



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 06:27 PM UTC

PDB ID : 6GPN / pdb_00006gpn
Title : Crystal Structure of the CsiD Glutarate Hydroxylase in complex with N-Oxalylglycine
Authors : Williams, R.M.; Mayans, O.; Hartig, J.S.
Deposited on : 2018-06-06
Resolution : 2.20 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

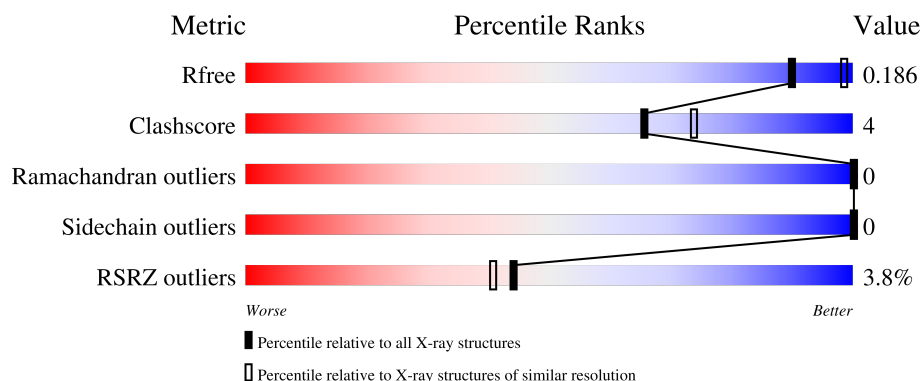
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	180053	6164 (2.20-2.20)
Clashscore	190562	6851 (2.20-2.20)
Ramachandran outliers	187476	6768 (2.20-2.20)
Sidechain outliers	187428	6769 (2.20-2.20)
RSRZ outliers	180081	6166 (2.20-2.20)

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 $\leq 5\%$. 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	353	 3% 74% 8% 18%
1	B	353	 3% 76% 7% 17%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4996 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 Glutarate 2-hydroxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	289	Total	C	N	O	S	0	1	0
			2336	1499	403	424	10			
1	B	292	Total	C	N	O	S	0	2	0
			2376	1524	411	431	10			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP P76621
A	-18	GLY	-	expression tag	UNP P76621
A	-17	SER	-	expression tag	UNP P76621
A	-16	SER	-	expression tag	UNP P76621
A	-15	HIS	-	expression tag	UNP P76621
A	-14	HIS	-	expression tag	UNP P76621
A	-13	HIS	-	expression tag	UNP P76621
A	-12	HIS	-	expression tag	UNP P76621
A	-11	HIS	-	expression tag	UNP P76621
A	-10	HIS	-	expression tag	UNP P76621
A	-9	SER	-	expression tag	UNP P76621
A	-8	SER	-	expression tag	UNP P76621
A	-7	GLY	-	expression tag	UNP P76621
A	-6	LEU	-	expression tag	UNP P76621
A	-5	VAL	-	expression tag	UNP P76621
A	-4	PRO	-	expression tag	UNP P76621
A	-3	ARG	-	expression tag	UNP P76621
A	-2	GLY	-	expression tag	UNP P76621
A	-1	SER	-	expression tag	UNP P76621
A	0	HIS	-	expression tag	UNP P76621
A	326	LEU	-	expression tag	UNP P76621
A	327	GLU	-	expression tag	UNP P76621
A	328	HIS	-	expression tag	UNP P76621
A	329	HIS	-	expression tag	UNP P76621
A	330	HIS	-	expression tag	UNP P76621

Continued on next page...

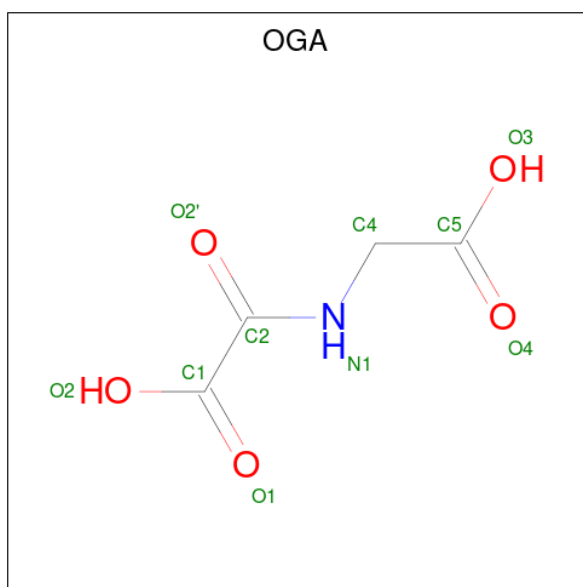
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	331	HIS	-	expression tag	UNP P76621
A	332	HIS	-	expression tag	UNP P76621
A	333	HIS	-	expression tag	UNP P76621
B	-19	MET	-	initiating methionine	UNP P76621
B	-18	GLY	-	expression tag	UNP P76621
B	-17	SER	-	expression tag	UNP P76621
B	-16	SER	-	expression tag	UNP P76621
B	-15	HIS	-	expression tag	UNP P76621
B	-14	HIS	-	expression tag	UNP P76621
B	-13	HIS	-	expression tag	UNP P76621
B	-12	HIS	-	expression tag	UNP P76621
B	-11	HIS	-	expression tag	UNP P76621
B	-10	HIS	-	expression tag	UNP P76621
B	-9	SER	-	expression tag	UNP P76621
B	-8	SER	-	expression tag	UNP P76621
B	-7	GLY	-	expression tag	UNP P76621
B	-6	LEU	-	expression tag	UNP P76621
B	-5	VAL	-	expression tag	UNP P76621
B	-4	PRO	-	expression tag	UNP P76621
B	-3	ARG	-	expression tag	UNP P76621
B	-2	GLY	-	expression tag	UNP P76621
B	-1	SER	-	expression tag	UNP P76621
B	0	HIS	-	expression tag	UNP P76621
B	326	LEU	-	expression tag	UNP P76621
B	327	GLU	-	expression tag	UNP P76621
B	328	HIS	-	expression tag	UNP P76621
B	329	HIS	-	expression tag	UNP P76621
B	330	HIS	-	expression tag	UNP P76621
B	331	HIS	-	expression tag	UNP P76621
B	332	HIS	-	expression tag	UNP P76621
B	333	HIS	-	expression tag	UNP P76621

