most of its time in this state, how can it act as an efficient catalyst? As the actual phosphoryl transfer reaction occurs extremely fast (13,14) the closed state does not need to be stabilised for too long. Therefore, in order for catalysis to be efficient, the open conformation needs to be dominant. The destabilising effect of the hydrophobic patch on the closed conformation may act like a powerful spring, allowing the protein to efficiently bind, and importantly, release substrates and products. As the catalytic trap will only be primed when all ligands are bound this prevents any time lost in futile stabilisation of the closed conformation.

The reaction pathway - Using the experimental evidence presented here a complete reaction cycle for PGK can be proposed. From the SAXS data it can be seen that the protein has a "normal state" with a large angle between the domains that maintains the substrates at a distance of 16 Å (nucleophile to phosphate atom) and that even during full catalytic turnover this conformation is preferred by the enzyme. Brownian motion will, presumably, provide energy to cycle the two domains between the open and closed conformations, but in the absence of substrates the interactions required to maintain the closed conformation away from the low energy open conformation are not present. However, upon substrate binding the charged residues that form the important salt bridges that stabilise the closed conformation are all positioned correctly. Therefore, the closed conformation is 'primed' on substrate binding. Substrate binding may also assists in the inter-domain communication and domain closure by the aid of well described interaction networks (17-19) and acts in concert with the internal dynamics of the protein. Therefore, as soon as the domains come together the salt bridges are formed and the closed conformation will be stabilised sufficiently for catalysis to occur. The tendency for the protein to return to the open conformation will avert any energy minima preventing substrate release or re-hydrolysis of the product. With all catalytic residues coordinating the transferring phosphate, the catalytic loop becomes an extension of helix-8 completing the continuing stabilisation of this region by the subsequent binding of substrates and its interaction with the N-domain. The substrates are now aligned for catalysis, and three positively charged residues (R38, K215 and K219) are positioned trigonally around the transferring phosphate, perfectly balancing the charge of the transition state. After phosphoryl transfer has occurred pressure to return to the open conformation from the exposed hydrophobic patch could destabilise them. As the domains begin the next cycle of opening, ATP bound to the C-domain and 3PG bound to the N-domain are exposed to the bulk solvent. The removal of the interaction with R38 and a destabilisation of helix-8, removing the interaction with K219, coupled with the lower affinity fully open conformation, leads to the release of ATP (movies defining all steps in the reaction cycle are available as supplementary materials). Competition between the higher affinities for ADP over ATP and 1,3BPG over 3PG may also play a part in the release of products (18).

The fast "spring-loaded" release mechanism with the majority of time spent in an open conformation would ensure the efficient

binding and release of substrates and products with short periods ensuring catalysis. Thus, the temperature-driven internal motion of protein structure seems to be essential for functioning of PGK, as also suggested for more complex allosteric proteins (42,43). While it has been assumed that the open, or unliganded, conformations of proteins that act through domain movements will be favoured (44), a mechanism for this equilibrium has not been proposed. Here we show that the equilibrium between open and closed is greatly biased towards the open conformation of PGK and a hydrophobic spring may be responsible. This mechanism may prove general for some of the large number of enzymes that act by domain movements.

