3620–3629 *Nucleic Acids Research, 2008, Vol. 36, No. 11* *Published online 7 May 2008*
doi:10.1093/nar/gkn212

Molecular basis for the lack of enantioselectivity of human 3-phosphoglycerate kinase

C. Gondeau[1], L. Chaloin[1], P. Lallemand[1], B. Roy[2], C. Périgaud[2], T. Barman[1], A. Varga[3], M. Vas[3], C. Lionne[1] and S. T. Arold[4,5,6,*]

[1]Centre d'études d'agents Pathogènes et Biotechnologies pour la Santé (CPBS), UMR 5236, CNRS–Universités Montpellier 1 et 2, Institut de Biologie, 4 bd Henri IV, CS69033, 34965 Montpellier cedex 2, [2]Institut des Biomolécules Max Mousseron (IBMM), UMR 5247, CNRS–Universités Montpellier 1 et 2, case courrier 1705, Université Montpellier 2, Place Eugène Bataillon, 34095 Montpellier cedex 5, France, [3]Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Karolina Strasse 29, H-1113 Budapest, Hungary, [4]CNRS, UMR5048, Centre de Biochimie Structurale, 34090 Montpellier, [5]INSERM, U554 and [6]Universités Montpellier 1 et 2, 34090 Montpellier, France

Received January 25, 2008; Revised March 18, 2008; Accepted April 9, 2008

ABSTRACT

Non-natural L-nucleoside analogues are increasingly used as therapeutic agents to treat cancer and viral infections. To be active, L-nucleosides need to be phosphorylated to their respective triphosphate metabolites. This stepwise phosphorylation relies on human enzymes capable of processing L-nucleoside enantiomers. We used crystallographic analysis to reveal the molecular basis for the low enantioselectivity and the broad specificity of human 3-phosphoglycerate kinase (hPGK), an enzyme responsible for the last step of phosphorylation of many nucleotide derivatives. Based on structures of hPGK in the absence of nucleotides, and bound to L and D forms of MgADP and MgCDP, we show that a non-specific hydrophobic clamp to the nucleotide base, as well as a water-filled cavity behind it, allows high flexibility in the interaction between PGK and the bases. This, combined with the dispensability of hydrogen bonds to the sugar moiety, and ionic interactions with the phosphate groups, results in the positioning of different nucleotides so to expose their diphosphate group in a position competent for catalysis. Since the third phosphorylation step is often rate limiting, our results are expected to alleviate *in silico* tailoring of L-type prodrugs to assure their efficient metabolic processing.

INTRODUCTION

Since the beginning of the 1990s and the discovery of lamivudine [3TC, (1,2)], L-nucleoside analogues, mirror images of the natural D-nucleosides analogues, emerged as a new class of antiviral agents (3). Two of the eight-nucleoside analogues targeting the HIV-1 reverse transcriptase and approved by the US Food and Drug Administration are L-type (3TC and FTC, emtricitabine). L-nucleosides are also used to treat cancer, and are emerging as agents against the hepatitis B virus (3). Their physiological activity however relies on the fortuitous capacity of a few human enzymes to phosphorylate these non-natural compounds into their triphosphate forms. The knowledge of the molecular basis for the relaxed enantioselectivity of these enzymes is therefore essential for the rational development of efficiently metabolized non-toxic L-type drugs for humans.

Phosphorylation of the non-natural L-nucleosides to the monophosphate counterpart involves the cytosolic deoxycytidine kinase (dCK) and the structurally homologous mitochondrial thymidine kinase 2 (TK2). For dCK, the L/D polyvalency results from malleable interactions between the active site and the nucleoside base. Combined with the intrinsic flexibility of the substrate, this allows the sugar moiety of L and D isoforms to establish similar polar interactions with the active site (4,5). Subsequent phosphorylation to diphosphate forms is carried out by the UMP/CMP kinase (UMP/CMPK) and the TMP kinase (TMPK) (6–8). Structures of UMP/CMPK and TMPK bound to D-, but not L-nucleosides, have been determined (9,10). Based on homology modelling of L-nucleoside-bound kinases, the relaxed enantioselectivity for UMP/CMPK and TMPK was suggested to arise from the lack of essential hydrogen bonds to the ribose, and the resulting conformational freedom of the sugar moiety (8).

The final step, the phosphorylation to L-nucleoside triphosphate homologues, is accomplished by the

*To whom correspondence should be addressed. Tel: +33 467 417 702; Fax: +33 467 417 913; Email: Stefan.Arold@cbs.cnrs.fr

3-phosphoglycerate kinase (PGK). This enzyme, which catalyses the reversible phospho-transfer between 1,3-bisphosphoglycerate (1,3-bPG) and MgADP, is active on a particularly broad spectrum of L- and D-nucleoside diphosphates, including antiviral agents (11–13). Structures of PGKs from a number of species have been determined (horse, pig, yeast, *Bacillus stearothermophilus, Thermotoga maritima* and *Trypanosoma brucei*), and structural data exist for unliganded PGKs (14–16), as well as their binary and ternary complexes with phosphoglycerate (PG) and/or D-ADP, D-ATP or stable analogues of D-ATP (17–25). Despite the wealth of information gained, these data fail to elucidate how human PGK (hPGK) is capable of recognizing such a broad spectrum of nucleosides, including purine and pyrimidine derivatives and their D- and L-enantiomers. However, because the accumulation of diphosphate analogues such as L-OddCDP, L-SddCDP or L-Fd4CDP in cells demonstrates that the third phosphorylation step is often rate limiting (26), the knowledge of the molecular basis for the broad specificity of hPGK would significantly alleviate *in silico* tailoring of L-type prodrugs to assure their efficient triphosphorylation. In view of this, we present here crystal structures of hPGK in its free state, or bound to L-ADP, D-ADP, D-CDP or L-CDP.

MATERIALS AND METHODS

Protein purification

Recombinant hPGK was produced in BL21 (DE3)/pDIA17 *Escherichia coli* cells harbouring the pgk-pET28a expression plasmid as previously described (12). Briefly, after lysis by sonication and ultracentrifugation, the supernatant was loaded onto a HisTrap FF column (GE Healthcare) and washed with buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8. The PGK was eluted with a linear gradient of imidazole (10–250 mM in the same buffer). Fractions containing the PGK were pooled, concentrated and dialysed against a buffer containing 20 mM Tris pH 7.5, 100 mM NaCl. Following cleavage of the N-terminal 6His-tag with thrombin (10 U/mg) at RT overnight, PGK was applied onto three columns in series: a Hitrap Benzamidine FF (GE Healthcare) to remove thrombin, a HisTrap FF (GE Healthcare) to remove the cleaved 6His-tag and the uncleaved 6His-tagged PGK, and finally a size exclusion chromatography column (HiLoad 16/60 Superdex 75 column, GE Healthcare). Fractions containing pure PGK in a 20 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT were pooled and concentrated to ~20–30 mg/ml. The mutant E343A of hPGK was prepared using the Quik Change mutagenesis kit (Stratagene) as decribed in Szabó *et al.* (27).

Enzyme activity measurements

The activity of the wild-type and E343A mutant hPGK were measured with either D- or L-MgADP and 1,3-bPG as substrates in an assay system coupled with GAPDH, as described earlier (11). Activity of the wild-type hPGK were also assayed in the reverse direction using the substrate D-MgATP, as described in the Supplementary Data.

