WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 23, 2023 – 12:24 PM EDT

PDB ID	:	3D8H
Title	:	Crystal structure of phosphoglycerate mutase from Cryptosporidium parvum, cgd7_4270
Authors	:	Wernimont, A.K.; Lew, J.; Wasney, G.; Alam, Z.; Kozieradzki, I.; Cossar, D.; Schapiro, M.; Bochkarev, A.; Arrowsmith, C.H.; Bountra, C.; Wilkstrom, M.; Edwards, A.M.; Hui, R.; Artz, J.D.; Hills, T.; Structural Genomics Consortium (SGC)
Deposited on	:	2008-05-23
Resolution	:	2.01 Å(reported)

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

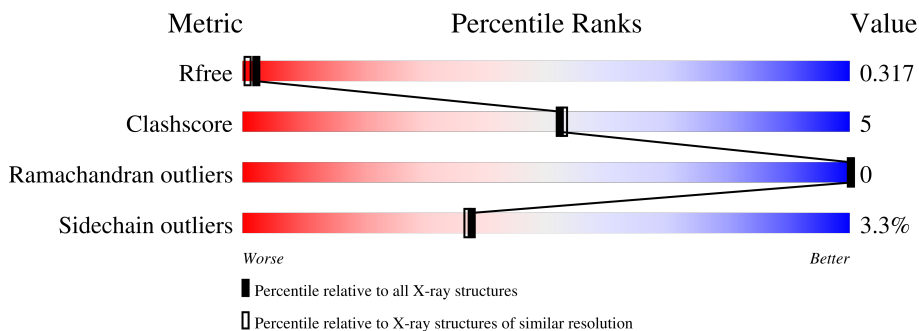|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.01 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	Percentile Ranks	Value
Rfree		0.317
Clashscore		5
Ramachandran outliers		0
Sidechain outliers		3.3%

Worse ◼ *Better*

◼ Percentile relative to all X-ray structures

◻ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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%

Mol	Chain	Length	Quality of chain
1	A	267	72% 12% • 14%
1	B	267	73% 15% 12%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 3871 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 Glycolytic phosphoglycerate mutase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	229	Total 1811	C 1167	N 296	O 339	S 9	0	1	0
1	B	235	Total 1854	C 1193	N 304	O 348	S 9	0	3	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP Q5CXZ9
A	2	GLY	-	expression tag	UNP Q5CXZ9
A	3	SER	-	expression tag	UNP Q5CXZ9
A	4	SER	-	expression tag	UNP Q5CXZ9
A	5	HIS	-	expression tag	UNP Q5CXZ9
A	6	HIS	-	expression tag	UNP Q5CXZ9
A	7	HIS	-	expression tag	UNP Q5CXZ9
A	8	HIS	-	expression tag	UNP Q5CXZ9
A	9	HIS	-	expression tag	UNP Q5CXZ9
A	10	HIS	-	expression tag	UNP Q5CXZ9
A	11	SER	-	expression tag	UNP Q5CXZ9
A	12	SER	-	expression tag	UNP Q5CXZ9
A	13	GLY	-	expression tag	UNP Q5CXZ9
A	14	LEU	-	expression tag	UNP Q5CXZ9
A	15	VAL	-	expression tag	UNP Q5CXZ9
A	16	PRO	-	expression tag	UNP Q5CXZ9
A	17	ARG	-	expression tag	UNP Q5CXZ9
A	18	GLY	-	expression tag	UNP Q5CXZ9
A	19	SER	-	expression tag	UNP Q5CXZ9
B	1	MET	-	expression tag	UNP Q5CXZ9
B	2	GLY	-	expression tag	UNP Q5CXZ9
B	3	SER	-	expression tag	UNP Q5CXZ9
B	4	SER	-	expression tag	UNP Q5CXZ9
B	5	HIS	-	expression tag	UNP Q5CXZ9
B	6	HIS	-	expression tag	UNP Q5CXZ9

