

STRUCTURE NOTE

Crystal structure of human phosphomevalonate kinase at 1.8 Å resolution

Qing Chang,^{1,2†} Xiao-Xue Yan,^{1†} Shen-Yan Gu,¹ Jun-Feng Liu,¹ and Dong-Cai Liang^{1*}

¹National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

²Graduate University of the Chinese Academy of Sciences, Beijing 100049, People's Republic of China

Key words: *isoprenoid biosynthesis pathway; nucleoside monophosphate kinase (NMP) family; X-ray crystallography.*

INTRODUCTION

The mevalonate pathway of isoprenoid biosynthesis is the most chemically diverse pathway in nature. This pathway produces isopentenyl diphosphate,¹ the fundamental five-carbon building block for the biosynthesis of over 23,000 isoprenoid compounds,² such as sterols,³ carotenoids,⁴ dolichols,⁵ ubiquinone,⁶ and some prominent classes of prenylated proteins.⁷ These compounds are essential for cell growth and differentiation, gene expression, protein glycosylation, and cytoskeletal assembly. In addition, isoprenoid compounds are essential for posttranslational modification of proteins involved in intracellular signaling.^{8,9}

Phosphomevalonate kinase (PMK, EC2.7.4.2) catalyzes the rate-limiting step for biosynthesis of isopentenyl diphosphate from mevalonate. Specifically, PMK catalyzes the transfer of the γ -phosphoryl group of ATP to (R)-5-phosphomevalonate (Pmev), which results in the formation of ADP and (R)-5-diphosphomevalonate (DPM).

There are two nonorthologous types of PMK. One orthologue of the *Saccharomyces cerevisiae* ERG8 gene is present in eubacteria, fungi, and plants, whereas orthologues of human PMK (hPMK) gene are found only in animals and low-homology invertebrates.¹⁰ Examination of the PMK crystal structure in *Streptococcus pneumoniae* confirmed that it belongs to the GHMP kinase (galactokinase/homoserine kinase/mevalonate kinase/phosphomevalonate) superfamily.¹¹ However, no empirical three-dimensional structural information is available for the divergent animal PMK proteins.

Animal PMKs have been purified from a variety of tissues, and several animal PMK genes have been iden-

tified. The hPMK gene was first cloned in 1996 when Chambliss *et al.* screened the human liver library with the porcine PMK clone.¹² They first expressed a GST-human PMK fusion construct in bacteria and did some preliminary characterization. Recently, Herdendorf *et al.* continued functional studies with the recombinant His-tag hPMK, alongside several mutant forms. Results from these kinetic and biophysical studies suggested that the hPMK belongs to the nucleoside monophosphate kinase (NMP) family, and possibly contains a "Walker A" motif.^{13,14}

The precise structure of hPMK is necessary to facilitate a more detailed understanding of diphosphomevalonate synthesis in humans. Here we report high-resolution crystal structure of hPMK at 1.76 Å. The three-dimensional structure of hPMK reveals that the enzyme is classified as a member of the NMP superfamily. Ours is the first report of hPMK crystal structure that suggests a potential substrate binding site and a possible enzyme catalytic mechanism.

Grant sponsor: National Basic Research Program; Grant numbers: 973, 2006CB806501 and 863, 2006AA02A316; Grant sponsor: National Protein Project of China; Grant numbers: 2006CB910902 and 2007CB914302.

[†]Qing Chang and Xiao-Xue Yan contributed equally to this work.

*Correspondence to: Dong-Cai Liang, National Laboratory of Biomacromolecules, Institute of Biophysics, CAS, 15 Datun Road, Chaoyang District, Beijing 100101, People's Republic of China. E-mail: dcliang@sun5.ibp.ac.cn

Received 27 March 2008; Accepted 19 May 2008

Published online 10 July 2008 in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/prot.22151