REFERENCES

1. Opperdoes, F. R. (1987) *Annu Rev Microbiol* **41**, 127-151
2. Mathe, C., and Gosselin, G. (2006) *Antiviral Res* **71**, 276-281
3. Krishnan, P., Gullen, E. A., Lam, W., Dutschman, G. E., Grill, S. P., and Cheng, Y. C. (2003) *J Biol Chem* **278**, 36726-36732
4. Lay, A. J., Jiang, X. M., Kisker, O., Flynn, E., Underwood, A., Condron, R., and Hogg, P. J. (2000) *Nature* **408**, 869-873
5. Lay, A. J., Jiang, X. M., Daly, E., Sun, L., and Hogg, P. J. (2002) *J Biol Chem* **277**, 9062-9068
6. Banks, R. D., Blake, C. C., Evans, P. R., Haser, R., Rice, D. W., Hardy, G. W., Merrett, M., and Phillips, A. W. (1979) *Nature* **279**, 773-777
7. May, A., Vas, M., Harlos, K., and Blake, C. (1996) *Proteins* **24**, 292-303
8. Szilágyi, A. N., Ghosh, M., Garman, E., and Vas, M. (2001) *J Mol Biol* **306**, 499-511
9. Harlos, K., Vas, M., and Blake, C. F. (1992) *Proteins* **12**, 133-144
10. Auerbach, G., Huber, R., Grattinger, M., Zaiss, K., Schurig, H., Jaenicke, R., and Jacob, U. (1997) *Structure* **5**, 1475-1483
11. Bernstein, B. E., Michels, P. A., and Hol, W. G. (1997) *Nature* **385**, 275-278
12. Cliff, M. J., Bowler, M. W., Hounslow, A. M., Baxter, N. J., Blackburn, G. M., Vas, M., and Waltho, J. P. (2010) *J Am Chem Soc* **132**, 6507-6516
13. Mildvan, A. S. (1997) *Proteins* **29**, 401-416
14. Bowler, M. W., Cliff, M. J., Waltho, J. P., and Blackburn, G. M. (2010) *New J Chem* **34**, 784-794
15. Balog, E., Laberge, M., and Fidy, J. (2007) *Biophys J* **92**, 1709-1716
16. Palmai, Z., Chaloin, L., Lionne, C., Fidy, J., Perahia, D., and Balog, E. (2009) *Proteins* **77**, 319-329
17. Varga, A., Flachner, B., Graczer, E., Osvath, S., Szilagyi, A. N., and Vas, M. (2005) *FEBS J* **272**, 1867-1885
18. Szabó, J., Varga, A., Flachner, B., Konarev, P. V., Svergun, D. I., Závodszky, P., and Vas, M. (2008) *Biochemistry* **47**, 6735-6744
19. Varga, A., Flachner, B., Konarev, P., Graczer, E., Szabó, J., Svergun, D., Závodszky, P., and Vas, M. (2006) *FEBS Lett* **580**, 2698-2706
20. Schroder, G. F., Brunger, A. T., and Levitt, M. (2007) *Structure* **15**, 1630-1641
21. Schroder, G. F., Levitt, M., and Brunger, A. T. (2010) *Nature* **464**, 1218-1222
22. Flachner, B., Varga, A., Szabo, J., Barna, L., Hajdu, I., Gyimesi, G., Zavodszky, P., and Vas, M. (2005) *Biochemistry* **44**, 16853-16865
23. Pernot, P., Theveneau, P., Giraud, T., Nogueira Fernandes, R., Nurizzo, D., Spruce, D., Surr, J., McSweeney, S., Round, A., Felisaz, F., Foedinger, L., Gobbo, A., Huet, J., Villard, C., and Cipriani, F. (2010) *J Phys: Conf Ser* **247**, 012009
24. Petoukhov, M. V., Konarev, P. V., Kikhney, A. G., and Svergun, D. I. (2007) *J App Cryst* **40**, S223-S228
25. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) *J App Cryst* **28**, 768-773
26. Sawyer, G. M., Monzingo, A. F., Poteet, E. C., O'Brien, D. A., and Robertus, J. D. (2008)

