wwPDB X-ray Structure Validation Summary Report (i)

Aug 22, 2023 – 02:17 AM EDT

PDB ID	:	2P9Q
Title	:	Crystal Structure of Phosphoglycerate Kinase-2
Authors	:	Sawyer, G.M.; Monzingo, A.F.; Poteet, E.C.; Robertus, J.D.
Deposited on	:	2007-03-26
Resolution	:	2.70 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

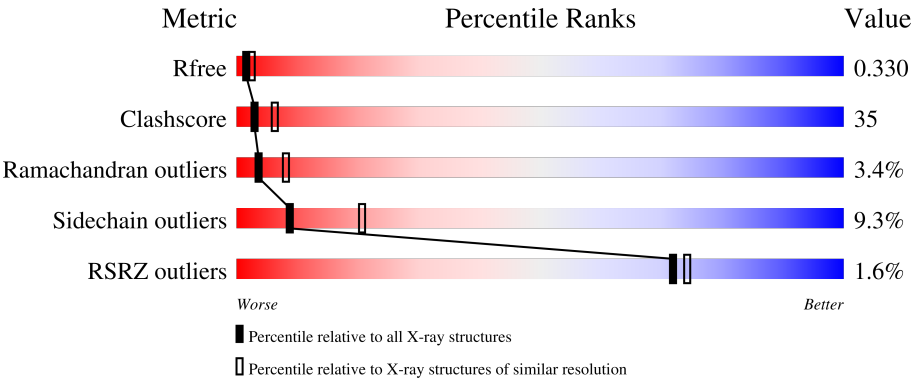
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	416	2% 45% 44% 8% .
1	B	416	% 47% 43% 7% .

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6125 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Phosphoglycerate kinase, testis specific.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	404	Total	C	N	O	S	0	0	0
			3043	1933	514	577	19			
1	B	403	Total	C	N	O	S	0	0	0
			3004	1907	506	572	19			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	150	ARG	GLN	SEE REMARK 999	UNP P09041
B	150	ARG	GLN	SEE REMARK 999	UNP P09041

- Molecule 2 is water.

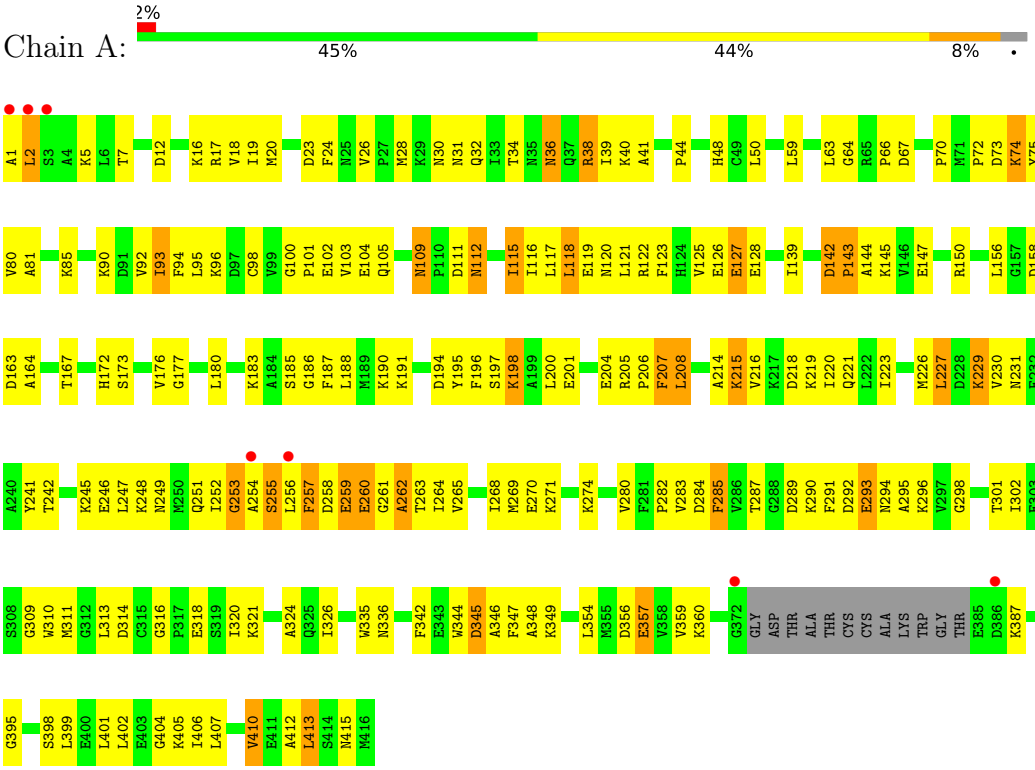
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	48	Total	O	0	0
			48	48		
2	B	30	Total	O	0	0
			30	30		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

