THE ROLE OF SUBSTRATE BINDING AND CHIRALITY IN THE DYNAMICS OF HUMAN 3-PHOSPHOGLYCERATE KINASE

PhD thesis

Zoltán Pálmai

Pharmaceutical Sciences Doctoral School

Semmelweis University

Supervisor: Dr. Erika Balog, Ph.D.

Official reviewers:
Dr. Zsuzsanna Dosztányi, Ph.D.
Dr. Szabolcs Béni, Ph.D.

Head of the Final Examination Committee:
Dr. Péter Mátyus, Ph.D., D.Sc.

Members of the Final Examination Committee:
Dr. András Czirók, Ph.D.
Dr. Attila Ambrus, Ph.D.

Budapest, 2012

1. INTRODUCTION

The biochemical function of the 3-phosphoglycerate kinase (PGK) enzyme is to catalyze the reversible transfer of phosphate between 1,3-bisphosphoglycerate (1,3-BPG) and adenosine diphosphate (ADP). It is ubiquitously expressed in all living organisms and plays a key role in glycolysis (for the production of adenosine triphosphate (ATP)) and carbon fixation in plants.

Human PGK (hPGK) was found to be involved in the activation of a broad spectrum of D- and L-nucleotide analogues, a new class of antiviral and anticancer agents. L-nucleotide analogues are the mirror images of the natural D-nucleotides analogues. Because nucleotides do not pass the cytoplasmic membrane, they cannot be used directly as drugs; instead they are given as the corresponding nucleosides because these "prodrugs" do pass. To be pharmacologically active nucleoside analogues need to be phosphorylated to their respective triphosphate metabolites. This stepwise phosphorylation is carried out by kinase enzymes. The last phosphorylation step is catalyzed by hPGK. Because the accumulation of diphosphate analogues in cells demonstrates that the third phosphorylation step is often rate limiting. Many times large amount (several hundreds of mg) is needed from a certain type of nucleoside analogue (e.g. Cidofovir, Valtorcitabine, Pentacept etc.) to achieve the effective dose of the pharmacologically active nucleotide form in the organism. Particular nucleoside analogues are phosphorylated by hPGK less efficiently. The knowledge of the atomic details of the dynamic properties of hPGK would significantly alleviate *in silico* tailoring of prodrugs to assure their efficient phosphorylation.

PGK is a monomer protein that contains two globular domains. The N-terminal domain binds 1,3-BPG or 3-PG, and the nucleotide binding site is located in the C-domain. In the early crystallographic studies the crystal structures were reported only in the open form for PGK for several species. In the open conformation of the enzyme a broad cleft separates the two domains and the binding sites are too far (12-15 Å) from each other for chemical interaction. Thus it was generally supposed that the reaction requires domain closure, a domain movement called hinge bending motion in the literature. The closed form was determined only in 1997. In the beginning of my doctoral studies the structure of both end states of the same species had not yet been reported, i.e. there was not a direct experimental evidence for the hinge bending functional domain motion.

In 2011 by the determination of the open and closed state of the human enzyme, the hinge bending hypothesis was confirmed. However the knowledge of both end states is not sufficient for the characterization of functional motions at the atomic scale. The molecular understanding of the mechanism of action of hPGK and the rational development of efficiently phosphorylated prodrugs necessitate further experimental and theoretical investigations.

During my doctoral studies I characterized the local and collective motions of hPGK and the structural-dynamic conditions of efficient enzyme activity using a variety of computational simulation methods.

2. OBJECTIVES

1. **The influence of substrate binding on protein dynamics:** What is the effect of binding of natural substrates on the dynamics of hPGK? Does the complex enzyme exhibit a significant difference on the ns time scale in its dynamic characteristics compared to the apo form?

2. **The influence of substrate chirality on protein dynamics:** What is the underlying atomistic phenomenon for the broad nucleotide specificity and the lack of enantioselectivity of hPGK? What is the structural-dynamical reason for the observed differences in the binding affinity of nucleotides differing in chirality and nucleobase?