Proteins **71**, 1134-1144
27. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) *J App Cryst* **36**, 1277-1282
28. Sanchez-Weatherby, J., Bowler, M. W., Huet, J., Gobbo, A., Felisaz, F., Lavault, B., Moya, R., Kadlec, J., Ravelli, R. B. G., and Cipriani, F. (2009) *Acta Cryst D* **65**, 1237-1246
29. Leslie, A. G. W. (1992) *Joint CCP4 and EACMB Newsletter on Protein Crystallography* **26**, Warrington: Daresbury Laboratory
30. Collaborative Computational Project Number 4. (1994) *Acta Cryst D* **50**, 760-763
31. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Cryst D* **53**, 240-255
32. Gondeau, C., Chaloin, L., Lallemand, P., Roy, B., Perigaud, C., Barman, T., Varga, A., Vas, M., Lionne, C., and Arold, S. T. (2008) *Nucleic Acids Res* **36**, 3620-3629
33. Emsley, P., and Cowtan, K. (2004) *Acta Cryst D* **60**, 2126-2132
34. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Cryst D* **49**, 129-147
35. DeLano, W. L. (2002)
36. Flores, S., Echols, N., Milburn, D., Hespenheide, B., Keating, K., Lu, J., Wells, S., Yu, E. Z., Thorpe, M., and Gerstein, M. (2006) *Nucleic Acids Res* **34**, D296-301
37. Flachner, B., Kovári, Z., Varga, A., Gugolya, Z., Vonderviszt, F., Náray-Szabó, G., and Vas, M. (2004) *Biochemistry* **43**, 3436-3449
38. Varga, A., Szabó, J., Flachner, B., Gugolya, Z., Vonderviszt, F., Závodszky, P., and Vas, M. (2009) *FEBS Lett* **583**, 3660-3664
39. Davies, G. J., Gamblin, S. J., Littlechild, J. A., Dauter, Z., Wilson, K. S., and Watson, H. C. (1994) *Acta Cryst D* **50**, 202-209
40. Kovári, Z., Flachner, B., Naray-Szabo, G., and Vas, M. (2002) *Biochemistry* **41**, 8796-8806
41. Bowler, M. W., Montgomery, M. G., Leslie, A. G., and Walker, J. E. (2007) *J Biol Chem* **282**, 14238-14242
42. Frederick, K. K., Marlow, M. S., Valentine, K. G., and Wand, A. J. (2007) *Nature* **448**, 325-329
43. Henzler-Wildman, K. A., Thai, V., Lei, M., Ott, M., Wolf-Watz, M., Fenn, T., Pozharski, E., Wilson, M. A., Petsko, G. A., Karplus, M., Hubner, C. G., and Kern, D. (2007) *Nature* **450**, 838-844
44. Gerstein, M., Lesk, A. M., and Chothia, C. (1994) *Biochemistry* **33**, 6739-6749

FOOTNOTES

The authors thank the Partnership for Structural Biology (PSB), Grenoble, for an integrated structural biology environment; Gordon Leonard and Sean McSweeney (ESRF) for helpful discussions and comments on the manuscript as well as Péter Závodszky (Institute of Enzymology, Hung. Acad. Sci.) for his continuous interest in the work. Assistance in expression and purification of the protein by Judit Szabó (Institute of Enzymology, Hung. Acad. Sci.) is also greatly appreciated. The financial support by the grant OTKA (NK 77978) of the Hungarian National Research Grant is gratefully acknowledged.

Data deposition: The atomic coordinates and structure factors for the half-open conformation PGK-3PG; PGK-3PG-ADP and PGK-3PG-AMP-PCP complexes have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 2xe6, 2xe7, and 2xe8 respectively). The coordinates for the fully open conformation defined by DEN refinement against the SAXS envelope has been deposited in the Protein Data Bank (PDB ID code 2yag) with the experimental data (scattering curve and pair distribution function). As the PDB SAXS task force is currently deliberating the requirements for depositions of this kind, the coordinates will not be released until their recommendations have been made. Therefore, the coordinates have been made available in supplementary materials.

The abbreviations used are: deformable elastic network, **DEN**; transition state analogue, **TSA**; human phosphoglycerate kinase, **HsPGK**; 3-phosphoglycerate, **3PG**; 1,3-bisphosphoglycerate, **1,3BPG**; β,γ-methyleneadenosine 5'-triphosphahte, **AMP-PCP**

FIGURE LEGENDS

Figure 1. Scattering curves and pair distance distribution functions. A. Comparison of SAXS Curves and the 2 component fits from OLIGOMER. The experimental curves for the samples are shown with the two component fits from OLIGOMER (blue lines). For the *apo* data the fit of the DEN refined structure as calculated using CRYSOL is also shown (green line). The curves are displaced for visualization and therefore log (I) values are relevant for the *apo* sample only. **B.** Pair distance distribution function, *P(r)*, plots calculated by an indirect Fourier transformation of the SAXS data using GNOM.