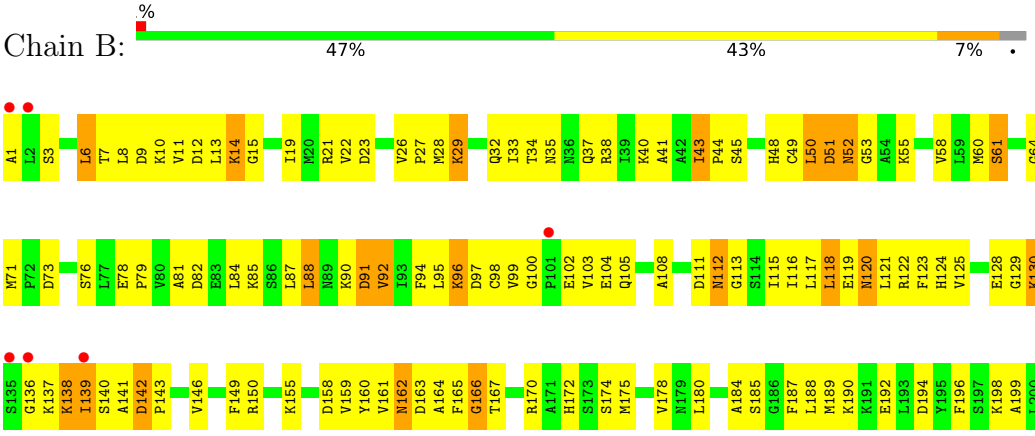
• Molecule 1: Phosphoglycerate kinase, testis specific

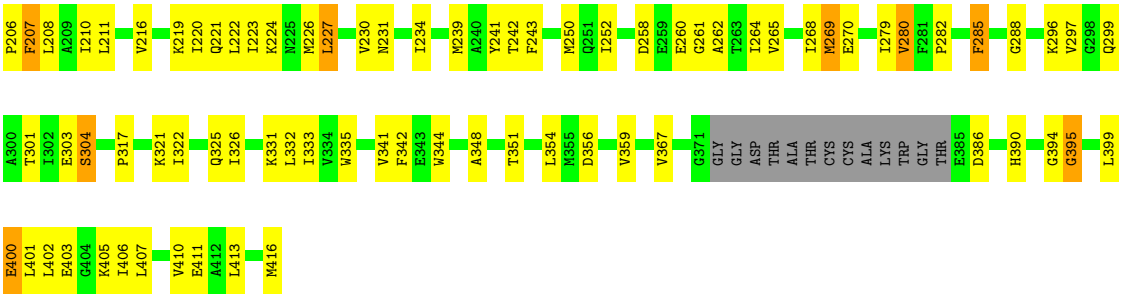
Chain A:



• Molecule 1: Phosphoglycerate kinase, testis specific

Chain B:




P206 F207 L208 A209 I210 L211 V216 K219 I220 Q221 L222 I223 K224 N225 M226 L227 V230 N231 I234 M239 A240 Y241 T242 F243 M250 Q251 I252 D258 E259 E260 G261 A262 T263 I264 V265 I268 M269 E270 I279 V280 F281 P282 F285 G288 K296 V297 G298 Q299
A300 T301 I302 E303 S304 P317 K321 I322 Q325 I326 K331 L332 I333 V334 W335 V341 F342 E343 W344 A348 T351 L354 M355 D356 V359 V367 G371 GLY GLY ASP THR ALA THR CYS CYS ALA LYS TRP GLY THR E385 D386 H390 G394 G395 L399
E400 L401 L402 E403 G404 K405 I406 L407 V410 E411 A412 L413 M416

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	52.04Å 80.85Å 92.90Å 90.00° 94.63° 90.00°	Depositor
Resolution (Å)	20.00 – 2.70 29.70 – 2.80	Depositor EDS
% Data completeness (in resolution range)	82.3 (20.00-2.70) 93.7 (29.70-2.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.98 (at 2.80Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.217 , 0.328 0.222 , 0.330	Depositor DCC
R_{free} test set	887 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	35.7	Xtriage
Anisotropy	0.217	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 44.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	6125	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 38.79 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 3.4985e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.43	0/3090	0.65	0/4158
1	B	0.41	0/3051	0.67	0/4115
All	All	0.42	0/6141	0.66	0/8273

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3043	0	3118	195	0
1	B	3004	0	3029	229	0
2	A	48	0	0	3	0
2	B	30	0	0	3	0
All	All	6125	0	6147	421	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 35.

The worst 5 of 421 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:215:LYS:H	1:A:215:LYS:HD2	1.05	1.08

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:6:LEU:HD13	1:B:184:ALA:HB2	1.46	0.97
1:B:139:ILE:HG22	1:B:140:SER:H	1.28	0.96
1:A:207:PHE:H	1:A:231:ASN:HD22	1.06	0.95
1:B:88:LEU:HD13	1:B:90:LYS:HB3	1.54	0.89

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	400/416 (96%)	337 (84%)	49 (12%)	14 (4%)	3	8
1	B	399/416 (96%)	344 (86%)	42 (10%)	13 (3%)	4	8
All	All	799/832 (96%)	681 (85%)	91 (11%)	27 (3%)	3	8

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	262	ALA
1	A	345	ASP
1	B	14	LYS
1	B	73	ASP
1	B	112	ASN

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	329/340 (97%)	296 (90%)	33 (10%)	7 18
1	B	319/340 (94%)	292 (92%)	27 (8%)	10 24
All	All	648/680 (95%)	588 (91%)	60 (9%)	9 21

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	357	GLU
1	B	297	VAL
1	B	29	LYS
1	B	285	PHE
1	B	402	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	30	ASN
1	B	120	ASN
1	B	109	ASN
1	B	162	ASN
1	A	179	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	404/416 (97%)	-0.08	7 (1%)	70	72	4, 24, 38, 48	0
1	B	403/416 (96%)	-0.07	6 (1%)	73	76	11, 27, 39, 45	0
All	All	807/832 (96%)	-0.07	13 (1%)	72	74	4, 25, 38, 48	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	SER	4.2
1	B	1	ALA	4.1
1	A	1	ALA	3.4
1	B	135	SER	3.2
1	A	372	GLY	3.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.