Synthesis of L-ADP and L-CDP

L-ADP was synthesized as previously described (11). L-CDP was synthesized starting from peracylated L-ribose. First of all, silylated 4-*N*-benzoylcytosine was treated with 1-*O*-acetyl-2,3,5-*tri*-*O*-benzoyl-β-L-ribofuranose in 1,2-dichloroethane in presence of trimethylsilyl triflate (TMSTf) to afford the fully protected L-cytidine (28). Debenzoylation of the protected nucleoside in methanolic ammonia yielded L-cytidine in 53% yield starting from peracylated L-ribose. Secondly, L-CMP was obtained in 48% yield by selective 5′-phosphorylation of L-cytidine with phosphorus oxychloride in triethylphosphate (8). In the last stage, 1,1′-carbonyldiimidazole (162 mg, 1 mmol) was added to a stirred solution of the tri-*n*-butylammonium salt of L-CMP (84 mg, 0.168 mmol) in anhydrous DMF (3.3 ml) at room temperature under argon. After 3 h of stirring at room temperature, methanol (80 µl) was added to the reaction mixture. After 30 min, a solution of tri-*n*-butylammonium phosphate (238 mg, 0.84 mmol) in 0.915 ml DMF was added. The reaction was controlled by TLC using as eluant 2-propanol/NH_4OH/water (11/7/2, v/v/v). The solution was stirred for 1 day at room temperature. The solvent was then evaporated under reduced pressure to give a residue which was diluted with water and purified by ion exchange chromatography using a DEAE-Sephadex A-25 (elution: gradient of TEAB pH 7.5 from 10 mM to 1 M) followed by chromatography on RP18 (elution with water). The sodium salt of the title compound was obtained after anionic exchange on DOWEX-AG 50WX2-400 column, the corresponding fractions were lyophilized to give 16 mg (21%) of L-CDP.

L-cytidine: $[\alpha]_D^{20}$ -28 (*c* 1.0, H_2O); ^{1}H NMR (D_2O, 300 MHz) δ 7.82 (d, 1 H, J = 7.6 Hz, H_6), 6.03 (d, 1 H, J = 7.6 Hz, H_5), 5.88 (d, 1 H, J = 3.9 Hz, $H_{1''}$), 4.30–4.27 (dd, 1 H, $H_{2'}$), 4.20–4.16 (dd, 1 H, $H_{3'}$), 4.13–4.09 (m, 1 H, $H_{4'}$), 3.94–3.76 (m, 2 H, $H_{5'}$, $H_{5''}$); ^{13}C NMR (D_2O, 300 MHz) δ 166.1, 157.5, 141.6, 96.1, 90.2, 83.7, 73.8, 69.2, 60.6; MS FAB$^+$ *m*/*z* 244 $(M + H)^+$; HRMS ($C_9H_{14}O_5N_3$), calcd 244.0933, found 244.0924. L-CMP: $[\alpha]_D^{20}$ -8 (*c* 1.0, H_2O); ^{1}H NMR (D_2O, 300 MHz) δ 8.01 (d, 1 H, J = 7.5 Hz, H_6), 6.09 (d, 1 H, J = 7.5 Hz, H_5), 5.97 (d, 1 H, J = 1.5 Hz, $H_{1'}$), 4.30 (sl, 2 H, $H_{2'}$, $H_{3'}$), 4.24 (sl, 1 H, $H_{4'}$), 4.06–3.99 (m, 2 H, $H_{5'}$, $H_{5''}$); ^{13}C NMR (D_2O, 300 MHz) δ 165.9, 157.4, 141.6, 96.4, 89.2, 83.1, 74.2, 69.4, 63.4; ^{31}P NMR (D_2O, 300 MHz) δ + 1.58 (s); MS FAB$^+$ *m*/*z* 346 $(M + H)^+$. L-CDP: $[\alpha]_D^{20}$ -12 (*c* 1.0, H_2O); ^{1}H NMR (D_2O, 300 MHz) δ 7.99 (d, 1 H, J = 7.5 Hz, H_6), 6.10 (d, 1 H, J = 7.5 Hz, H_5), 5.95 (d, 1 H, J = 3.5 Hz, $H_{1'}$), 4.41–4.39 (m, 1 H, $H_{3'}$), 4.31–4.28 (m, 1 H, $H_{2'}$), 4.22 (m, 3 H, $H_{4'}$ $H_{5'}$, $H_{5''}$); ^{13}C NMR (D_2O, 300 MHz) δ 165.5, 156.9, 140.8, 95.7, 88.7, 81.8, 73.6, 67.9, 63.0; ^{31}P NMR (D_2O, 300 MHz) δ –6.1 (d, $J_{P–P}$ = 21.8 Hz), –10.7 (d, $J_{P–P}$ = 21.8 Hz).

Crystallization

Crystals of human PGK were obtained by vapour diffusion in hanging drops, within 1–3 weeks at 18°C. The protein solution (10–15 mg/ml) in Tris-HCl pH 7.0 contained 10 mM ADP, 25 mM $MgCl_2$, 50 mM PG, 10 mM DTT and 0.02% NaN_3. The drop contained 1 µl of the protein solution mixed with 1 µl of the reservoir solution. Reservoir solutions consisted of 2.5–2.7 M $NaKPO_4$ pH 8.1–8.4. Crystals of hPGK complexed with PG, or with PG and Mg-D-ADP or Mg-L-ADP were obtained by co-crystallization, whereas crystals of hPGK·PG·Mg-D/L-CDP were produced by soaking hPGK·PG with 75–150 mM of Mg-D/L-CDP for 1–5 days.

Data collection and processing, and structure determination and refinement

X-ray data were collected at beamlines ID14 and ID29 of the European Synchrotron Radiation Facility in Grenoble, France, at 100 K. Prior to data collection, crystals were cryoprotected in mother liquid supplemented with 25% glycerol, and then flash-cooled in liquid nitrogen. For details see Table 1. Data were integrated using MOSFLM and reduced with SCALA. The structure was determined by molecular replacement [MOLREP (29)] using pig muscle PGK as template (PDB entry 1VJC (20). Refinement was carried out with REFMAC using translation-liberation-screw (TLS) refinement (29). After initial rounds of restrained refinement with unliganded PGK, ligand densities were best visible using 3fo-2fc maps calculated with CNS (30), and 2fo-fc maps after ARP/WARP (31). Corresponding ligands were built into these maps using COOT (32), and refined using TLS refinement (REFMAC). The final models show good stereochemistry, as assessed using MolProbity (http://molprobity.biochem.duke.edu/). Coordinates were deposited in the Protein Data Bank (PDB IDs 2ZGV, 3C39, 3C3A, 3C3B, 3C3C for hPGK·D-ADP, hPGK·PG, hPGK·PG·L-ADP, hPGK·PG·D-CDP and hPGK·PG·L-CDP, respectively).

Molecular modelling

Manual docking was carried out by substituting L-CDP of the hPGK·PG·L-CDP structure with the nucleotide analogue of choice. The resulting complex was subjected to mild minimization/structure idealization using REFMAC (29). Automated *in silico* docking of L-type nucleotide analogues into hPGK was carried out as described (33).