Figure 2. A The structure of the fully open *apo* HsPGK (ribbon, the N-domain is coloured green, the C-domain red and the inter-domain region cyan) determined by rigid body and DEN refinement against SAXS data. **B** The structure of the SAXS-observed fully open *apo* HsPGK aligned with the structure of the fully closed HsPGK-MgF_3^--ADP complex (2wzb). The N-domain of the closed conformation is coloured orange for clarity. Structures were aligned on residues 200-400 of the C-domain and are related by a 56° rotation of the N-domain. The substrates 3PG, MgF_3^- and ADP are shown as spheres (carbon atoms green). **C** The crystal structure of the half-open conformation of the HsPGK-3PG binary complex (ribbon, coloured as above) aligned as before with that of the fully closed HsPGK-MgF_3^--ADP complex. The half-open and fully closed structures are related by a 28° rotation of the N-domain. Substrates are shown as spheres (carbon atoms green).

Figure 3. Conformational changes induced by substrate binding in PGK. A: The binding of 3PG. An alignment of residues 10-150 of the half-open conformation binary complex (yellow) and the mouse *apo* protein. 3PG is bound by three arginines and a histidine. Residues R65 and R170 are crucial in stabilizing the closed conformation and are only positioned correctly on ligand binding. **B:** The binding of ADP to the binary complex. An alignment of residues 200-400 of the half-open conformation binary (grey) and 3PG ADP ternary (yellow) complexes. On binding of ADP the catalytic loop (Residues 214-219) adopts 2 conformations, the end of helix 8 moves 2 Å towards the nucleotide leading to the preparation of K215 for catalysis (arrow). The closed conformation is also prepared for by the positioning of D218 to form an essential salt bridge with R65 from the N-domain. **C**: An alignment of residues 200-400 of the ADP (grey) and AMP-PCP (yellow) half-open conformation ternary complexes. The catalytic loop resumes the conformation observed in the binary complex and the ATP analogue is ready for release. **D**. Comparison of the half-open 3PG ADP ternary complex (grey) and the fully closed 3PG-MgF_3^--ADP TSA complex (yellow). On domain closure a salt bridge is formed between D218 and R65, K215 is positioned for catalysis and helix 8 is extended. The alternative position of K215 is very similar to that observed in the closed conformation and prepares the enzyme to form an interaction with the 3-phosphate of 3PG to stabilise the closed conformation. In the half-open conformation the catalytically essential residue K219 remains coordinated to the α-phosphate of ADP as there is no γ-phosphate or 3-phosphate present and is not involved in stabilizing the closed conformation.

Figure 4. Changes in the solvent exposed surface upon domain closure. **A**. Electrostatic potential representation (blue, positive; red, negative; white, neutral, values from +2 kcal/mol to -2 kcal/mol) of the solvent accessible surface (probe radius 1.4 Å) of the fully open *apo* conformation, the half-open binary complex and the fully closed conformation. In the fully open conformation the binding sites for ligands are well separated and exposed to the bulk solvent. The complementary charged areas that stabilize the closed conformation can easily be seen. Ligands are shown as spheres (carbon atoms green) **B.** Surface representation of the transition from fully open to fully closed. Residues involved in hydrophobic interactions are coloured yellow. As the enzyme moves to the fully closed a hydrophobic patch is exposed.

Figure 5. Multiple sequence alignment of PGK from all kingdoms of life. Residues coloured yellow are the same as those shown in figure 4B. Residues that form the hydrophobic patch are almost completely conserved from archaea to man. Residues involved in catalysis are shown in red. Species are as follows: Sus – *Sus scrofa*; Homo – *Homo sapiens*; Saccharomy – *Saccharomyces cerevisiae*; Arabidopsi – *Arabidopsis thaliana*; Trypanosom – *Trypanosoma brucei*; Escherichi – *Escherichia coli*; Pyrococcus – *Pyrococcus furiosus*. The trypanosome specific insertion (residues 69-84) was removed for clarity.

