



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 2, 2017 – 06:21 AM EDT

PDB ID : 2ZYA  
Title : Dimeric 6-phosphogluconate dehydrogenase complexed with 6-phosphogluconate  
Authors : Chen, Y.-Y.; Ko, T.-P.; Lo, L.-P.; Lin, C.-H.; Wang, A.H.-J.  
Deposited on : unknown  
Resolution : 1.60 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

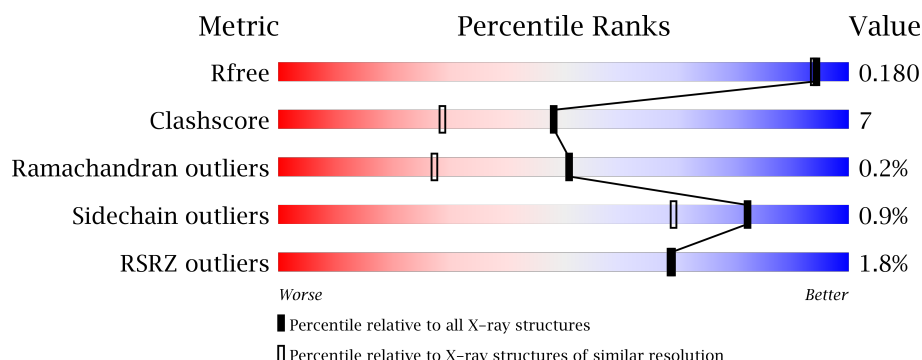
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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))
$R_{free}$	100719	2696 (1.60-1.60)
Clashscore	112137	2967 (1.60-1.60)
Ramachandran outliers	110173	2887 (1.60-1.60)
Sidechain outliers	110143	2886 (1.60-1.60)
RSRZ outliers	101464	2714 (1.60-1.60)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	480	<div> <div>2%</div> <div> <div></div> <div>87%</div> <div>9%</div> <div>..</div> </div> </div>
1	B	480	<div> <div>%</div> <div> <div></div> <div>88%</div> <div>9%</div> <div>..</div> </div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 8643 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, decarboxylating.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	465	Total	C	N	O	S	0	0	0
			3602	2290	606	697	9			
1	B	466	Total	C	N	O	S	0	0	0
			3610	2296	607	698	9			

There are 26 discrepancies between the modelled and reference sequences:

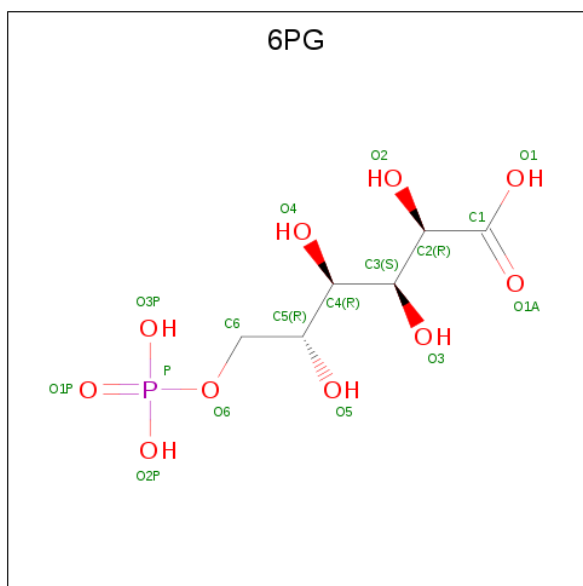
Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	MET	-	EXPRESSION TAG	UNP P00350
A	-10	GLY	-	EXPRESSION TAG	UNP P00350
A	-9	HIS	-	EXPRESSION TAG	UNP P00350
A	-8	HIS	-	EXPRESSION TAG	UNP P00350
A	-7	HIS	-	EXPRESSION TAG	UNP P00350
A	-6	HIS	-	EXPRESSION TAG	UNP P00350
A	-5	HIS	-	EXPRESSION TAG	UNP P00350
A	-4	HIS	-	EXPRESSION TAG	UNP P00350
A	-3	HIS	-	EXPRESSION TAG	UNP P00350
A	-2	HIS	-	EXPRESSION TAG	UNP P00350
A	-1	HIS	-	EXPRESSION TAG	UNP P00350
A	0	HIS	-	EXPRESSION TAG	UNP P00350
A	378	CYS	TYR	ENGINEERED	UNP P00350
B	-11	MET	-	EXPRESSION TAG	UNP P00350
B	-10	GLY	-	EXPRESSION TAG	UNP P00350
B	-9	HIS	-	EXPRESSION TAG	UNP P00350
B	-8	HIS	-	EXPRESSION TAG	UNP P00350
B	-7	HIS	-	EXPRESSION TAG	UNP P00350
B	-6	HIS	-	EXPRESSION TAG	UNP P00350
B	-5	HIS	-	EXPRESSION TAG	UNP P00350
B	-4	HIS	-	EXPRESSION TAG	UNP P00350
B	-3	HIS	-	EXPRESSION TAG	UNP P00350
B	-2	HIS	-	EXPRESSION TAG	UNP P00350
B	-1	HIS	-	EXPRESSION TAG	UNP P00350
B	0	HIS	-	EXPRESSION TAG	UNP P00350