RESULTS AND DISCUSSION

Overall structure, lobe orientation and ligand binding

Purified recombinant hPGK was assessed for enzymatic activity (Supplementary Data), and crystallized. Structures of hPGK were determined by molecular replacement, to resolutions of 2.4 Å or better (Table 1). As described for PGKs of other species, hPGK forms two lobes, each composed of a central parallel β-sheet sandwiched by helices, and connecting loops. The C-terminal lobe (residues 186–416) harbours also a short β-barrel domain. The C-domain binds the nucleotide, the N-domain the PG (Figure 1, Table 2). D-ADP-bound hPGK crystallized in the space group $P2_1$, with one molecule per asymmetric unit (AU). All other hPGK crystals (hPGK·PG·L-ADP, hPGK·PG·L-CDP, hPGK·PG·D-CDP and hPGK·PG) belonged to space group P1, with two molecules per AU. The P1 crystals do not simply constitute a $P2_1$ form with a symmetry break, but have a distinct protein packing arrangement, despite very similar crystallization conditions. The relative position of the two lobes in the $P2_1$ hPGK·D-ADP structure differs as compared to those of the

Table 1. Data collection, phasing and refinement statistics[a]

	hPGK·PG	hPGK·D-ADP:	hPGK·PG·L-ADP:	hPGK·PG·D-CDP:	hPGK·PG·L-CDP:
Data collection					
Space group	P1	$P2_1$	P1	P1	P1
Cell dimensions					
a, *b*, *c* (Å)	35.61 55.79 93.23	35.96 106.63 50.35	35.63 55.95 95.44	35.58 55.59 93.80	35.79 55.82 94.50
α, β, γ (°)	79.18 84.33 83.23	90.0 97.30 90.00	101.88 95.20 82.91	78.91 84.31 82.95	78.94 84.51 83.22
Wavelength (Å)	0.933	0.934	0.934	0.933	0.933
Resolution (Å) [b]	1.85 [1.95]	2.00 [2.24]	2.30 [2.30]	1.80 [1.80]	2.40 [2.4]
R_{merge}	0.048 (0.340)	0.065 (0.323)	0.081 (0.318)	0.042 (0.370)	0.126 (0.595)
$I/\sigma I$	14.0 (1.9)	11.4 (2.9)	9.3 (2.7)	14.6 (3.0)	11.0 (2.0)
Completeness (%)	77.1 (50.4)	69.0 (31.0)	84.9 (55.3)	87.2 (50.2)	92.3 (92.3)
Redundancy (%)	1.9 (1.8)	2.6 (2.0)	2.1 (2.1)	1.9 (1.8)	3.1 (3.1)
Refinement					
Resolution (Å)	1.85	2.00	2.30	1.80	2.40
No. reflections	44 499	16 987	26 390	54 699	24 806
R_{work}/R_{free}	0.19/0.27	0.17/0.26	0.18/0.28	0.17/0.24	0.19/0.28
R.m.s. deviation					
Bond lengths (Å)	0.021	0.017	0.014	0.025	0.014
Bond angles (°)	1.9	1.7	1.7	2.0	1.5
Composition of ASU	PG: 2, hPGK: 2	D-ADP: 1, hPGK: 1	PG: 2, L-ADP: 1, hPGK: 2	PO_4: 2, D-CDP: 1, hPGK: 2	PG: 2, L-CDP: 1, hPGK: 2

[a]Values in parentheses correspond to the last resolution shell.
[b]Values in brackets correspond to high resolution limit with cumulative completeness >80%.

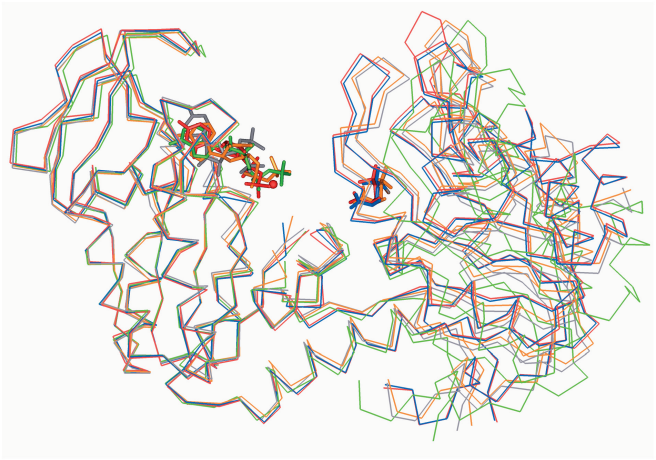

Figure 1. Superimposition of hPGK structures. Cα trace representations of hPGK·L-ADP (molecule B), orange; hPGK (PG-bound, nucleotide-free molecule A), blue; hPGK·L-CDP (molecule A), red; hPGK·D-ADP, green; hPGK·D-CDP (molecule B), grey. Structures were superimposed on their C-terminal lobe.

P1 crystals, grown in presence of L-ADP or without nucleotide (CDP derivatives were obtained by soaking of hPGK·PG crystals) (Figure 1). In the $P2_1$ form PG was absent in the PGK-binding site. In this structure, the relative lobe position is accompanied by a ~2.4 Å shift of the helix formed by residues 164 to 170, as compared to structures where PG was bound. Because R170 is an important constituent of the PG-binding site, this shift destroyed the PG-binding pocket. Since the $P2_1$ crystals were grown in the presence of PG, the observed form without PG is likely due to the antagonizing effect which occurs between PG and D-ADP, but not between PG and L-ADP (33). This connection between lobe opening and PG binding corroborates observations by Harlos *et al.* (21).

The lobe opening of both molecules in the P1 AUs differs also to some extent, whereas corresponding molecules of the different P1 AUs have similar lobe constellations. In the other structures, PG was found in all molecules, except in both molecules of the AU of hPGK·D-CDP, where it was replaced by a phosphate group from the crystallization buffer. The PG-binding site was not altered, inferring that this was due to changes in the relative concentrations of PG and phosphate from the buffer. In all P1 forms, only one hPGK molecule of the AU showed a bound nucleotide. The structural explanation for this might be that the loop connecting the first β-strand of the C-lobe with the following helix is moved consistently towards the nucleotide by about 0.6–0.8 Å in the nucleotide-bound molecules. Together, our observations support that lobe opening allosterically introduces deformations in the ligand binding sites, which influence ligand binding. Indeed, lobe closure has been shown to partake in the catalytic mechanism of PGK (18,19).