Table 1. SAXS experimental data, oligomer fits and comparison with crystal structures

Protein complex	**Rg (nm)**	**Percentage open**	**Percentage closed**	**Chi square of fit**
Apo	2.42	-	-	1.08 (DEN model)
3PG-AlF_4^--ADP TSA complex	2.24	39.3	60.6	1.36
3PG.ATP	2.34	92.5	7.5	7.55
3PG.ADP	2.37	96.5	3.5	0.9
3PG	2.40	96.6	3.4	2.75
ADP	2.42	96.7	3.3	5.15
3PG-AlF_4^--ADP TSA crystal structure	2.17	-	-	-
Binary complex 'open conformation' crystal structure	2.33	-	-	-

Table 2 Data processing and refinement statistics

Structure	hsPGK 3PG	hsPGK 3PG-ADP	hsPGK 3PG-AMP-PNP
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$
Wavelength	0.933 Å	0.933 Å	0.933 Å
Unit cell dimensions (Å) a, b, c	61.6, 73.0, 93.6	58.6, 73.4, 93.4	59.3, 71.2, 91.5
Resolution range (Å)	20-1.74	20-2.2	20-1.79
Number of unique reflections	43561	20841	36815
Multiplicity[1]	4.1 (3.7)	3.9 (4.0)	2.9 (2.7)
Completeness[1] (%)	98.9 (100.0)	99.1 (99.8)	98.6 (97.1)
Rmerge[1,2]	0.10 (0.29)	0.06 (0.29)	0.07 (0.29)
$\langle I/\sigma(I)\rangle$[1]	8.1 (3.2)	12.8 (4.0)	10.1 (3.6)
Wilson B factor	22.6 Å^2	32.8 Å^2	19.4 Å^2
Water molecules	363	65	140
R-factor[3] (%)	20.6	22.4	19.8
Free R-factor[4] (%)	25.3	28.8	24.3
RMS deviations:			
Bonds (Å) Angles (°)	0.023 1.856	0.017 1.61	0.023 1.96

[1] Statistics for the highest resolution bin (1.83 – 1.74 - 3PG; 2.32 – 2.20 – 3PG-ADP; and 1.88 – 1.79 3PG-AMP-PNP) are shown in parenthesis

[2] $R_{merge} = \sum_h \sum_i |I(h) - I(h)_i| \Big/ \sum_h \sum_i I(h)_i$, where I(h) is the mean weighted intensity after rejection of outliers.

[3] $R = \sum_{hkl} \left| |F_{obs}| - k|F_{calc}| \right| \Big/ \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes.

[4] $R_{free} = \sum_{hkl \subset T} \left| |F_{obs}| - k|F_{calc}| \right| \Big/ \sum_{hkl \subset T} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes and T is the test set of data omitted from refinement (5% in this case).