3. METHODS

Molecular dynamics simulations were carried out on apo hPGK and on its four ternary complexes: D-/L-ADP*1,3-BPG*Mg*hPGK (D-/L-ADP complex), D-/L-CDP*1,3-BPG*Mg*hPGK (D-/L-CDP complex). The structure of hPGK was constructed from the open conformation of a pig muscle PGK crystal structure (PDB code: 1VJC) by homology modeling using the Modeller program, release 8v2. For determining the ternary complexes we docked the 1,3-BPG and the various nucleotides to the already constructed human structure. The docking procedure was carried out with the GOLD program version 3.2.

MD runs were carried out with NAMD (CHARMM all-atom parameter set 22) on an SGI Altix 350 machine. The preparation and the analysis of the simulations were performed with the CHARMM program. The energy of the model structures was minimized by 1500 steps of steepest descent, 200 steps of conjugate gradient and 1000 steps with adopted basis Newton-Raphson method. Each system was immersed in a rectangular water box of thickness of 12 Å measured from the protein surface using periodic boundary conditions. Na- and Cl-ions were added to random positions and the solvated systems were energy minimized using the same protocol as described earlier. For energy calculations a dielectric constant of 1 was taken. The PME method was used to calculate the electrostatic interactions with a grid spacing of 1 Å and order of 6, a real space summation cut-off of 12.0 Å and a width of Gaussian of 0.34 $Å^{-1}$. Van der Waals interactions were reduced to zero by ‘‘switch’’ truncation operating between 10.0 and 12.0 Å. The MD simulation was carried out with an integration time step of 1 fs. The energy minimized structures were each first heated to 300 K with a temperature increment of 10 K after each 1000 steps. In each case a 50 ps canonical equilibration was followed by a 500 ps *NpT* (isobar-isotherm) equilibration. After this a 20 ns *NpT* (p=1 atm, T=300 K) production run was carried out in each case and the coordinates of the trajectories were saved every 1 ps.

The local motions were analyzed by calculating the RMSF and the cross-correlation values of the atomic displacements.

The collective motions were analyzed by PCA method and calculating the cross-correlation values of the atomic displacements.

For hinge point identification time evolution of dihedral angles formed by consecutive C_α atoms in the protein chain were calculated.

Frequency distribution of protein conformations. Frequency distribution of projections of the vector linking the centers of mass of domains onto the three principal axes of inertia were computed. The corresponding histograms of these projections were drawn to characterize the differently populated conformational states.

4. RESULTS

4.1. Characterization of the dynamics of apo hPGK and D-ADP complex

The time evolution of the distance between the centers of mass of the two domains during the 20 ns production run revealed that both in the absence and in the presence of the natural substrates the enzyme exhibits a hinge bending type motion but the characteristics of this motion vary considerably. In the apo form, the enzyme exhibits a hinge bending motion of a relatively small amplitude with a time period around 20 ns, while the time period of the complexed form is much longer than our simulation time. The apo enzyme explores larger number of conformational states than the D-ADP complex.

Fluctuation analysis of local motions. In general, the highly flexible regions are the loop regions in both forms of the enzyme. In addition the binding of 1,3-BPG increases the flexibility of the binding range in the N-domain. The effect of D-ADP binding is more complex. D-ADP binding rigidifies the phosphate and ribose binding site, while increases the flexibility of some residues in the base binding site. The nucleobase is inserted into a flexible hydrophobic binding pocket. In contrast the negatively charged phosphate chain is kept rigidly by the electrostatic interactions of the surrounding residues, which explains the reduced flexibility of the phosphate and ribose binding residues. The fluctuations of the β sheets constituting the ‘‘cores’’ of both of the domains are very low, acting as a ‘‘rigid core’’ of the enzyme.

Correlation analysis of intra- and interdomain motions. The interdomain motions exhibit collective motion characteristics in both systems. Our results demonstrate that one domain moves with respect to the other as a whole unit. The C-terminal part of the protein, which bends back to the N-terminal and is taken as part of the N-domain moves together with the N-domain.