METHODS

Protein expression and purification

The human phosphomevalonate kinase gene (GeneID:10654) was amplified from the human liver genomic cDNA template by PCR. The amplified PCR product was annealed into the bacterial expression vector pET-22b(+) (Novagen) that expresses the cloned gene fused with a C-terminal hexa-histidine tag. The resultant recombinant protein was over-expressed in *E. coli* Rosetta(DE3) in Luria Broth containing 100 mg/L ampicillin and 34 mg/L chloramphenicol. Cells were incubated at 37°C until a density of 0.6 (A600), and then induced with 0.5 mM isopropyl β -D-thiogalactoside at 28°C overnight. Afterward, cells were harvested by centrifugation, resuspended in buffer [500 mM $(\text{NH}_4)_2\text{SO}_4$, 50 mM HEPES pH 7.6], and disrupted by sonication. The extract was centrifuged at 4°C for 30 min at 10,000g. Proteins were separated from the extract by passage through a nickel-affinity column, and purified with size-exclusion chromatography on a Superdex200 column (GE Healthcare). The protein was concentrated to 10 mg/mL, and was kept in 100 mM $(\text{NH}_4)_2\text{SO}_4$, 20 mM Hepes pH 7.6, and stored at -80°C .

Crystallization and data collection

The crystals of hPMK were grown at 15°C using the sitting-drop method. Drops consisted of 2 μL of protein solution and 2 μL of reservoir solution [15% (v/v) Pentaerythritol ethoxylate (15/4 EO/OH), 100 mM Hepes pH 7.6, 15% (v/v) MPD]. Optimal brick-like crystals were obtained in 2–3 days. Mercury derivatives were obtained by soaking the crystals in the mother liquor supplemented with 10 mM Ethylmercurithiosalicylate (EMTS) for 6 h. Native and Hg-derivative crystals in the mother liquor were flash-cooled directly in liquid nitrogen. Data were collected with a Rigaku R-AXIS IV++ image plate detector, and were integrated and scaled with DENZO and SCALEPACK.¹⁵ The statistics of data collection are summarized in Table I. The Matthews coefficient (VM) of $2.0 \text{ \AA}^3 \text{ Da}^{-1}$ for both crystals suggests the presence of a monomer in an asymmetric unit with an estimated solvent content of 40% (v/v).¹⁶

Structure determination and refinement

hPMK structure was determined by SAD method using the SHARP program.¹⁷ Three heavy atom sites were independently located by the heavy atom search routine, yielding an overall figure of merit (F.O.M.) of 0.50 and 0.84 after DM.¹⁸ Automatic model building was performed with ARP/wARP.¹⁹ Within the resolution range 8 – 1.76 \AA , a combination of COOT²⁰ and CNS²¹ was adopted for careful revision of the initial model obtained by automatic modeling. The native structure was refined

Table I
Data Collection and Refinement Statistics for hPMK Structures

	Native	Hg-derivative
Data collection statistics		
Wavelength (\AA)	1.5418	1.5418
Resolution region (\AA)	15–1.76	15–2.1
The highest resolution shells (\AA)	1.80–1.76	2.17–2.1
Space group	C2	C2
Unit cell parameters	$a = 93.14$, $b = 32.62$, $c = 62.78 \text{ \AA}$, $\beta = 111.01^\circ$	$A = 93.43$, $b = 33.18$, $c = 61.87 \text{ \AA}$, $\beta = 107.33^\circ$
No. unique reflections	17,386	10,777
Redundancy	9.8 (9.0)	13.3 (10.2)
Completeness (%) ^a	99.4 (98.2)	99.5 (95.5)
R_{merge} (%) ^{a,b}	5.5 (44.2)	8.3 (19.4)
Mean $I/\sigma(I)$	16.9 (5.4)	13.5 (12.6)
Refinement statistics		
Resolution region (\AA)	8–1.76	
Total no. of reflections used	16,987	
Reflections in working set	15,289	
Reflections in test set	1698	
Nonhydrogen atoms		
Protein	1516	
Water	149	
SO_4^{2-}	2	
MPD	2	
R_{work} (%) ^c	19.9	
R_{free} (%) ^d	22.1	
RMS deviations from ideal geometry		
Bond length (\AA)	0.005	
Bond angle ($^\circ$)	1.523	
Overall average B factor (\AA^2)	19.08	
Main chain B factor (\AA^2)	17.25	
Side chain B factor (\AA^2)	20.88	
Water B factor (\AA^2)	28.92	
Ramachandran plot		
Most favorable (%)	90.1	
Additional allowed (%)	9.9	

^aNumbers in parentheses correspond to the highest-resolution shell.

^b $R_{\text{merge}} = \sum_i |I_i - \langle I \rangle| / \sum_i I_i$, where I_i is an individual intensity measurement and $\langle I \rangle$ is the average intensity for all measurements of the reflection i .

^c $R_{\text{work}} = \sum_i ||F_o| - |F_c|| / \sum_i |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively.