Invariably in all our structures, electron density for helix 13 was absent, indicating unfolding and/or very high flexibility of this region, as already observed by Szilágyi *et al.* (25). The propensity to unfold might have been reinforced by crystal contacts pushing onto residues 386–387, immediately upstream of helix 13. In all

Table 2. Distances (Å) between substrates and protein residues

Nucleotide groups	N6-Adenine or N4- Cytidine	O2-Cytidine	C2-Adenine or C6-Cytidine	C2-Adenine or C4-Cytidine	C2-Adenine or C4-Cytidine	C6-Adenine or C2-Cytidine	2′-OH (ribose)	3′-OH (ribose)	O4′ (ribose)	α-Phosphate (oxygen)	β-Phosphate (oxygen)	–	β-Phosphate
PG	–	–	–	–	–	–	–	–	–	–	–	–	1-C of PG
hPGK residues	Gly312 (β strand Q)	Gly238 (β strand H)	Val341 (loop after β strand J)	Phe291 (β strand O)	Leu256 (loop between α helices 9 and 10)	Gly238 (β strand H)	Glu343 (loop after β strand J)	Glu343 (loop after β strand J)	Ala214 (loop before α helix 8)	Lys219 (α helix 8)	Asn336 (β strand J)	Gly372(N)-Ser392(O)	–
D-ADP	2.89	–	4.08	4.11	3.30	3.62	2.62	2.31	3.49	3.62	7.04	4.03	–
L-ADP	3.58	–	3.89	3.80	4.33	4.01	2.51	3.16	4.01	3.39	9.68	4.09	10.46
D-CDP	5.36	2.74	3.59	4.01	3.58	3.80	**10.80**	**9.08**	5.03	5.95	8.95	3.27	–
L-CDP	3.50	3.21	4.13	5.15	4.50	4.11	2.89	2.73	3.43	2.76	7.46	2.82	11.13

Values in bold correspond to the distance between E343 and the hydroxyl groups of the D-CDP ribose.

molecules B of the P1 form, electron density for residues 99–106 was missing. In molecules B of hPGK·PG·L-ADP, hPGK·D-ADP, hPGK·PG·D-CDP but not in hPGK·PG, a disulfide bridge formed between cysteine residues 98 and 107. The differences in this N-lobe region, which is situated on the far side of the PG-binding site, are not expected to be connected with ligand binding, but probably with crystal packing, and potency of reducing agents present in the crystallization liquor.

Molecular basis for nucleotide promiscuity

D-ADP bound to loops above the N-termini of the central β-strands of the C-lobe of hPGK, as described for PGKs from other species (Figures 2, 3A and B, Table 2) (23,25). The adenine moiety adopts an *anti* conformation, and inserts into a conserved hydrophobic cleft formed by A214, G238, L256, F291, G312, L313 and V341. Insertion of the adenine encloses a number of water molecules in a cavity behind and above the base. The sugar, which is mostly solvent accessible, is held in place by hydrophobic contacts with P338, and two hydrogen bonds between its hydroxyl groups and E343. Both phosphates lie on the surface of the C-terminal lobe, with the β-phosphate pointing towards the PG-binding site on the N-lobe. The α-phosphate is positioned by ionic interactions with K219, and possibly K215. The β-phosphate is repulsed from the C-lobe, and directed towards the PG by D218. The interaction of the ADP β-phosphate with the NH_2 group of N336, observed in other ADP-bound PGK structures (17,19) is absent in the present structure, possibly due to the high mobility of the nearby helix 13.

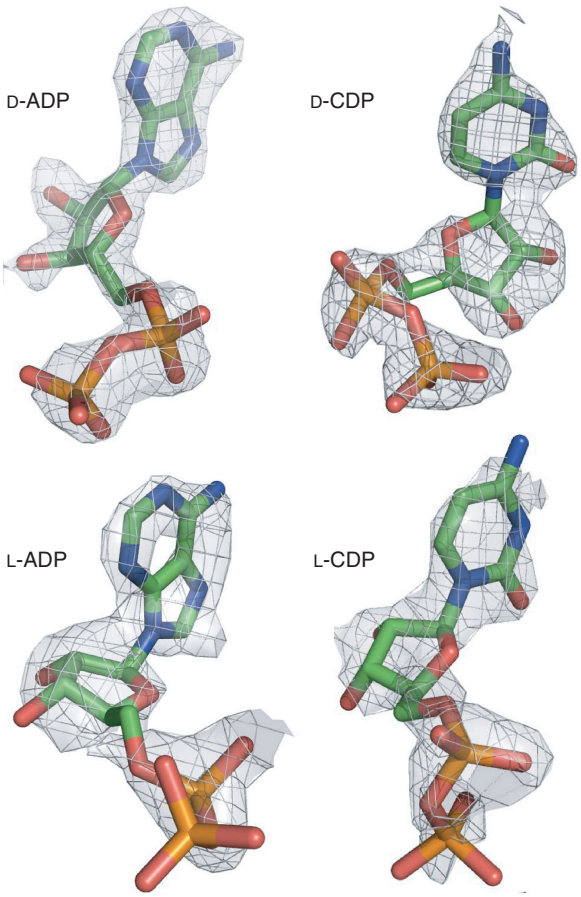

Figure 2. Superimposition of nucleotides with their 2fofc omit maps. Final models are shown. Maps were produced by omitting the nucleotide from the model. Prior to map calculation, model bias was removed by subjecting all atoms of the remaining model to 0.15 Å random shifts, resetting their B factors to the value suggested by the Wilson plot, and refinement of the resulting model without nucleotide.

In the hPGK·PG·L–ADP complex, the position of the β-phosphate was only poorly defined in the omit electron density maps (Figure 2). The hydrophobic interactions with the base and sugar, hydrogen bonds to sugar hydroxyls and the position of the α-phosphate are conserved. For this, the base of L-ADP adopts a *syn* conformation and is tilted by about 40° in the binding site, inserting its amine 1.5 Å deeper into the base-binding pocket than D-ADP (Supplementary Figure). This adjustment is rendered possible by the malleable water network in the cavity behind the base, and the promiscuous hydrophobic clamp to the adenine (Figure 3A and B, Table 2). As a result of the *syn* conformation, the hydroxyls of the L-ADP ribose can still form D-ADP-like hydrogen bonds with E343. The 1.5-Å difference in hydroxyl position is absorbed by a concomitant movement of the E343 side chain, but nevertheless leads to an increase in the hydrogen-bond length between E343 and the 3′-OH of the L-ADP ribose (Table 2). Due to the flexibility of the sugar-phosphate link, the L-ADP α-phosphate is placed in the same position as its D-ADP counterpart, despite differences in the sugar orientation. The comparison of L- and D-ADP-bound hPGK suggested that the difference in catalytic efficiency between both enantiomers stemmed mainly from differences in E343-hydrogen bonding. This was corroborated by our kinetic measurements with a hPGK E343A mutant: While D-ADP displayed an over 2-fold lower K_m value for wild-type hPGK than L-ADP (0.12 and 0.27 mM, respectively), K_m values for D- and L-ADP were very similar when measured with the E343A mutant (1.24 and 1.35 mM, respectively). Also the k_{cat} values of hPGK E343A for both enantiomers were highly comparable (41 and 40 s^{-1} for the D- and L-ADP, respectively), inferring that different catalytic processing of D- and L-ADP is at least in part due to the differences in E343-ribose hydrogen bond strength. While these data confirm that the hydrogen bonds between E343 and the hydroxyl groups of the ribose are not essential for catalysis, as suggested previously by Gallois-Montbrun *et al.* (12), they also affirm that E343-sugar hydrogen bonds are potent enhancers of the catalytic efficiency.

