Full wwPDB X-ray Structure Validation Report (i)

May 22, 2020 – 01:35 am BST

PDB ID	:	2PGD
Title	:	THE STRUCTURE OF 6-PHOSPHOGLUCONATE DEHYDROGENASE REFINED AT 2 ANGSTROMS RESOLUTION
Authors	:	Adams, M.J.; Phillips, C.; Gover, S.
Deposited on	:	1994-07-18
Resolution	:	2.00 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

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 2.00 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.
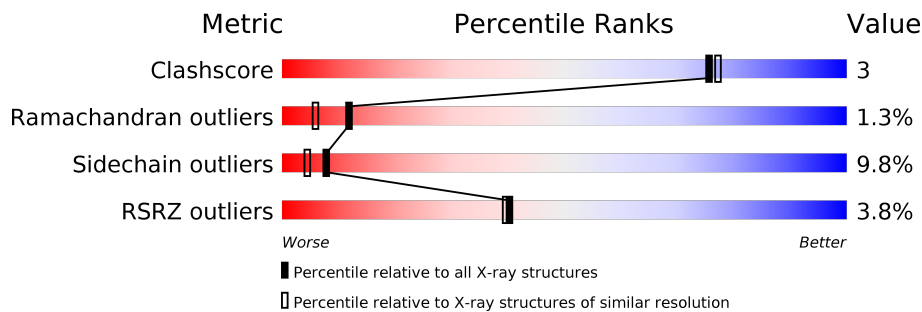
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 on 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	482	4% 81% 15%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4015 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 6-PHOSPHOGLUCONATE DEHYDROGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	473	Total 3654	C 2332	N 634	O 666	S 22	0	0	0

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O$_4$S).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0
2	A	1	Total 5	O 4	S 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	346	Total 346	O 346	0	0

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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: 6-PHOSPHOGLUCONATE DEHYDROGENASE
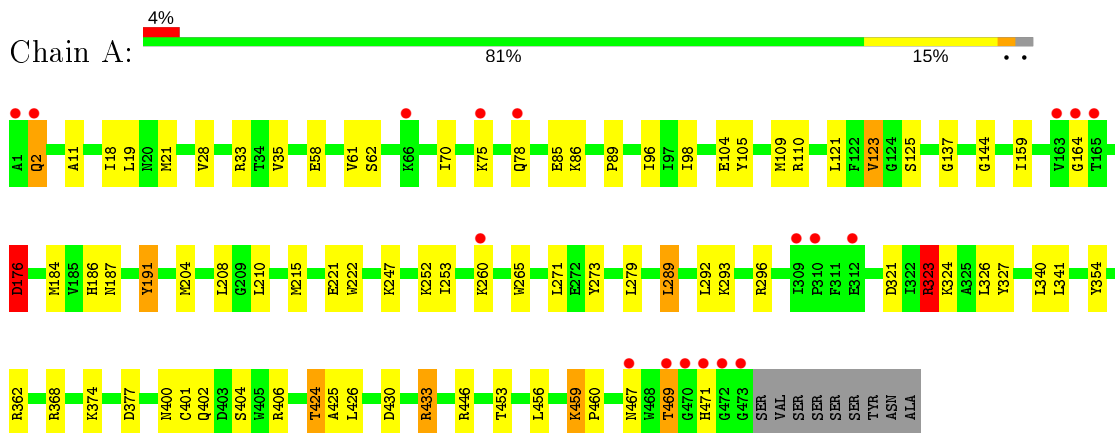
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	72.74Å 148.40Å 102.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.00 19.73 – 1.99	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-2.00) 91.6 (19.73-1.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.66 (at 1.99Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.198 , (Not available) 0.199 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	28.1	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 73.2	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4015	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.27% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.53	0/3731	1.16	16/5028 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	3

There are no bond length outliers.

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	433	ARG	CA-CB-CG	7.15	129.13	113.40
1	A	433	ARG	NE-CZ-NH2	-6.56	117.02	120.30
1	A	433	ARG	NE-CZ-NH1	6.56	123.58	120.30
1	A	424	THR	CA-CB-CG2	-6.54	103.25	112.40
1	A	424	THR	N-CA-CB	6.36	122.38	110.30
1	A	323	ARG	CA-CB-CG	5.94	126.47	113.40
1	A	368	ARG	NE-CZ-NH1	5.91	123.26	120.30
1	A	424	THR	CB-CA-C	-5.72	96.14	111.60
1	A	323	ARG	NE-CZ-NH1	5.69	123.14	120.30
1	A	110	ARG	NE-CZ-NH1	5.68	123.14	120.30
1	A	406	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	A	191	TYR	CB-CG-CD2	-5.55	117.67	121.00
1	A	433	ARG	CB-CA-C	-5.51	99.38	110.40
1	A	446	ARG	NE-CZ-NH1	5.41	123.00	120.30
1	A	327	TYR	CB-CG-CD2	-5.17	117.90	121.00
1	A	453	THR	CA-CB-CG2	-5.10	105.25	112.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	191	TYR	Sidechain
1	A	273	TYR	Sidechain
1	A	354	TYR	Sidechain

5.2 Too-close contacts (i)

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	3654	0	3661	23	0
2	A	15	0	0	0	0
3	A	346	0	0	3	0
All	All	4015	0	3661	23	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 3.

