



wwPDB X-ray Structure Validation Summary Report

May 26, 2020 – 05:55 am BST

PDB ID : 2GP4
Title : Structure of [FeS]cluster-free Apo Form of 6-Phosphogluconate Dehydratase from *Shewanella oneidensis*
Authors : Schormann, N.; Symersky, J.; Southeast Collaboratory for Structural Genomics (SECSG)
Deposited on : 2006-04-16
Resolution : 2.49 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  symbol.

The following versions of software and data (see [references](#) ) 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

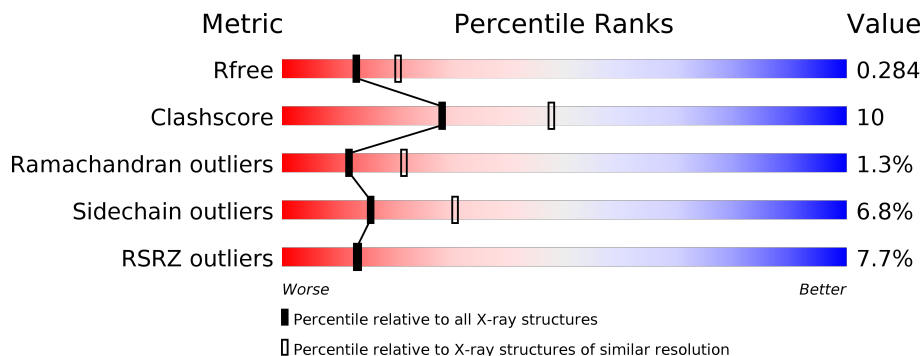
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.49 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 $\leq 5\%$. 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	628	 6% 64% 15% 17%
1	B	628	 6% 65% 12% 21%

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 8116 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 dehydratase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	523	3892	2445	685	743	5	14	0	0	0
1	B	495	3671	2307	649	697	4	14	0	0	0

There are 74 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MSE	-	EXPRESSION TAG	GB 24348501
A	-18	GLY	-	EXPRESSION TAG	GB 24348501
A	-17	SER	-	EXPRESSION TAG	GB 24348501
A	-16	SER	-	EXPRESSION TAG	GB 24348501
A	-15	HIS	-	EXPRESSION TAG	GB 24348501
A	-14	HIS	-	EXPRESSION TAG	GB 24348501
A	-13	HIS	-	EXPRESSION TAG	GB 24348501
A	-12	HIS	-	EXPRESSION TAG	GB 24348501
A	-11	HIS	-	EXPRESSION TAG	GB 24348501
A	-10	HIS	-	EXPRESSION TAG	GB 24348501
A	-9	SER	-	EXPRESSION TAG	GB 24348501
A	-8	SER	-	EXPRESSION TAG	GB 24348501
A	-7	GLY	-	EXPRESSION TAG	GB 24348501
A	-6	LEU	-	EXPRESSION TAG	GB 24348501
A	-5	VAL	-	EXPRESSION TAG	GB 24348501
A	-4	PRO	-	EXPRESSION TAG	GB 24348501
A	-3	ARG	-	EXPRESSION TAG	GB 24348501
A	-2	GLY	-	EXPRESSION TAG	GB 24348501
A	-1	SER	-	EXPRESSION TAG	GB 24348501
A	0	HIS	-	EXPRESSION TAG	GB 24348501
A	1	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	77	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	111	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	122	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	134	MSE	MET	MODIFIED RESIDUE	GB 24348501

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	144	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	173	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	181	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	233	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	237	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	260	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	305	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	353	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	484	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	507	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	512	MSE	MET	MODIFIED RESIDUE	GB 24348501
A	579	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	-19	MSE	-	EXPRESSION TAG	GB 24348501
B	-18	GLY	-	EXPRESSION TAG	GB 24348501
B	-17	SER	-	EXPRESSION TAG	GB 24348501
B	-16	SER	-	EXPRESSION TAG	GB 24348501
B	-15	HIS	-	EXPRESSION TAG	GB 24348501
B	-14	HIS	-	EXPRESSION TAG	GB 24348501
B	-13	HIS	-	EXPRESSION TAG	GB 24348501
B	-12	HIS	-	EXPRESSION TAG	GB 24348501
B	-11	HIS	-	EXPRESSION TAG	GB 24348501
B	-10	HIS	-	EXPRESSION TAG	GB 24348501
B	-9	SER	-	EXPRESSION TAG	GB 24348501
B	-8	SER	-	EXPRESSION TAG	GB 24348501
B	-7	GLY	-	EXPRESSION TAG	GB 24348501
B	-6	LEU	-	EXPRESSION TAG	GB 24348501
B	-5	VAL	-	EXPRESSION TAG	GB 24348501
B	-4	PRO	-	EXPRESSION TAG	GB 24348501
B	-3	ARG	-	EXPRESSION TAG	GB 24348501
B	-2	GLY	-	EXPRESSION TAG	GB 24348501
B	-1	SER	-	EXPRESSION TAG	GB 24348501
B	0	HIS	-	EXPRESSION TAG	GB 24348501
B	1	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	77	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	111	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	122	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	134	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	144	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	173	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	181	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	233	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	237	MSE	MET	MODIFIED RESIDUE	GB 24348501