Figure 1

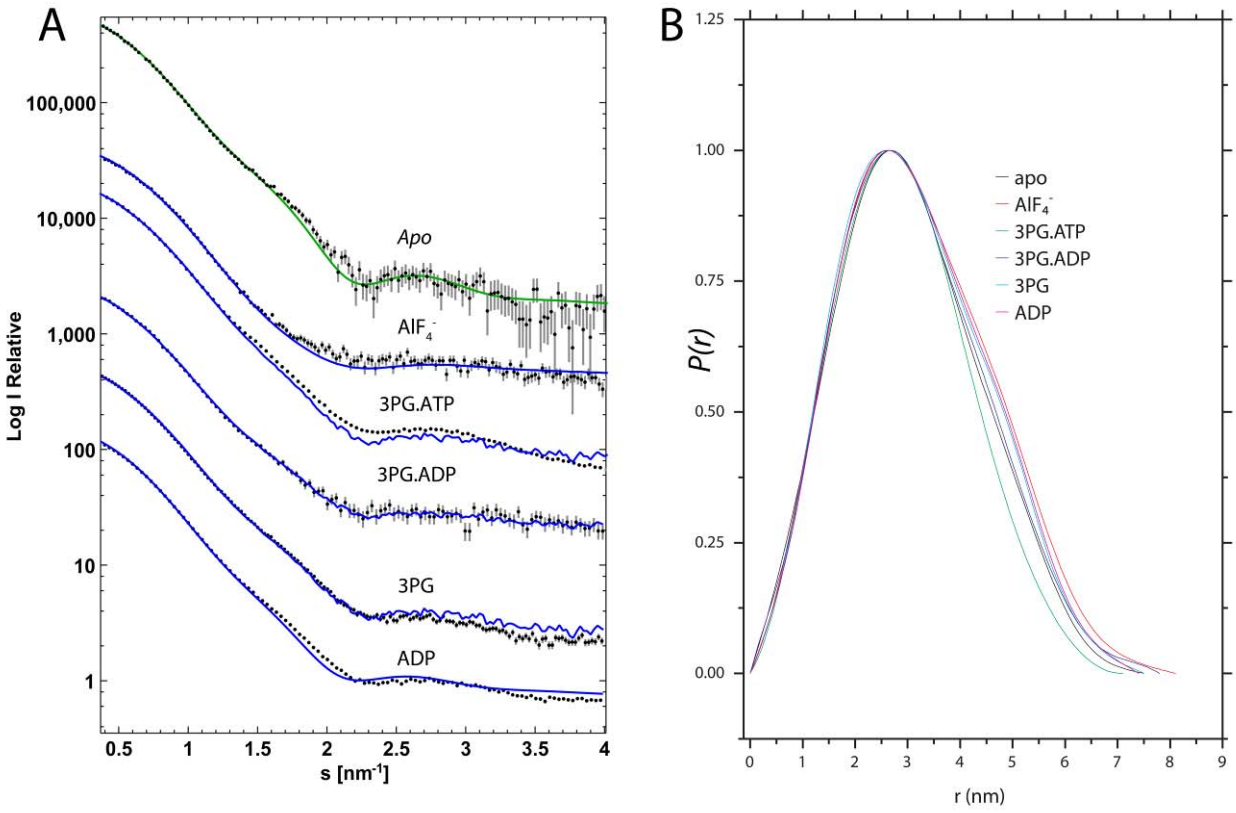

Figure 2