Although the L-ADP β-phosphate was not well resolved in the electron density, its most probable position infers an Mg^{2+}-mediated link to the carboxylate of D218. This binding mode resembles to the one observed in pig muscle PGK for the ATP-analogue AMP-PCP (22). In this previous work, it was concluded that D218, present in helix 8, represents an alternative binding site for the nucleotide phosphates (in addition to D374, located in helix 13), which might be crucial for catalysis. In agreement, the present structure supports an interaction of

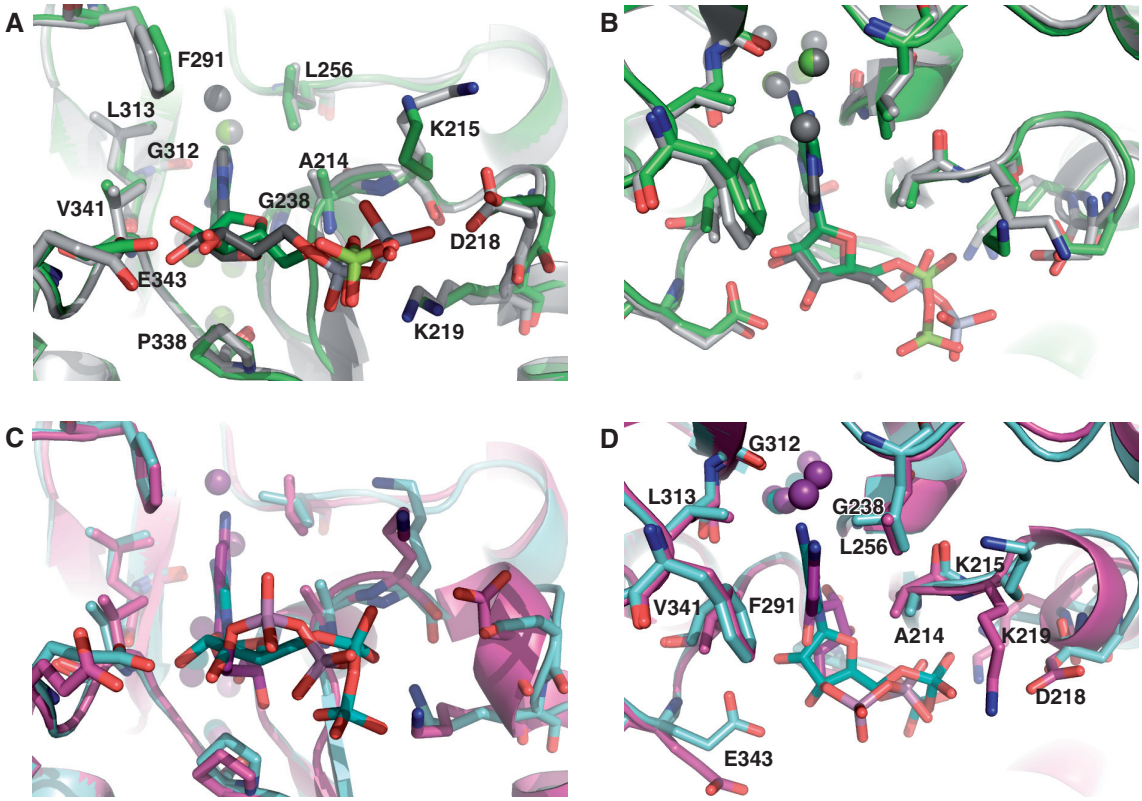

Figure 3. Details of the nucleotide interactions. (**A**, **B**) Two 90° views of hPGK·L-ADP (molecule B, light grey), hPGK·D-ADP (green). Waters are presented as green and grey spheres. (**C**, **D**) Two 90° views of hPGK·L-CDP (light blue), and hPGK·D-CDP (magenta). Waters from are presented as blue and magenta spheres. hPGKs were superimposed on their C-terminal lobe. (C) Two 90°-views as compared to (A). Only the nucleotides are shown to illustrate their spatial dispersion in the binding site. The β-phosphates are indicated.

Table 3. Catalytic efficiency (k_{cat}/K_m, $M^{-1}s^{-1}$) of hPGK, or hPGK E343A(*) with different substrates

Substrate	D-enantiomer	L-enantiomer	Ratio L/D	Reference
Purine				
ADP(*)	9.0×10^6	2.5×10^6	0.28	This study and (11)
ADP(*)	3.3×10^4	3.0×10^4	0.90	This study
ATP	1.7×10^6	2.0×10^5	0.12	(12)
2′dATP	3.0×10^4	2.6×10^4	0.87	(12)
2′3′ddATP	2.3×10^4	–	–	(12)
Pyrimidine				
CTP	150	5200	35	(12)
2′dCTP	<5	1200	>240	(12)
2′3′ddCTP	<5	–	–	(12)

D218 with K215, one of the essential catalytic residues of PGK (34).

PGK is able to act on D- and L-pyrimidine nucleotides, but less efficiently than on purine nucleotides (Table 3). In contrast to ATP however, L-CTP is a much better substrate than D-CTP (12). Based on the available structural and biochemical data, the same relationship is expected for L- and D-CDP. Our crystal structure of D-CDP-bound hPGK shows that the base of D-CDP is held in a similar position as the D/L-ADP adenines (Figures 2, 3C and D, Supplementary Figure). However, the pyrimidine is shifted about 2.5 Å further away from the central β-sheet than the purine. This displacement is buffered by the water-filled cavity above the base, whereas shape differences between adenine and cytosine are accommodated by the water network behind the base. Consequently, the sugar moiety is sucked 2 Å deeper into the binding cleft, and, in turn, the sugar hydroxyls can not longer interact with E343 (Table 2), explaining the low catalytic efficiency of D-CDP (Table 3). The inward motion also positions the D-CDP β-phosphate at roughly the same position as the D-ADP α-phosphate. As a result, the distance between the PG and the β-phosphate increases by about 3 Å in hPGK·PG·D-CDP as compared to hPGK(·PG)·D/L-ADP. Thus, the low catalytic turnover of D-CDP is explained by the lack of hydrogen bonds to

E343 and the enhanced distance between β-phosphate and PG.

In the hPGK·PG·L-CDP complex, the position of E343 and conformation of the nucleotide inferred the presence of a hydrogen bond to the sugar hydroxyls, as seen in the ADP structures (Figure 3C and D, and Table 2). Nevertheless, based on the kinetic results with L-CTP and L-ATP (Table 3), L-CDP is expected to be a less efficient substrate than D- or L-ADP. This might be explained by the (i) weaker hydrogen bond of E343 to the ribose hydroxyls (in the L-CDP structure, both hydrogen bond lengths are over 2.7 Å, and the angles less favourable than in the ADP structures), and (ii) by the altered positioning of the β-phosphate, placing it about 1.5 Å further from the PG, than in D- or L-ADP.

Our structural results correlate well with published biochemical and functional data (Table 3), and explain how L and D forms of purine and pyrimidine nucleotides can be used by hPGK as substrates: The broad specificity of PGK results mainly from fixing the nucleotide base with an unspecific hydrophobic clamp, which allows large movements of the base in directions of the base plane. The nucleotide-binding site provides enough space around that clamp to allow interactions with different bases, and different base orientations of L- and D-enantiomers of the same nucleotide. Additionally, the sugar moiety is mostly solvent accessible and thus free to rotate and engage hydrogen bonds with E343 in all tested nucleotides except D-CDP. D-CDP is still catalysed, although at a very low rate, because the E343-ribose hydrogen bonds are not strict requirements for catalysis, and even without them phosphate groups can be oriented similarly by electrostatic guidance. The difference in the position of the D-CDP β-phosphate can probably be accounted for by an increased domain–domain closure during catalysis, owing to the marked flexibility of the PGK two-domain fold.