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	7	HIS	-	expression tag	UNP Q5CXZ9
B	8	HIS	-	expression tag	UNP Q5CXZ9
B	9	HIS	-	expression tag	UNP Q5CXZ9
B	10	HIS	-	expression tag	UNP Q5CXZ9
B	11	SER	-	expression tag	UNP Q5CXZ9
B	12	SER	-	expression tag	UNP Q5CXZ9
B	13	GLY	-	expression tag	UNP Q5CXZ9
B	14	LEU	-	expression tag	UNP Q5CXZ9
B	15	VAL	-	expression tag	UNP Q5CXZ9
B	16	PRO	-	expression tag	UNP Q5CXZ9
B	17	ARG	-	expression tag	UNP Q5CXZ9
B	18	GLY	-	expression tag	UNP Q5CXZ9
B	19	SER	-	expression tag	UNP Q5CXZ9

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	92	Total 92	O 92	0	0
2	B	114	Total 114	O 114	0	0

3 Residue-property plots ⓘ

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. 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. 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.
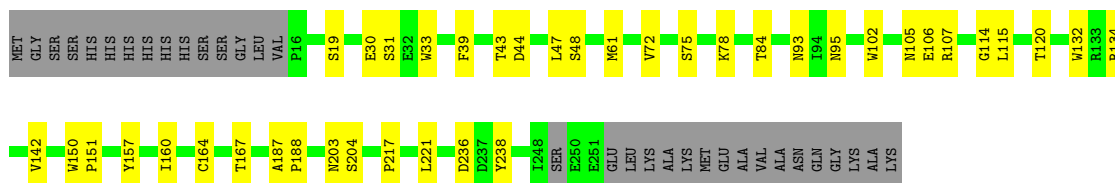
• Molecule 1: Glycolytic phosphoglycerate mutase

Chain A:
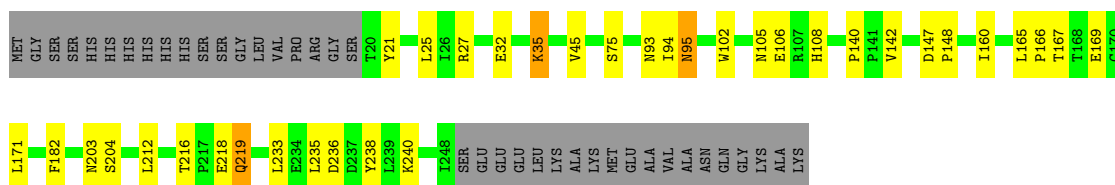
• Molecule 1: Glycolytic phosphoglycerate mutase

Chain B:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	52.25Å 158.37Å 140.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.50 – 2.01 40.50 – 2.01	Depositor EDS
% Data completeness (in resolution range)	99.7 (40.50-2.01) 99.7 (40.50-2.01)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	2.23 (at 2.01Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.231 , 0.256 0.301 , 0.317	Depositor DCC
R_{free} test set	1958 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	38.4	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 40.2	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	3871	wwPDB-VP
Average B, all atoms (Å^2)	36.0	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 36.29 % 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 5.1092e-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	0.55	1/1857 (0.1%)	0.57	0/2526
1	B	0.45	0/1906	0.58	0/2590
All	All	0.50	1/3763 (0.0%)	0.58	0/5116

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	219	GLN	CD-OE1	10.65	1.47	1.24

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	1811	0	1798	19	0
1	B	1854	0	1826	21	0
2	A	92	0	0	1	0
2	B	114	0	0	0	0
All	All	3871	0	3624	40	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 5.

The worst 5 of 40 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:95:ASN:H	1:A:95:ASN:HD22	1.46	0.64
1:B:236:ASP:HB3	1:B:238:TYR:H	1.61	0.64
1:A:95:ASN:H	1:A:95:ASN:ND2	1.97	0.62
1:A:142:VAL:HG13	1:A:167:THR:HB	1.85	0.58
1:B:115:LEU:HB3	1:B:120:THR:HG23	1.86	0.58

There are no symmetry-related clashes.

5.3   Torsion angles ⓘ

5.3.1   Protein backbone ⓘ

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	228/267 (85%)	225 (99%)	3 (1%)	0	100	100
1	B	234/267 (88%)	227 (97%)	7 (3%)	0	100	100
All	All	462/534 (86%)	452 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2   Protein sidechains ⓘ

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	198/236 (84%)	189 (96%)	9 (4%)	27	24
1	B	200/236 (85%)	196 (98%)	4 (2%)	55	58
All	All	398/472 (84%)	385 (97%)	13 (3%)	38	37

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	204	SER
1	A	218	GLU
1	B	95	ASN
1	B	61	MET
1	B	93	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	95	ASN
1	A	116	ASN
1	A	128	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.