^d R_{free} was calculated as R_{work} using 10% of the randomly selected unique reflections that were omitted from structure refinement.

to a final $R_{\text{work}} = 19.9\%$ and $R_{\text{free}} = 22.1\%$ and confirmed to have good stereochemistry from a Ramachandran plot calculated by PROCHECK.²² Detailed crystallography and statistical parameters of the final model are shown in Table I. The atomic coordinates and structure factors were deposited in the RCSB Protein data bank with accession number 3CH4.

RESULTS AND DISCUSSION

The crystal structure of hPMK was determined to 1.76 \AA resolution using the SAD method (Table I). The hPMK molecule exhibits a compact α/β structure with dimen-

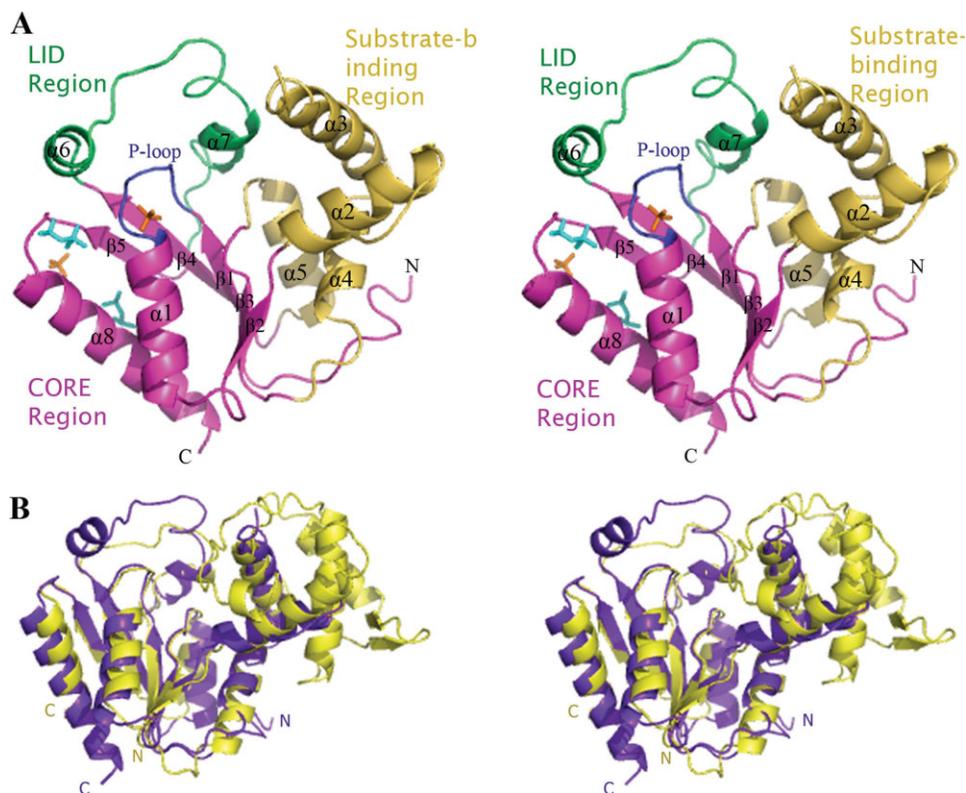


Figure 1

A: Stereo ribbon representation of hPMK. The different regions of the enzyme are color-coded. CORE region is shown in magenta, LID region in green, and substrate binding region in light brown. MPD (cyan) and sulfate ions (orange) are shown in stick models. The P-loop is shown in dark blue. **B:** Stereo ribbon representation of the superposition of DNK (deoxynucleoside monophosphate kinase from Bacteriophage T4, PDB: 1dek) with hPMK. DNK is shown in yellow and PMK in purple.

sions of about $58 \text{ E} \times 45 \text{ \AA} \times 30 \text{ \AA}$ [Fig. 1(A)]. The asymmetric unit includes a monomer with 202 residues. The final model includes 188 amino acid residues, 2 MPD molecules, 2 sulfate ions, and 149 water molecules. Residues 61–67 are absent in the model because of their poor electron density. As shown in Figure 1(A), hPMK is comprised of a central five-stranded parallel β -sheet, with strand order 23,145 and eight α -helices ($\alpha 1$ – $\alpha 8$). Helices $\alpha 1$ and $\alpha 8$ are packed against the β -sheet and $\alpha 2$, $\alpha 3$, $\alpha 4$, $\alpha 5$ are on the opposite side. Helices $\alpha 6$ and $\alpha 7$ are situated on the top of the C-terminal end of the parallel β strands. The overall topology structure of human PMK is similar to NMP kinase fold family. The protein structure can be divided into three parts, the CORE region (residues 1–42, 101–108, 122–132, 166–192), the LID region (residues 133–165), and the acceptor substrate binding region (residues 43–100, 109–121) [Fig. 1(A)].