Modeling prodrug–hPGK interactions

Structural information would be of great help to improve the poor processing of several L-type prodrugs by hPGK. However, producing these crystal structures is not trivial. Firstly, the diphosphate compounds are not commercially available and need to be synthesized, and secondly, due to their poor affinity for hPGK, extremely high prodrug concentrations need to be used for crystallization, causing most crystals to disintegrate. By providing structural insights into hPGK bound to L- and D-type purine and pyrimidine-based diphosphate nucleosides, we now offer enough information to allow straightforward homology modeling of the complexes formed between hPGK and nucleotide analogues; in fact, insightful models can simply be obtained by replacing the nucleotide of the closest experimental structure with the nucleotide analogue. Optionally, freely available software such as REFMAC (29) could be used to subject the model to mild energy minimization. Accordingly, we modelled the complexes formed between hPGK and antiviral diphosphate analogues such as L-OddCDP, L-SddCDP, L-Fd4CDP or L-FMAU (Supplementary Figure 2). The validity of this uncomplicated manual approach was corroborated by the fact that the obtained structures were very similar to the ones obtained by a full-scale *in silicio* docking approach, using genetic algorithms. Inspection of the obtained complexes in the light of the experimental results obtained here strongly suggested that the limiting factor for catalysis of L-OddCDP, L-SddCDP or L-Fd4CDP is not the modified base, which can be accommodated by the non-specific base-clamp of hPGK, but the absence of E343-ribose hydrogen bonds. The anti-HBV drug candidate L-FMAU preserves the O3′ hydrogen bond to E343, explaining why this nucleoside is efficiently metabolized to its triphosphate state (35).

Comparison of PGK with dCK, UMP/CMPK and TMPK: different mechanisms, but common fold

The molecular basis for the relaxed enantioselectivity of hPGK contrasts in several aspects with the one described for dCK, UMP/CMPK and TMPK. These three enzymes consist of only one domain, and substrate (phosphate acceptor) and ATP (phosphate donor) need to be fixed on this single domain in an orientation readily competent for catalysis (Figure 4). For this, the entire substrate is completely engulfed in a deep pocket, and oriented towards the γ-phosphate of bound ATP. In dCK, hydrophobic interactions tolerate a $\sim 10^{\circ}$ tilt of the cytosine moiety perpendicular to the base plane, which, concomitant with changes in the exocyclic and glycosidic bond torsion angles, allows similar positioning of the sugar hydroxyls of the L- and D-enantiomers, and hence similar hydrogen bonds between dCK and the ribose moieties (4,5).

In the UMP/CMPK and TMPK the sugar moiety of bound substrates is placed inside a water-filled cavity (9,36). The nucleoside is mainly being held in place by malleable hydrophobic interactions with its base, and ion bonds to its α-phosphate. Accordingly, a recent *in silico* docking analysis (8) suggested that D/L polyvalency of UMP/CMPK and TMPK arises from the positional freedom of the sugar moiety. This allows absorbing the structural differences between D and L forms, while keeping the base and phosphate in similar positions in both enantiomers. In view of this, the cavity around the sugar appears to provide multiple possibilities for hydrogen bonds to the ribose hydroxyls in D and L orientations, and, while enhancing catalytic activity, these bonds are not essential for it.

Importantly, nucleoside (mono/diphosphate) kinases such as dCK, UMP/CMPK, TMPK or the nucleoside diphosphate kinase (NDPK) use ATP or other nucleoside triphosphates as a phosphate donor to phosphorylate a nucleoside (mono/diphosphate) substrate. In the process of nucleoside prodrug metabolism, the non-natural nucleoside analogues take the place of the nucleoside substrate in these enzymes. Conversely, PGK uses 1,3-bPG as phosphate donor to phosphorylate ADP, and nucleoside diphosphate analogues take the place of ADP in PGK. Since ADP is abundant in the cell, PGK does not need, and probably has not evolved, stringent selection criteria for it. Thus, the relaxed specificity of PGK towards non-natural nucleoside diphosphate

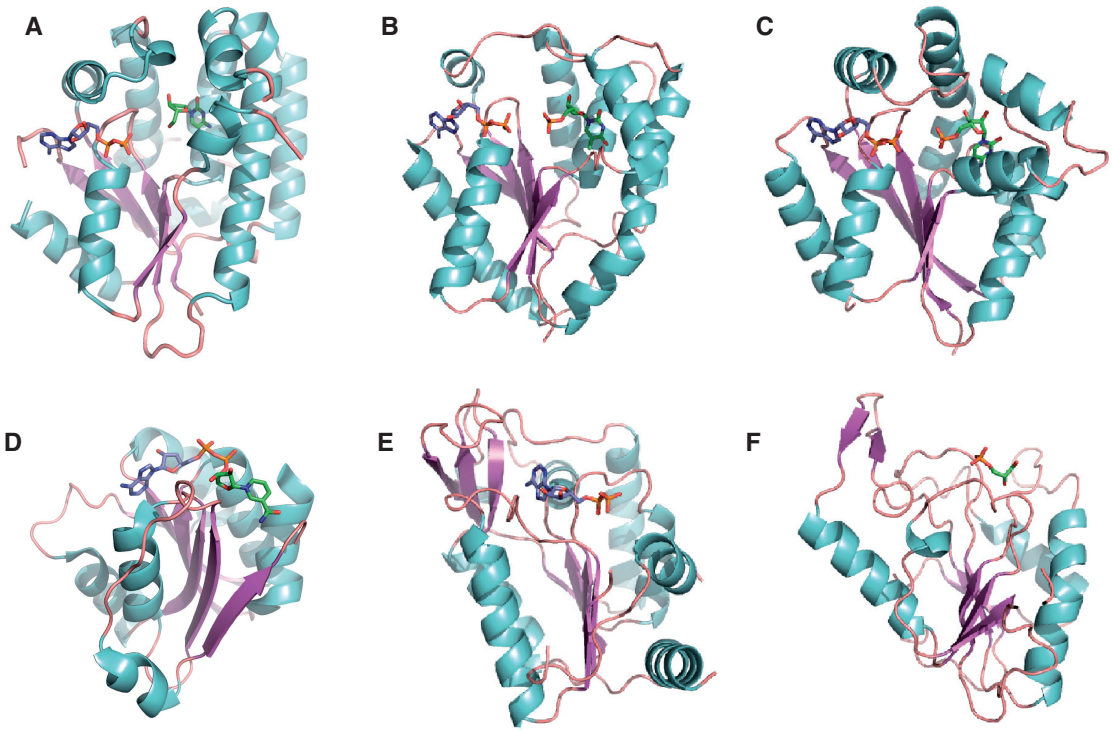

Figure 4. Comparison of hPGK, dCK, TMPK, ADH and UMP/CMPK. Secondary structures are coloured according to magenta: β-strands, cyan: α-helices, salmon: loops. Ligands are shown in stick representation, with carbon atoms of ADP moieties coloured in blue, and of substrates or nicotinamide moiety in green. (**A**) dCK (PDB entry 2NO1), (**B**) TMPK (PDB entry 1E2D), (**C**) UMP/CMPK (PDB entry 2UKD), (**D**) NADH-bound liver alcohol dehydrogenase [PDB entry 2OHX (38)], (**E**) hPGK C-domain and (**F**) hPGK N-domain.

analogues might be caused by the absence of a stringent selection mechanism towards its natural ADP substrate.