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	378	CYS	TYR	ENGINEERED	UNP P00350

- Molecule 2 is 6-PHOSPHOGLUCONIC ACID (three-letter code: 6PG) (formula:  $C_6H_{13}O_{10}P$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	O	P	0	0
			17	6	10	1		
2	B	1	Total	C	O	P	0	0
			17	6	10	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	680	Total	O	0	0
			680	680		
3	B	717	Total	O	0	0
			717	717		



- Molecule 1: 6-phosphogluconate dehydrogenase, decarboxylating



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.00Å 108.72Å 133.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 1.60 31.98 – 1.60	Depositor EDS
% Data completeness (in resolution range)	93.6 (50.00-1.60) 93.6 (31.98-1.60)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.68 (at 1.60Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.153 , 0.181 0.153 , 0.180	Depositor DCC
$R_{free}$ test set	6192 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.7	Xtriage
Anisotropy	0.281	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 52.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	8643	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.61% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: 6PG

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	0.82	1/3670 (0.0%)	0.84	2/4969 (0.0%)
1	B	0.89	1/3678 (0.0%)	0.91	3/4980 (0.1%)
All	All	0.86	2/7348 (0.0%)	0.87	5/9949 (0.1%)

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	1
1	B	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	263	GLY	C-O	10.02	1.39	1.23
1	A	263	GLY	C-O	9.81	1.39	1.23

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	141	MET	CG-SD-CE	-8.35	86.84	100.20
1	A	261	GLY	N-CA-C	6.44	129.19	113.10
1	A	111	ARG	NE-CZ-NH2	-5.30	117.65	120.30
1	B	111	ARG	NE-CZ-NH2	-5.14	117.73	120.30
1	B	111	ARG	NE-CZ-NH1	5.12	122.86	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	427	TYR	Sidechain
1	B	353	TYR	Sidechain

## 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	3602	0	3569	57	0
1	B	3610	0	3580	48	0
2	A	17	0	10	0	0
2	B	17	0	10	0	0
3	A	680	0	0	20	1
3	B	717	0	0	12	1
All	All	8643	0	7169	100	1

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 7.

The worst 5 of 100 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:309:GLN:HG2	1:A:414:ASN:ND2	1.83	0.93
1:A:410:TYR:CZ	1:A:414:ASN:OD1	2.22	0.92
1:A:262:THR:HG22	3:A:947:HOH:O	1.70	0.90
1:A:413:GLN:HE22	1:B:402:GLN:HE22	1.16	0.90
1:A:262:THR:CG2	3:A:947:HOH:O	2.22	0.84

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:678:HOH:O	3:B:948:HOH:O[2_564]	2.18	0.02



## 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	463/480 (96%)	451 (97%)	11 (2%)	1 (0%)	51	27
1	B	464/480 (97%)	452 (97%)	11 (2%)	1 (0%)	51	27
All	All	927/960 (97%)	903 (97%)	22 (2%)	2 (0%)	51	27

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	245	ASP
1	B	129	GLY

### 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	373/387 (96%)	370 (99%)	3 (1%)	85	73
1	B	374/387 (97%)	370 (99%)	4 (1%)	78	61
All	All	747/774 (96%)	740 (99%)	7 (1%)	82	69

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	5	GLN
1	B	414	ASN
1	B	106	GLN
1	A	106	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	141	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	414	ASN
1	B	112	ASN
1	B	401	GLN
1	B	5	GLN
1	B	106	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 ⓘ

2 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	6PG	A	2002	-	13,16,16	2.01	6 (46%)	18,23,23	1.32	1 (5%)
2	6PG	B	2001	-	13,16,16	1.88	6 (46%)	18,23,23	1.27	3 (16%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	6PG	A	2002	-	-	0/18/22/22	0/0/0/0
2	6PG	B	2001	-	-	0/18/22/22	0/0/0/0

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2001	6PG	P-O3P	2.01	1.63	1.54
2	A	2002	6PG	C4-C3	2.08	1.57	1.53
2	A	2002	6PG	P-O3P	2.23	1.64	1.54
2	B	2001	6PG	P-O2P	2.25	1.64	1.54
2	B	2001	6PG	O4-C4	2.30	1.48	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2002	6PG	P-O6-C6	-3.81	107.79	118.30
2	B	2001	6PG	P-O6-C6	-2.32	111.92	118.30
2	B	2001	6PG	O3P-P-O6	2.18	112.55	106.73
2	B	2001	6PG	O2-C2-C3	2.81	113.75	107.52

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

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	465/480 (96%)	-0.27	11 (2%) 59 59	8, 16, 30, 52	0
1	B	466/480 (97%)	-0.42	6 (1%) 77 78	7, 13, 24, 46	0
All	All	931/960 (96%)	-0.35	17 (1%) 69 69	7, 15, 28, 52	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	244	GLU	5.3
1	A	2	SER	4.9
1	A	244	GLU	4.5
1	A	243	ASP	4.1
1	B	467	LEU	3.9

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	6PG	A	2002	17/17	0.83	0.18	1.25	25,30,35,42	0
2	6PG	B	2001	17/17	0.86	0.14	0.94	19,27,29,29	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.