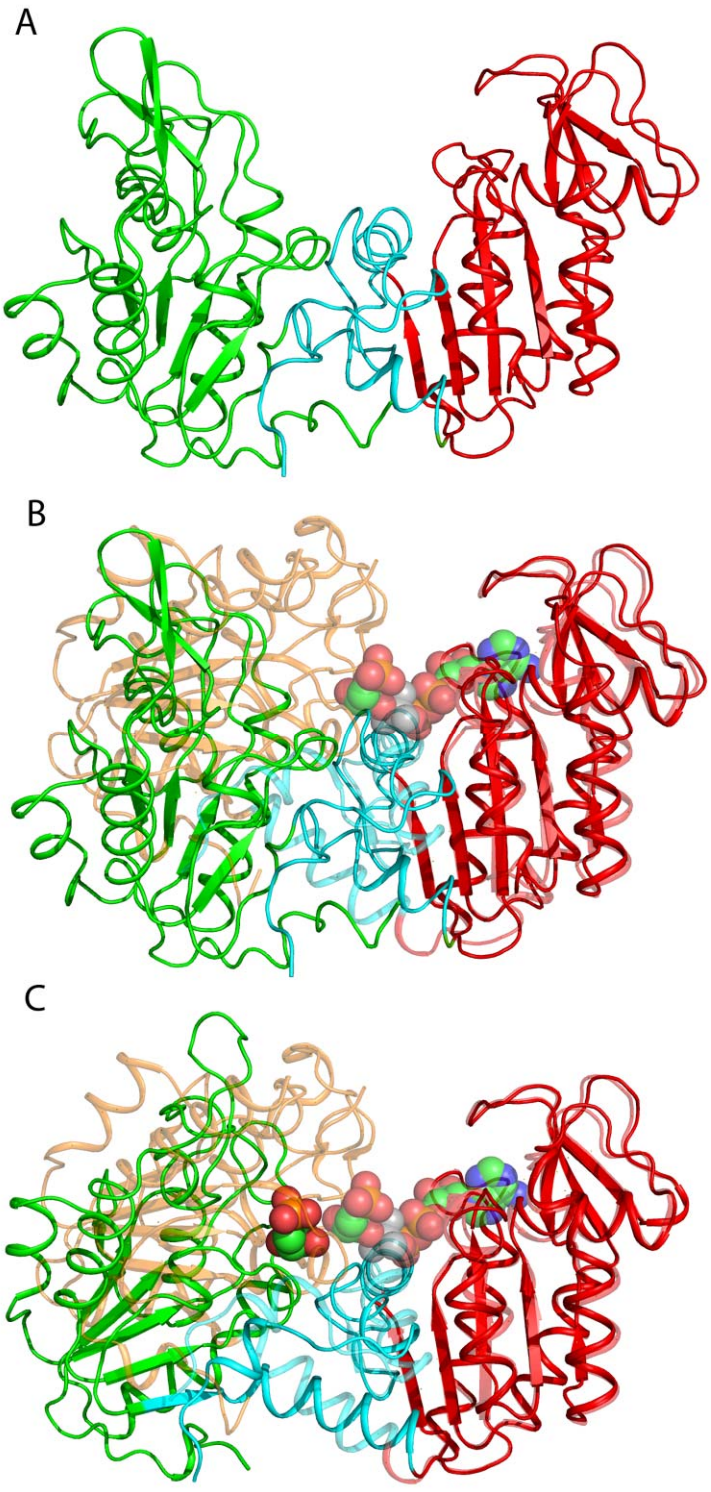

Figure 3

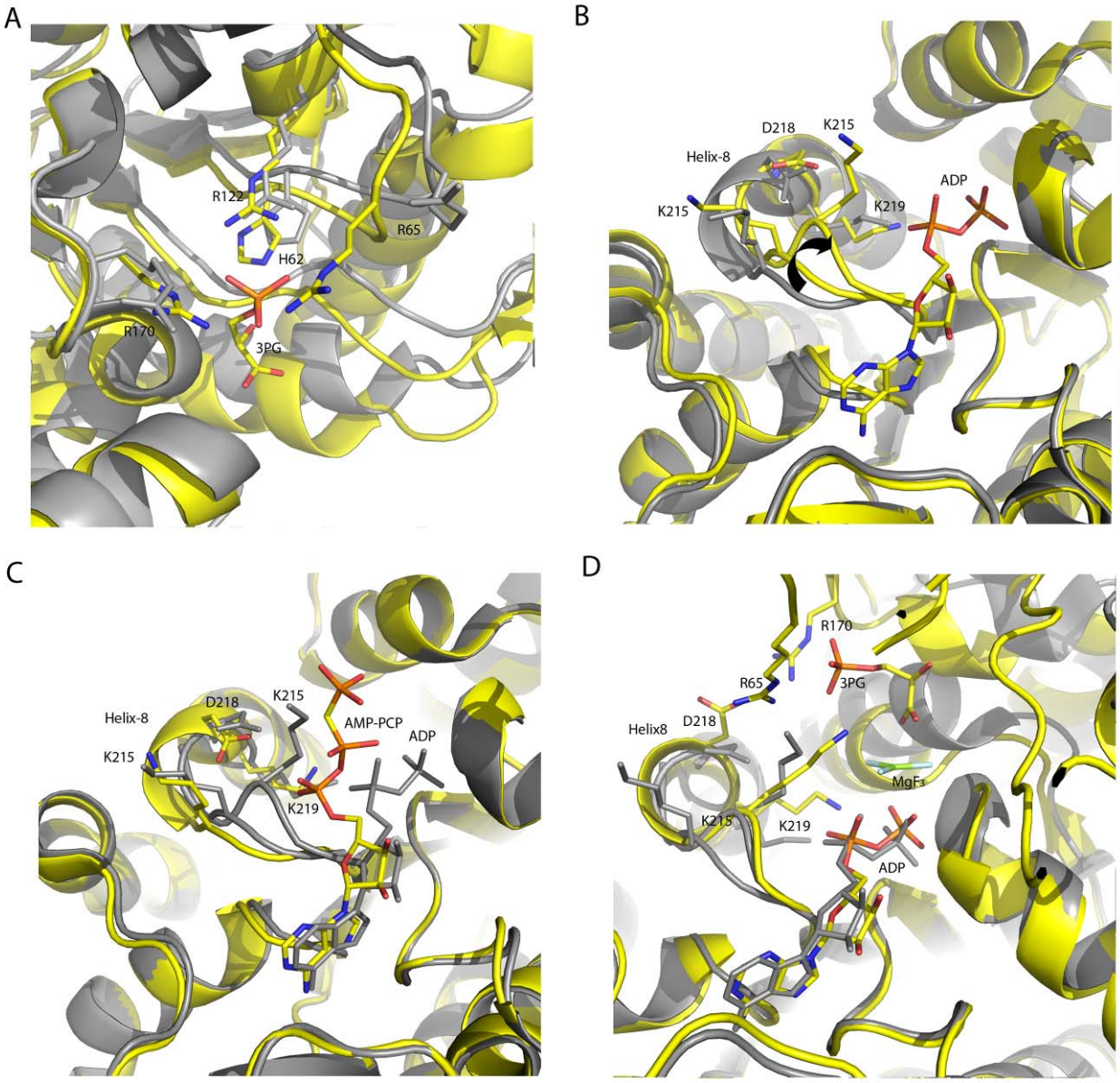