Despite functional and structural differences, dCK, UMP/CMPK, TMPK and PGK display intriguing resemblances in their fold and position of the substrate-binding site. Each of the two lobes of PGK, as well as the individual domains of dCK, UMP/CMPK and TMPK, can be related to the same basic fold: a central five-stranded β-sheet sandwiched by helices and loops (Figure 4). Admitting deletions or insertions of one strand–loop/helix unit, the topology of all domains is the same. Also, the ligand-binding positions, in loop regions over the central β-sheet, are very comparable in all these domains. Moreover, the ADP and PG molecules bound to PGK are oriented in the same way as ADP and substrate in dCK, UMP/CMPK and TMPK, respectively. This fold and nucleotide orientation are also found in other nucleotide-binding enzymes, such as NADH in alcohol dehydrogenase (ADH). This could indicate that these enzymes have evolved from the same one-domain nucleotide-binding ancestor. Thus, although not apparent in its intron/exon pattern (37), the 3D structure comparison supports that PGK has originated by gene duplication and separation of substrate and phosphate-donor-binding sites onto different lobes.

CONCLUSION

For an L-nucleoside prodrug to be physiologically active, it needs to be a substrate for three different enzymes to accomplish the three subsequent phosphorylation steps. With the study presented here, structural explanations do now exist for the relaxed enantioselectivity of enzymes from all three phosphorylation steps, even though the modelling study on UMP/CMPK and TMPK would require an experimental confirmation. This knowledge should help conceiving efficiently metabolized compounds, and to screen and optimize existing pharmacophors *in silico*. Nevertheless, because the selection mechanisms differ for hPGK, dCK and probably UMP/CMPK and TMPK, it will be challenging to produce compounds that are excellent substrates for all of them, and still act on their viral or cellular target. Given (i) the preference of dCK, UMP/CMPK and TMPK for pyrimidine-based compounds, (ii) the malleability of the interactions with the base, especially in hPGK, (iii) the importance of interactions with the base and α-phosphate in UMP/CMPK, TMPK and hPGK and (iv) that hydrogen bonds to the ribose are non-essential, but catalytically favourable, it appears a good strategy to base prodrugs on L-type pyrimidine analogues which preserve at least one ribose oxygen for hydrogen bonding with E343, rather than using dideoxyribose nucleosides. L-FMAU illustrates that preservation of one ribose oxygen—and hence good catalytic efficiency—are compatible with antiviral activity. Metabolic bottlenecks could thus be amended by reinforcing hydrogen bonds to the ribose in the concerned enzyme. In view of this, modification of the base might help by reorienting the sugar moiety in the binding site.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank ESRF beamline staff, as well as Gilles Labesse and Marie-Line Garron for their assistance in data recording. We are grateful to Dominique Deville-Bonne for her gift of the E. coli cells transformed with the pgk-pET28a plasmid, help in the production in purification of hPGK and stimulating discussions. We thank Gábor Szabó and Giulia Formichi for critical comments on the manuscript, and Lionel Verdoucq for the use of the Cell Disruptor and technical advice. The present publication was prepared within the framework of the Hungarian-French Intergovernmental Scientific and Technological Cupertino Program and was supported by the grant (OMFB-00493/2007, Project No. F-43/2006) from the Hungarian Foundation of Research and Innovative Technology as well as by Egide (Balaton program No. 14100ZF), and by financial support from the 'Agence Nationale de Recherches sur le Sida'. Funding to pay the Open Access publication charges for this article was provided by CNRS, INSERM and Genopole Languedoc-Roussillon.

Conflict of interest statement. None declared.