A search for the 3D homologs using DALI²³ shows that hPMK has structural similarity with various kinases in the NMP family, although the overall amino acid sequence identity is less than 30%. The best match was deoxynucleoside monophosphate kinase (DNK) from

*Bacteriophage T4*²⁴ with a Z-score of 13.4. This structure could be superimposed on hPMK with an rms deviation of 2.9 \AA for 152 C α atoms representing the CORE region [Fig. 1(B)]. However, the LID region and substrate binding region are remarkably different. In hPMK, the LID region consists of residues 133–165, forming helix $\alpha 6$ followed by a loop that connects to helix $\alpha 7$. By contrast, there is only a short 15 residues loop in DNK that cannot completely cover the active center. Moreover, the substrate-binding region in DNK, which is much larger than in hPMK, contains five α -helices folding in a myoglobin-like manner, and two β -sheets forming a β -hairpin. These differences likely support the different binding behaviors of the acceptor substrates.

hPMK contains the Walker A motif, a short conserved stretch of sequence GXXXXGK[T/S],²⁵ which is positioned between the $\beta 1$ and the $\alpha 1$, and forms the characteristic P-loop [Fig. 1(A)]. The P-loop is the key element in binding the triphosphate in many nucleotide-binding proteins. It forms a large anion hole, with a positive electrostatic potential provided by the amino groups of the loop and the dipole of helix $\alpha 1$.