In general it can be noted that the correlation of intradomain atomic movements became more pronounced upon binding the natural substrates. The N-domain displays very low correlations in each system. In contrast high correlations appear between the nucleotide binding residues of the C-domain in both forms of the enzyme. The nucleotide binding residues move together in a concerted fashion. Upon D-ADP binding a further increase in the correlation of particular base and ribose binding residues can be seen. Furthermore, anticorrelated motions on the opposite sides of the C-domain are present, which likely indicates a twisting domain motion.

PCA analysis of collective motions. Three modes are sufficient to describe the collective motions of both systems. In case of the apo enzyme the first three modes characterize shear, hinge and twisting domain motions respectively. It indicates that the shear domain motion (40.1 % of total protein fluctuations) is more dominant than the hinge motion (30.3 % of total protein fluctuations), nevertheless latter is also important component of the protein dynamics. In case of D-ADP complex the above mentioned three types of motions determine the collective dynamics also, however the order of motions according to their contributions to the total fluctuations is: twisting, hinge and shear domain motions. Thus, the twisting motion (40.3 % of total protein fluctuations) is the dominant component of the dynamics, nevertheless the hinge type of motion (24.1 % of total protein fluctuations) is also important.

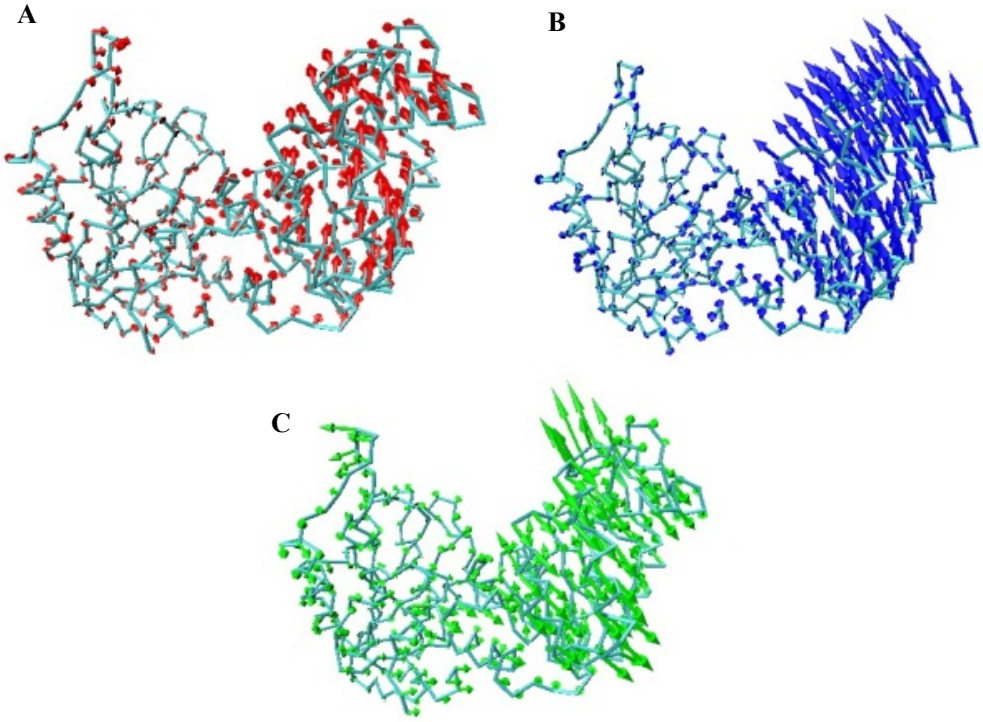

Figure 1 The C_α PCA component vectors of (**A**) shear, (**B**) hinge and (**C**) twisting modes shown on the 3-dimensional trace representation of hPGK.

Hinge point identification. In the case of apo enzyme there are two hinge points at the C-terminal side of α7 helix (Val199, Ser202), while in the region before the C-terminal – near to the binding site – three hinge points can be found (Gly372, Asn383, Thr393). Upon substrate binding the number of hinges decrease, one hinge remains at the C-terminal side of α7 helix (Pro203) and at the proximity of the substrate binding site (Gly372), which becomes dominant.