REFERENCES

1. Coates,J.A., Cammack,N., Jenkinson,H.J., Mutton,I.M., Pearson,B.A., Storer,R., Cameron,J.M. and Penn,C.R. (1992) The separated enantiomers of 2′-deoxy-3′-thiacytidine (BCH 189) both inhibit human immunodeficiency virus replication *in vitro*. *Antimicrob. Agents Chemother.*, **36**, 202–205.
2. Schinazi,R.F., Chu,C.K., Peck,A., McMillan,A., Mathis,R., Cannon,D., Jeong,L.S., Beach,J.W., Choi,W.B., Yeola,S. *et al.* (1992) Activities of the four optical isomers of 2′,3′-dideoxy-3′-thiacytidine (BCH-189) against human immunodeficiency virus type 1 in human lymphocytes. *Antimicrob. Agents Chemother.*, **36**, 672–676.
3. Mathé,C. and Gosselin,G. (2006) L-nucleoside enantiomers as antivirals drugs: a mini-review. *Antiviral Res.*, **71**, 276–281.
4. Sabini,E., Hazra,S., Konrad,M., Burley,S.K. and Lavie,A. (2007) Structural basis for activation of the therapeutic L-nucleoside analogs 3TC and troxacitabine by human deoxycytidine kinase. *Nucleic Acids Res.*, **35**, 186–192.
5. Sabini,E., Hazra,S., Konrad,M. and Lavie,A. (2007) Nonenantioselectivity property of human deoxycytidine kinase explained by structures of the enzyme in complex with L- and D-nucleosides. *J. Med. Chem.*, **50**, 3004–3014.
6. Ostermann,N., Segura-Pena,D., Meier,C., Veit,T., Monnerjahn,C., Konrad,M. and Lavie,A. (2003) Structures of human thymidylate kinase in complex with prodrugs: implications for the structure-based design of novel compounds. *Biochemistry*, **42**, 2568–2577.
7. Pasti,C., Gallois-Montbrun,S., Munier-Lehmann,H., Véron,M., Gilles,A.M. and Deville-Bonne,D. (2003) Reaction of human UMP-CMP kinase with natural and analog substrates. *Eur. J. Biochem.*, **270**, 1784–1790.
8. Alexandre,J.A., Roy,B., Topalis,D., Pochet,S., Périgaud,C. and Deville-Bonne,D. (2007) Enantioselectivity of human AMP, dTMP and UMP-CMP kinases. *Nucleic Acids Res.*, **35**, 4895–4904.
9. Schlichting,I. and Reinstein,J. (1997) Structures of active conformations of UMP kinase from *Dictyostelium discoideum* suggest phosphoryl transfer is associative. *Biochemistry*, **36**, 9290–9296.
10. Segura-Pena,D., Sekulic,N., Ort,S., Konrad,M. and Lavie,A. (2004) Substrate-induced conformational changes in human UMP/CMP kinase. *J. Biol. Chem.*, **279**, 33882–33889.
11. Varga,A., Szabo,J., Flachner,B., Roy,B., Konarev,P., Svergun,D., Zavodszky,P., Périgaud,C., Barman,T., Lionne,C. *et al.* (2008) Interaction of human 3-phosphoglycerate kinase with L-ADP, the mirror image of D-ADP. *Biochem. Biophys. Res. Commun.*, **366**, 994–1000.
12. Gallois-Montbrun,S., Faraj,A., Seclaman,E., Sommadossi,J.P., Deville-Bonne,D. and Véron,M. (2004) Broad specificity of human phosphoglycerate kinase for antiviral nucleoside analogs. *Biochem. Pharmacol.*, **68**, 1749–1756.
13. Krishnan,P., Gullen,E.A., Lam,W., Dutschman,G.E., Grill,S.P. and Cheng,Y.C. (2003) Novel role of 3-phosphoglycerate kinase, a glycolytic enzyme, in the activation of L-nucleoside analogs, a new class of anticancer and antiviral agents. *J. Biol. Chem.*, **278**, 36726–36732.
14. Banks,R.D., Blake,C.C.F., Evans,P.R., Haser,R., Rice,D.W., Hardy,G.W., Merrett,M. and Phillips,A.W. (1979) Sequence, structure and activity of phosphoglycerate kinase: a possible hinge-bending enzyme. *Nature*, **279**, 773–777.
15. Blake,C.C. and Evans,P.R. (1974) Structure of horse muscle phosphoglycerate kinase. Some results on the chain conformation, substrate binding and evolution of the molecule from a 3 angstrom Fourier map. *J. Mol. Biol.*, **84**, 585–601.
16. Watson,H.C., Walker,N.P., Shaw,P.J., Bryant,T.N., Wendell,P.L., Fothergill,L.A., Perkins,R.E., Conroy,S.C., Dobson,M.J., Tuite,M.F. *et al.* (1982) Sequence and structure of yeast phosphoglycerate kinase. *EMBO J.*, **1**, 1635–1640.
17. Davies,G.J., Gamblin,S.J., Littlechild,J.A., Dauter,Z., Wilson,K.S. and Watson,H.C. (1994) Structure of the ADP complex of the 3-phosphoglycerate kinase from Bacillus stearothermophilus at 1.65 A. *Acta Crystallogr.*, **50**, 202–209.
18. Auerbach,G., Huber,R., Grattinger,M., Zaiss,K., Schurig,H., Jaenicke,R. and Jacob,U. (1997) Closed structure of phosphoglycerate kinase from *Thermotoga maritima* reveals the catalytic mechanism and determinants of thermal stability. *Structure*, **5**, 1475–1483.
19. Bernstein,B.E., Michels,P.A.M. and Hol,W.G.J. (1997) Synergistic effects of substrate-induced conformational changes in phosphoglycerate kinase activation. *Nature*, **385**, 275–278.
20. Flachner,B., Kovári,Z., Varga,A., Gugolya,Z., Vonderviszt,F., Naray-Szabó,G. and Vas,M. (2004) Role of phosphate chain mobility of MgATP in completing the 3-phosphoglycerate kinase catalytic site: binding, kinetic, and crystallographic studies with ATP and MgATP. *Biochemistry*, **43**, 3436–3449.
21. Harlos,K., Vas,M. and Blake,C.C.F. (1992) Crystal structure of the binary complex of pig muscle phosphoglycerate kinase and its substrate 3-phospho-D-glycerate. *Proteins*, **12**, 133–144.
22. Kovári,Z., Flachner,B., Naray-Szabó,G. and Vas,M. (2002) Crystallographic and thiol-reactivity studies on the complex of pig muscle phosphoglycerate kinase with ATP analogues: correlation between nucleotide binding mode and helix flexibility. *Biochemistry*, **41**, 8796–8806.
23. Kovári,Z. and Vas,M. (2004) Protein conformer selection by sequence-dependent packing contacts in crystals of 3-phosphoglycerate kinase. *Proteins*, **55**, 198–209.
24. May,A., Vas,M., Harlos,K. and Blake,C. (1996) 2.0 Å resolution structure of a ternary complex of pig muscle phosphoglycerate kinase containing 3-phospho-D-glycerate and the nucleotide Mn adenylylimidodiphosphate. *Proteins*, **24**, 292–303.
25. Szilágyi,A.N., Ghosh,M., Garman,E. and Vas,M. (2001) A 1.8 Å resolution structure of pig muscle 3-phosphoglycerate kinase with bound MgADP and 3-phosphoglycerate in open conformation: new insight into the role of the nucleotide in domain closure. *J. Mol. Biol.*, **306**, 499–511.
26. Focher,F., Spadari,S. and Maga,G. (2003) Antivirals at the mirror: the lack of stereospecificity of some viral and human enzymes offers novel opportunities in antiviral drug development. *Curr. Drug Targets Infect. Disord.*, **3**, 41–53.
27. Szabó,J., Varga,A., Flachner,B., Konarev,P., Svergun,D., Závodszky,P. and Vas,M. (2008) Communication between the nucleotide site and the main molecular hinge of 3-Phosphoglycerate Kinase. *Biochem.*, in press.

28. Périgaud,C., Gosselin,G. and Imbach,J.-L. (1992) Potential antiviral agents. stereospecific synthesis of purines and pyrimidines substituted with chiral acyclic chains by sugar-ring opening of α-L-arabinopyranosyl nucleosides. *Chem. Soc. Perkin Trans*, **1**, 1943–1952.
29. CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 760–763.
30. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
31. Lamzin,V.S. and Wilson,K.S. (1993) Automated refinement of protein models. *Acta Crystallogr. D Biol. Crystallogr.*, **49**, 129–147.
32. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
33. Gondeau,C., Chaloin,L., Varga,A., Roy,B., Lallemand,P., Perigaud,C., Barman,T., Vas,M. and Lionne,C. (2008) Differences in the Transient Kinetics of the Binding of D-ADP and its Mirror Image L-ADP to Human 3-Phosphoglycerate Kinase Revealed by the Presence of 3-Phosphoglycerate. *Biochemistry*, **47**, 3462–3473.
34. Flachner,B., Varga,A., Szabo,J., Barna,L., Hajdu,I., Gyimesi,G., Zavodszky,P. and Vas,M. (2005) Substrate-assisted movement of the catalytic Lys 215 during domain closure: site-directed mutagenesis studies of human 3-phosphoglycerate kinase. *Biochemistry*, **44**, 16853–16865.
35. Yao,G.Q., Liu,S.H., Chou,E., Kukhanova,M., Chu,C.K. and Cheng,Y.C. (1996) Inhibition of Epstein-Barr virus replication by a novel L-nucleoside, 2′-fluoro-5-methyl-beta-L-arabinofuranosyluracil. *Biochem. Pharmacol.*, **51**, 941–947.
36. Ostermann,N., Schlichting,I., Brundiers,R., Konrad,M., Reinstein,J., Veit,T., Goody,R.S. and Lavie,A. (2000) Insights into the phosphoryltransfer mechanism of human thymidylate kinase gained from crystal structures of enzyme complexes along the reaction coordinate. *Structure*, **8**, 629–642.
37. Michelson,A.M., Blake,C.C., Evans,S.T. and Orkin,S.H. (1985) Structure of the human phosphoglycerate kinase gene and the intron-mediated evolution and dispersal of the nucleotide-binding domain. *Proc. Natl Acad. Sci. USA*, **82**, 6965–6969.
38. Al-Karadaghi,S., Cedergren-Zeppezauer,E.S. and Hovmoller,S. (1994) Refined crystal structure of liver alcohol dehydrogenase-NADH complex at 1.8 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.*, **50**, 793–807.