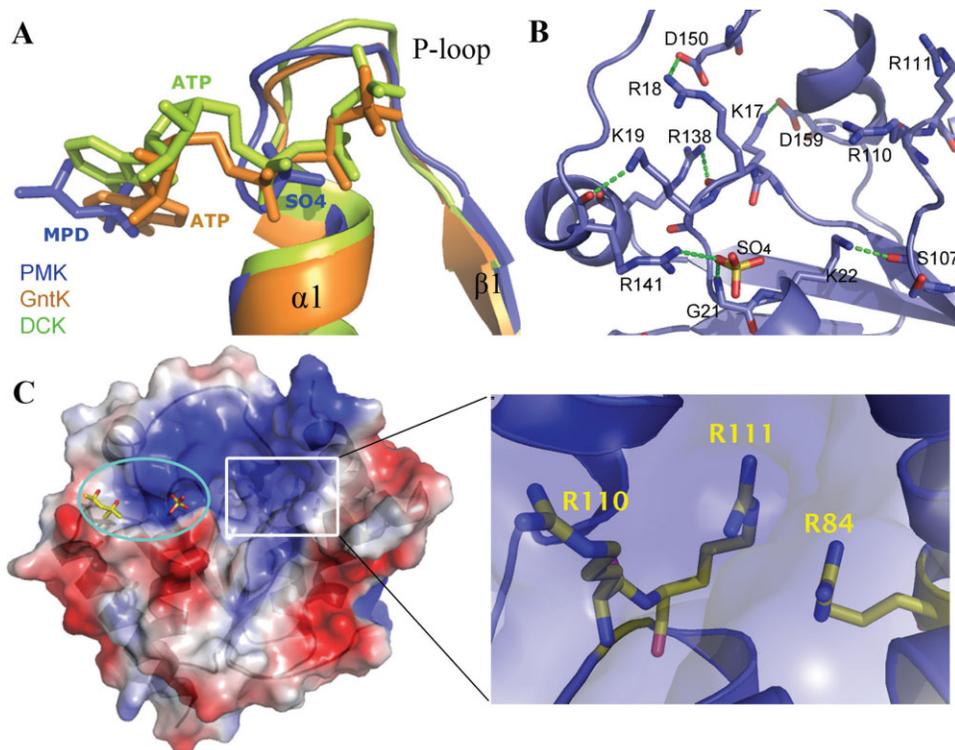


Figure 2

A: A superimposed view of the ATP binding site of sulfate-bound PMK (purple), ATP-bound GntK (orange), and ATP-bound DCK (green). The superimposition is based on C α atoms of the P-loop, helix α 1 and strand β 1. ATP, MPD and sulfate ion are shown in stick models. **B:** The interactions with the bounded sulfate ion and the interactions between residues located in the P-loop and LID region. The sulfate ion (SO $_4$) and correlative residues are represented by stick structures, and the remainder of the structure is shown in cartoon representation. Carbon atoms are colored in purple, oxygen atoms in red, nitrogen atoms in blue and the sulfur atom is yellow. Hydrogen bond pairs are connected with green dashed lines. **C:** Electrostatic surface representation of hPMK with the transparent ribbon presentation. Red represents a negative charge, and blue represents a positive charge. MPD and sulfate ion are represented by sticks. The ATP-binding site is the region within the cyan ellipse, and the substrate-binding site is the region within the white rectangle. An enlarged view of the detailed putative substrate-binding site is shown on the right, and the conserved residues that may interact with mevalonate 5-phosphate are shown by stick structures.

hPMK three-dimensional structure shows a sulfate molecule tightly bound to the P-loop. This sulfate ion forms two hydrogen bonds: one oxygen points toward the amino group of G21 and the other oxygen is hydrogen-bonded to the side chain of R141 [Fig. 2(B)]. R141 is located at helix α 6 of the LID region and is known to primarily effect nucleotide affinity.¹⁴

The superimposition of the hPMK structure, the GntK-ATP complex (gluconate kinase from *E. coli*, PDB: 1ko5),²⁶ and the DCK-ATP complex (dephospho-coenzyme A kinase from *Haemophilus influenzae*, PDB: 1jjv)²⁷ reveals that the sulfate ion most likely occupies the position corresponding to the α -phosphate of ATP [Fig. 2(A)]. Residues K17, R18, K19 with long side chains form hydrogen bonds with D159, D150, R138, and R141 of the LID region, respectively [Fig. 2(B)]. These links result in a more compact conformation of the enzyme, and thus brings the LID region in close proximity to the ATP binding site. The purpose of such a mechanism in

substrate binding has been suggested as the prevention of premature ATP hydrolysis. K22 forms hydrogen bonds to the oxygen atom of S107 close to R110 and R111, which are important for catalysis and binding of mevalonate 5-phosphate [Fig. 2(B)].¹⁴ This interaction not only stabilizes the tight structure of the loop but also ensures adjacency of the donor and acceptor substrates. With no ligand bound to it, hPMK has an exposed MPD molecule located in the binding site of the adenine moiety of ATP. We assert that Q140 of helix α 6, N170-H171-G172-V173 in the loop between helix α 6 and α 7, and E174 of α 8 may take part in the binding of the ribose sugar of ATP.

The acceptor substrate binding region of hPMK presents an open conformation because of the absence of substrate Pmev. There is a deep cavity with a positive electrostatic potential on the side of the β -sheet opposite from the P-loop, which is a putative Pmev binding site [Fig. 2(C)]. The conserved residues R84, R110, and R111 in the cavity may form electrostatic interaction with the

5-phosphate and hydroxyl groups of the substrate. It has been reported that R84 and R111 significantly contribute to the binding of hPMK nonnucleotide substrate. Also, R110 is important to hPMK catalysis, which is also influenced by K48 and R73.¹⁴ By analyzing the enzyme structure, we found that in the acceptor substrate binding region there is a flexible loop between helix $\alpha 2$ and $\alpha 3$ missed because of the unaccountable electron density. We guessed that once the substrate binds to the cavity, helix $\alpha 2$, $\alpha 3$, and the loop between them may move toward the top of the cavity to make a compact conformation, and as a result of this change, catalytic residues may be correctly positioned to participate in catalysis, and the reactive environment may be protected. At present, the basis of recognition of both substrates remains unclear. Attempts to crystallize hPMK complexes with ATP and/or Pmev are in progress.

In summary, we have obtained the crystal structure of hPMK in the absence of substrate. We are the first to report three-dimensional structure of hPMK protein that differs from *Streptococcus pneumoniae* PMK encoded by ERG8 gene. Despite low sequence identity, the topological and structural similarities provide strong evidence that this enzyme is a member of the NMP kinase superfamily. Most of the CORE domain and specifically the central parallel β -strands of this superfamily are in a very similar orientation. The ATP binding manners of these enzymes are also quite similar. Therefore, we infer that these kinases have the identical enzyme ancestor and similar catalytic mechanism. Our definition of the crystal structure of hPMK provides the structural foundation for studying the catalytic mechanism of animal PMKs.