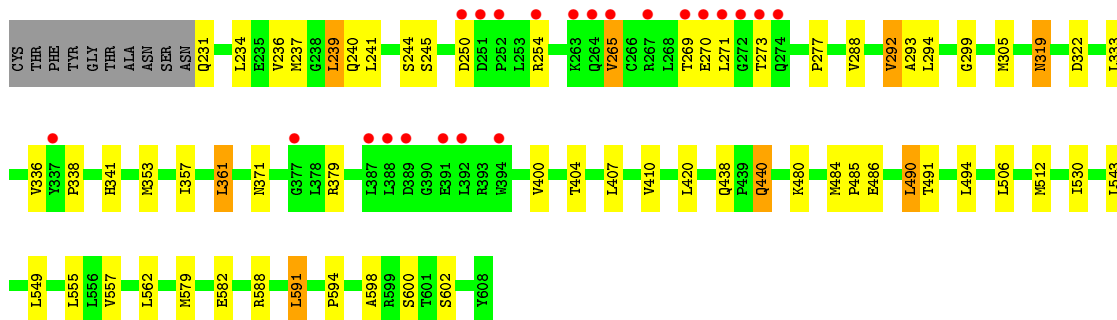
Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	260	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	305	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	353	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	484	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	507	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	512	MSE	MET	MODIFIED RESIDUE	GB 24348501
B	579	MSE	MET	MODIFIED RESIDUE	GB 24348501

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	292	Total O 292 292	0	0
2	B	261	Total O 261 261	0	0



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	61.36Å 118.65Å 160.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.68 – 2.49 47.65 – 2.49	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.68-2.49) 99.5 (47.65-2.49)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.39 (at 2.48Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.230 , 0.291 0.226 , 0.284	Depositor DCC
R_{free} test set	2086 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	31.8	Xtriage
Anisotropy	0.077	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 53.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8116	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.77% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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	0.37	0/3934	0.55	0/5304
1	B	0.36	0/3710	0.56	0/5003
All	All	0.37	0/7644	0.55	0/10307

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 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	3892	0	3949	85	0
1	B	3671	0	3735	65	0
2	A	292	0	0	7	0
2	B	261	0	0	2	0
All	All	8116	0	7684	150	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 10.

The worst 5 of 150 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:B:122:MSE:HE2	1:B:486:GLU:H	1.17	1.08

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:305:MSE:HE2	1:B:598:ALA:HB2	1.34	1.05
1:A:119:GLN:HB2	1:A:120:PRO:HD3	1.40	1.04
1:B:305:MSE:CE	1:B:598:ALA:HB2	1.89	1.02
1:B:305:MSE:HE2	1:B:598:ALA:CB	1.97	0.94

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	509/628 (81%)	486 (96%)	16 (3%)	7 (1%)	11	20
1	B	483/628 (77%)	461 (95%)	16 (3%)	6 (1%)	13	24
All	All	992/1256 (79%)	947 (96%)	32 (3%)	13 (1%)	12	21

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	97	GLU
1	A	158	VAL
1	B	98	VAL
1	B	153	ILE
1	A	153	ILE

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	416/482 (86%)	385 (92%)	31 (8%)	13	26
1	B	392/482 (81%)	368 (94%)	24 (6%)	18	36
All	All	808/964 (84%)	753 (93%)	55 (7%)	16	30

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	490	LEU
1	A	591	LEU
1	B	490	LEU
1	A	494	LEU
1	A	512	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	6	GLN
1	B	96	GLN
1	B	248	ASN
1	A	438	GLN
1	B	319	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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 [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(Å ²)	Q < 0.9
1	A	509/628 (81%)	0.19	40 (7%) 12 12	16, 26, 49, 70	0
1	B	481/628 (76%)	0.17	36 (7%) 14 14	16, 25, 48, 62	0
All	All	990/1256 (78%)	0.18	76 (7%) 13 13	16, 25, 49, 70	0

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	215	TYR	6.5
1	B	389	ASP	6.5
1	B	273	THR	6.3
1	B	271	LEU	6.2
1	A	91	LEU	6.0

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.