Figure 4

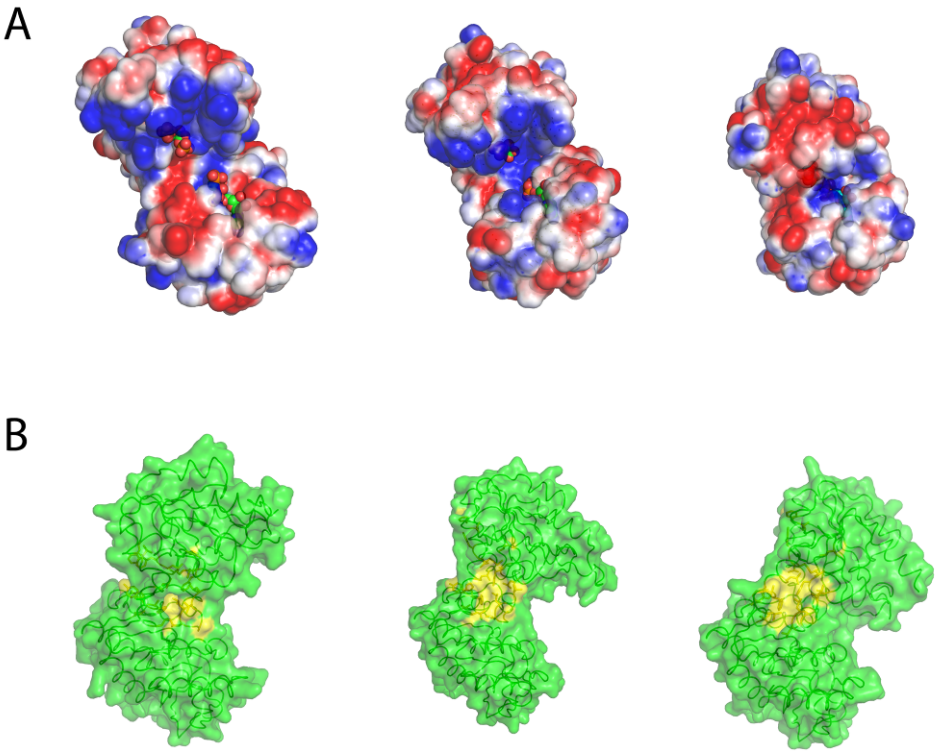

Figure 5