REFERENCES

- Cornforth RH, Popjak G. Chemical synthesis of substrates of sterol biosynthesis. *Meth Enzymol* 1969;15:359–371.
- Spurgeon SL, Porter JW. Biosynthesis of isoprenoid compounds. New York: John Wiley and Sons; 1981. pp 1–46.
- Morton RA. Natural substances formed biologically from mevalonic acid. Chairman's opening remarks. *Biochem Soc Symp* 1970;29:1–3.
- Goodwin TW. Biosynthesis of carotenoids and plant triterpenes. *Biochem J* 1971;123:293–329.
- Matsuoka S, Sagami H, Kurisaki A, Ogura K. Variable product specificity of microsomal dehydrodolichyl diphosphate synthase from rat liver. *J Biol Chem* 1991;266:3464–3468.
- Ashby MN, Edwards PA. Elucidation of the deficiency in two yeast coenzyme Q mutants. Characterization of the structural gene encoding hexaprenyl pyrophosphate synthetase. *J Biol Chem* 1990; 265:13157–13164.
- Clarke S. Protein isoprenylation and methylation at carboxyl-terminal cysteine residues. *Annu Rev Biochem* 1992;61:355–386.
- Goldstein JL, Brown MS. Regulation of the mevalonate pathway. *Nature* 1990;343:425–430.
- Hinson DD, Chambliss KL, Toth MJ, Tanaka RD, Gibson KM. Post-translational regulation of mevalonate kinase by intermediates of the cholesterol and nonsterol isoprene biosynthetic pathways. *J Lipid Res* 1997;38:2216–2223.
- Houten SM, Waterham HR. Nonorthologous gene displacement of phosphomevalonate kinase. *Mol Genet Metab* 2001;72:273–276.
- Romanowski MJ, Bonanno JB, Burley SK. Crystal structure of the *Streptococcus pneumoniae* phosphomevalonate kinase, a member of the GHMP kinase superfamily. *Proteins* 2002;47:568–571.
- Chambliss KL, Slaughter CA, Schreiner R, Hoffmann GF, Gibson KM. Molecular cloning of human phosphomevalonate kinase and identification of a consensus peroxisomal targeting sequence. *J Biol Chem* 1996;271:17330–17334.
- Herdendorf TJ, Mizioro HM. Phosphomevalonate kinase: functional investigation of the recombinant human enzyme. *Biochemistry* 2006;45:3235–3242.
- Herdendorf TJ, Mizioro HM. Functional evaluation of conserved basic residues in human phosphomevalonate kinase. *Biochemistry* 2007;46:11780–11788.
- Outwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Meth Enzymol* 1997;276:307–326.
- Matthews BW. Solvent content of protein crystals. *J Mol Biol* 1968;33:491–497.
- Vonrhein C, Blanc E, Roversi P, Bricogne G. Automated structure solution with autoSHARP. *Methods Mol Biol* 2007;364:215–230.
- Cowtan K. dm: An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM. Newsl Protein Crystallogr* 1994;31:34–38.
- Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 1999;6:458–463.
- Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004;60 (Part 12 Part 1):2126–2132.
- Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 1998;54 (Part 5):905–921.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 1993;26:283–291.
- Holm L, Sander C. Touring protein fold space with Dali/FSSP. *Nucleic Acids Res* 1998;26:316–319.
- Tepljakov A, Sebastiao P, Obmolova G, Perrakis A, Brush GS, Bessman MJ, Wilson KS. Crystal structure of bacteriophage T4 deoxynucleotide kinase with its substrates dGMP and ATP. *EMBO J* 1996;15:3487–3497.
- Walker JE, Saraste M, Runswick MJ, Gay NJ. Distantly related sequences in the α - and β -subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* 1982;1:945–951.
- Kraft L, Sprenger GA, Lindqvist Y. Conformational changes during the catalytic cycle of gluconate kinase as revealed by X-ray crystallography. *J Mol Biol* 2002;318:1057–1069.
- Obmolova G, Tepljakov A, Bonander N, Eisenstein E, Howard AJ, Gilliland GL. Crystal structure of dephospho-coenzyme A kinase from *Haemophilus influenzae*. *J Struct Biol* 2001;136:119–125.