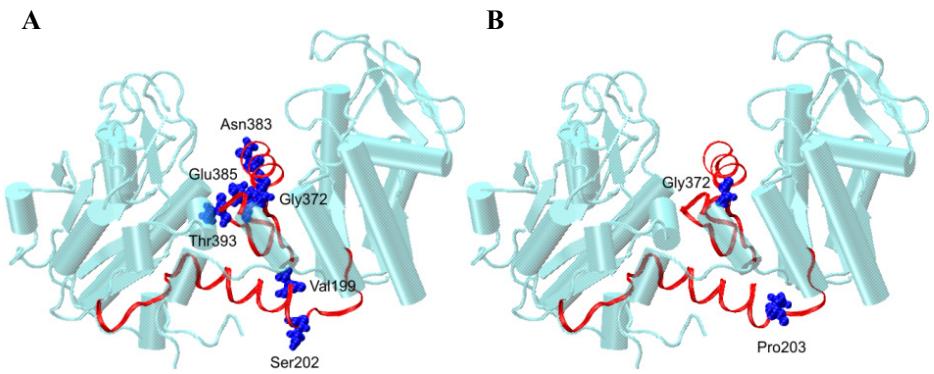

Figure 2 The location of hinge points in the structure of the (**A**) apo enzyme and (**B**) the D-ADP complex. Hinge points are indicated by blue van der Waals spheres. The interdomain regions are denoted by red ribbon representation. In the case of the complex substrates are not shown for clarity.

4.2. Characterization of the dynamics of D-/L-ADP - and D-/L-CDP complexes

Characterization of nucleotide binding. Based on the simulation results, relative binding affinities of the observed nucleotides differing in nucleobase (ADP, CDP) and chirality (D, L enantiomers) were predicted. D-/L-ADP exhibited similar binding affinities towards the protein (slightly lower for L-ADP). In contrast D-/L-CDP showed much lower affinity than D-/L-ADP. D-/L-CDP substrates form weak interactions with hPGK residues making the binding of the nucleotide quite unstable as observed during the dynamics. According to the analysis of the non-bonded energy, this fragile binding is explained by a reduction of the van der Waals contribution in the interaction energy between hPGK residues and CDP. Binding of the phosphate chains was stable during the simulation for both CDPs, while the bases escaped from the binding pocket in a short time. CDP nucleotides were not further analyzed in details since they were not fully bound to the protein during the entire simulation.

Characterization of 1,3-BPG binding in the presence of D- or L-ADP. According to our data the binding affinity of 1,3-BPG towards the enzyme is lower in the presence of D-ADP than in the case of L-ADP. The altering binding modes of the phosphate chains of D- and L-ADP are likely responsible for the differences in the binding affinities of 1,3-BPG. Due to the slightly alternative orientation of the phosphate chain of L-ADP, the influence of its electrostatic repulsion is less pronounced at the 1,3-BPG binding site. To prove this hypothesis further investigations are required.

Collective motions. According to the frequency distributions of the principal axes of inertia projections the D-ADP complex exhibits high rigidity in the directions of both shear and

hinge motions. Thus it explores low number of conformational states. Based on the PCA results, in the case of the D-ADP complex the dominant collective motion is the twisting domain motion, which explains the restricted motions in the shear and hinge directions. During twisting the distance between the centers of mass of domains changes to a very less extent in the shear and hinge directions.