All (23) 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:204:MET:HB2	1:A:215:MET:HE3	1.57	0.86
1:A:430:ASP:HA	1:A:433:ARG:HD3	1.73	0.70
1:A:186:HIS:HD2	1:A:187:ASN:OD1	1.90	0.54
1:A:104:GLU:HG3	3:A:615:HOH:O	2.09	0.51
1:A:324:LYS:HD3	1:A:401:CYS:HA	1.94	0.49
1:A:86:LYS:O	1:A:89:PRO:HD2	2.13	0.49
1:A:321:ASP:HA	1:A:404:SER:HB3	1.97	0.46
1:A:98:ILE:HG12	1:A:123:VAL:HG13	1.96	0.46
1:A:279:LEU:HD12	1:A:425:ALA:HB2	1.98	0.45
1:A:105:TYR:O	1:A:109:MET:HG3	2.16	0.45
1:A:184:MET:HE2	1:A:265:TRP:HE3	1.81	0.45
1:A:222:TRP:CE2	1:A:323:ARG:HG2	2.54	0.42
1:A:11:ALA:HA	3:A:687:HOH:O	2.19	0.42
1:A:58:GLU:O	1:A:61:VAL:HG12	2.20	0.42
1:A:289:LEU:HD22	1:A:296:ARG:HD3	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:21:MET:HG2	1:A:159:ILE:HG21	2.02	0.41
1:A:323:ARG:HG3	1:A:323:ARG:HH11	1.85	0.41
1:A:144:GLY:HA2	1:A:176:ASP:HA	2.03	0.41
1:A:204:MET:CB	1:A:215:MET:HE3	2.40	0.41
1:A:18:ILE:HG12	1:A:28:VAL:HG11	2.02	0.41
1:A:467:ASN:HB2	3:A:1063:HOH:O	2.21	0.41
1:A:459:LYS:N	1:A:460:PRO:HD3	2.35	0.41
1:A:96:ILE:HA	1:A:121:LEU:O	2.20	0.40

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	471/482 (98%)	447 (95%)	18 (4%)	6 (1%)	12 6

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	GLN
1	A	176	ASP
1	A	469	THR
1	A	164	GLY
1	A	137	GLY
1	A	459	LYS

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	378/386 (98%)	341 (90%)	37 (10%)	8	4

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	GLN
1	A	19	LEU
1	A	33	ARG
1	A	35	VAL
1	A	62	SER
1	A	70	ILE
1	A	75	LYS
1	A	78	GLN
1	A	85	GLU
1	A	123	VAL
1	A	125	SER
1	A	176	ASP
1	A	208	LEU
1	A	210	LEU
1	A	221	GLU
1	A	247	LYS
1	A	252	LYS
1	A	253	ILE
1	A	260	LYS
1	A	271	LEU
1	A	289	LEU
1	A	292	LEU
1	A	293	LYS
1	A	323	ARG
1	A	326	LEU
1	A	340	LEU
1	A	341	LEU
1	A	362	ARG
1	A	374	LYS
1	A	377	ASP
1	A	400	ASN
1	A	402	GLN
1	A	424	THR
1	A	426	LEU
1	A	456	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	469	THR
1	A	471	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	78	GLN
1	A	180	HIS
1	A	186	HIS
1	A	445	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 carbohydrates in this entry.

5.6　Ligand geometry ⓘ

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
2	SO4	A	505	-	4,4,4	0.72	0	6,6,6	0.29	0
2	SO4	A	508	-	4,4,4	0.55	0	6,6,6	0.37	0
2	SO4	A	507	-	4,4,4	0.40	0	6,6,6	0.30	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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(Å²)	Q<0.9
1	A	473/482 (98%)	0.05	18 (3%)	40	39	13, 30, 54, 89	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	473	GLY	6.9
1	A	472	GLY	6.5
1	A	469	THR	6.1
1	A	471	HIS	5.7
1	A	470	GLY	5.5
1	A	1	ALA	4.9
1	A	310	PRO	4.2
1	A	165	THR	3.7
1	A	312	GLU	3.6
1	A	78	GLN	3.4
1	A	309	ILE	2.9
1	A	66	LYS	2.6
1	A	260	LYS	2.4
1	A	75	LYS	2.3
1	A	164	GLY	2.3
1	A	2	GLN	2.2
1	A	467	ASN	2.2
1	A	163	VAL	2.0

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

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	SO4	A	508	5/5	0.87	0.30	60,60,60,61	5
2	SO4	A	507	5/5	0.93	0.22	60,60,62,62	0
2	SO4	A	505	5/5	0.99	0.07	27,30,33,33	0

6.5 Other polymers (i)

There are no such residues in this entry.