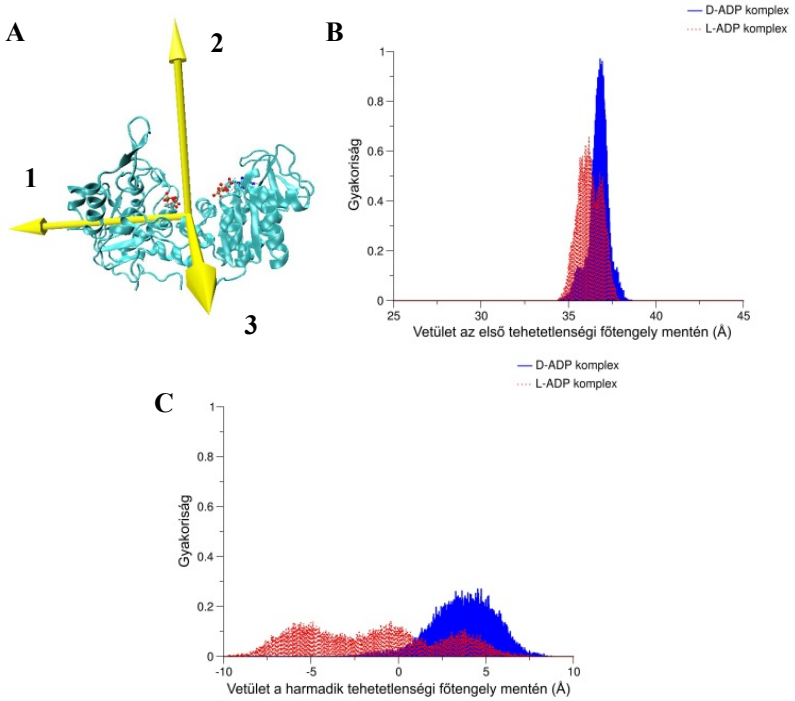

Figure 3 (**A**) The principal axes of inertia of hPGK. (**B, C**) Frequency distribution of the projection of the vector linking the centers of mass of domains onto the (**B**) first and (**C**) third principal axes of inertia. The conformational distribution broadens to a great extent for the L-ADP complex.

In the case of the L-ADP complex the conformational distribution broadens to a great extent in both directions, mainly in the shear direction. It indicates the intensive flexibility of the L-ADP complex in this direction. The system explores a wide range of conformational states during the simulation. The PCA results are in agreement with these findings. Two modes are responsible for the collective motions of the L-ADP complex: the first describes the shear domain motions, while the second mode characterizes the hinge domain motions. As the first mode comprises the 80.7 % of the total protein fluctuations, it dominates the protein dynamics. Surprisingly the hinge domain motions compose only 7 % of the total fluctuations, i.e. the hinge motions act a very minor role in the dynamics of the L-ADP complex.

Local motions. The binding site of L-ADP is more rigid compared to that of D-ADP. E.g. the hydrophobic binding pocket which becomes flexible upon D-ADP binding, remains rigid in the L-ADP complex. Regarding the phosphate chain binding residues, the influences of D- and L-ADP on the protein dynamics are similar: both enantiomers rigidify the phosphate binding site. It indicates the similar tight binding mode of both phosphate chains. (However the phosphate binding of L-ADP is slightly weaker.)

Correlations of the inter- and intradomain motions. Very similar patterns can be observed on the correlation maps of D- and L-ADP complexes. Regarding the intradomain motions, high correlations can be found between the ADP binding residues. In the D-ADP complex these correlations are even more pronounced. Furthermore, the anticorrelated motions on the opposite sides of the D-ADP complex C-domain, do not appear in the case of the L-ADP complex. It proves that the domain motions of L-ADP complex do not contain twisting components in agreement with the PCA results.

The correlations of interdomain motions also exhibit great similarities for the two complexes. The correlation maps confirm that the domains move as a whole entity, a rigid body relative to each other in both complexes.

Hinge point identification. Both complexes possess a hinge point at the C-terminal side of α7 helix: Pro203 (D-ADP) and Glu201 (L-ADP). The similar location of these hinges and corresponding dihedral angle change indicate that the influence of the nucleotide chirality is not significant on these hinge points. In contrast the altering chirality exerts a more important effect on the hinge points at the C-terminal, which is in close proximity of the binding site. While one dominant hinge point (Gly372) is identified in this interdomain region for the D-ADP complex, two less important hinges (Ile370 and Trp382) can be found for the L-ADP complex. The latter two hinges coordinate the domain motions with similar contributions.

5. CONCLUSIONS

5.1. Effects of substrate binding

- in the apo form, the enzyme exhibits a hinge bending motion of a relatively small amplitude with a time period around 20 ns, while the time period of the D-ADP complex is much longer than our simulation time
- in the apo enzyme the collective interdomain motion is preponderantly the superposition of shear and hinge motions, while in the D-ADP complex the collective interdomain motion is composed of mainly twisting and hinge motions
- the apo form is more flexible, there are more hinge points that contribute with similar significance to the collective interdomain motions, while in the D-ADP complex only one dominant hinge point is identified
- upon substrate binding the collective interdomain motions become more directional
- the 1,3-BPG and nucleobase binding site become more flexible, while the phosphate and ribose binding site become more rigid upon substrate binding
- the local intradomain motions become more correlated upon substrate binding

5.2. Effects of substrate chirality

- the similar binding mode of phosphate chains of various nucleotide analogues may be responsible for the broad nucleotide specificity of hPGK
- the catalytic efficiency is related to the fluctuation properties and hydrophobic contacts of the nucleobase
- the ranking of nucleotides with respect to binding affinity towards hPGK is: D-CDP < L-CDP << L-ADP << D-ADP
- the lower catalytic efficiency of L-ADP complex (compared to the D-ADP complex) is related to the large flexibility of the collective interdomain motions in the shear direction
- the correlations of the local intradomain motions are stronger in the D-ADP complex compared to the L-ADP complex
- the similar binding modes of the various nucleotides differing in chirality is responsible for the low enantioselectivity of the enzyme

- the dynamic conditions for efficient catalytic activity: (i) stable nucleotide binding, (ii) significant hinge bending component of collective interdomain motions, (iii) correlating intra- and interdomain motions, (iv) low number of hinges
- observing the above mentioned conditions during a simulation, it can be established whether the query nucleotide is phosphorylated by hPGK efficiently

6. LIST OF OWN PUBLICATIONS

Publications in the topic of the thesis:

1. **Palmai Z**, Perahia D, Lionne C, Fidy J, Balog E, Chaloin L. (2011) Ligand chirality effects on the dynamics of human 3-phosphoglycerate kinase: comparison between D- and L-nucleotides. Arch Biochem Biophys, 511: 88-100.
2. **Palmai Z**, Chaloin L, Lionne C, Fidy J, Perahia D, Balog E. (2009) Substrate binding modifies the hinge bending characteristics of human 3-phosphoglycerate kinase: a molecular dynamics study. Proteins, 77: 319-329.

Poster presentations:

1. Influence of ligand chirality on the catalytic efficiency of human 3-phosphoglycerate kinase (Zoltan Palmai, David Perahia, Corinne Lionne, Judit Fidy, Erika Balog, Laurent Chaloin) - 55th Annual Meeting of the Biophysical Society, March 5-9 2011, Baltimore, Maryland, USA
2. Substrate binding directs the functional hinge bending motion of human 3-phosphoglycerate kinase (Erika Balog, Zoltan Palmai, Laurent Chaloin, Corinne Lionne, Judit Fidy, David Perahia) - PhD Tudományos Napok 2010, 2010. április 15-16. Semmelweis Egyetem, Budapest
3. A szubsztrát kötődés irányítottá teszi a humán foszfoglicerát kináz (PGK) funkcionális csuklómozgását (Pálmai Z., Chaloin L., Lionne C., Fidy J., Perahia D., Balog E.) – A magyar biofizika társaság XXIII. Kongresszusa, 2009. augusztus 23-26, Pécs
4. The influence of substrate binding on the functional hinge bending motion of human 3-phosphoglycerate kinase (Z. Palmai, L. Chaloin, C. Lionne, J. Fidy, D. Perahia, E. Balog) – Methods in Molecular Simulation Summer School 2009, 5-14 July 2009, Sheffield, UK
5. Interdomain interactions influence the subunit dynamics of PGK (Balog, E.,

Palmai, Z. and Fidy, J.) – Regional Biophysics Conference 21-25 August 2007 Balatonfüred, Hungary

6. The influence of interdomain interactions on intradomain motions in yeast phosphoglycerate kinase (E. Balog, Z. Palmai, J. Fidy) – 6th EBSA & British Biophysical Society Congress July 14-18 2007, Imperial College London, UK

7. Interdomain interactions influence the subunit dynamics of PGK (Erika Balog, Zoltan Palmai, Judit Fidy) – Joint meeting of Hungarian and German Biophysicists, Time and space resolved methods in molecular biophysics May 17 – May 20, 2007 St Bonifatius Monastery, Hünfeld, Germany