- Molecule 2 is FE (II) ION (CCD ID: FE2) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Fe 1 1	0	0
2	B	1	Total Fe 1 1	0	0

- Molecule 3 is N-OXALYLGLYCINE (CCD ID: OGA) (formula: C₄H₅NO₅).

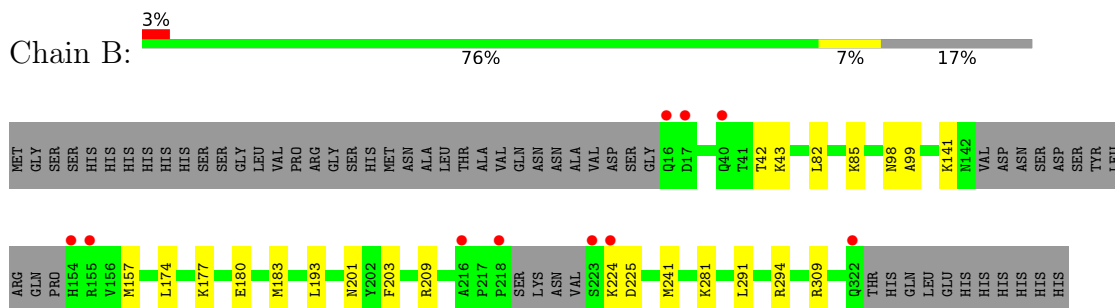


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	N	O	0	0
			10	4	1	5		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	100	Total	O	0	0
			100	100		
4	B	172	Total	O	0	0
			172	172		

- Molecule 1: Glutarate 2-hydroxylase



4 Data and refinement statistics

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, α , β , γ	121.38Å 121.38Å 136.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.91 – 2.20 42.91 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (42.91-2.20) 100.0 (42.91-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.18	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.33 (at 1.98Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.187 , 0.211 0.189 , 0.186	Depositor DCC
R_{free} test set	2180 reflections (3.00%)	wwPDB-VP
Wilson B-factor (Å ²)	37.1	Xtriage
Anisotropy	0.225	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 47.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	4996	wwPDB-VP
Average B, all atoms (Å ²)	52.0	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 28.91 % 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 1.6983e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: OGA, FE2

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	0.35	1/2395 (0.0%)	0.55	1/3248 (0.0%)
1	B	0.35	0/2440	0.57	2/3308 (0.1%)
All	All	0.35	1/4835 (0.0%)	0.56	3/6556 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	40	GLN	CD-NE2	6.39	1.46	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	40	GLN	CB-CA-C	6.23	120.78	110.81
1	B	201	ASN	CA-C-N	-5.17	111.90	120.68
1	B	201	ASN	C-N-CA	-5.17	111.90	120.68

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	2336	0	2265	17	0
1	B	2376	0	2322	18	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	B	10	0	3	1	0
4	A	100	0	0	1	0
4	B	172	0	0	2	0
All	All	4996	0	4590	35	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 4.

The worst 5 of 35 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:157:MET:HE1	3:B:402:OGA:H4C1	1.52	0.92
1:B:180:GLU:OE2	1:B:183:MET:HG2	1.73	0.88
1:B:241:MET:HE3	1:B:291:LEU:HD11	1.59	0.82
1:A:182:ASN:HD22	1:A:302:ASP:HB2	1.51	0.74
1:A:203:PHE:O	1:A:209:ARG:NH2	2.33	0.60

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	284/353 (80%)	278 (98%)	6 (2%)	0	100	100
1	B	288/353 (82%)	283 (98%)	5 (2%)	0	100	100
All	All	572/706 (81%)	561 (98%)	11 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	247/312 (79%)	247 (100%)	0	100	100
1	B	255/312 (82%)	255 (100%)	0	100	100
All	All	502/624 (80%)	502 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
1	A	44	GLN
1	A	98	ASN
1	A	182	ASN
1	A	187	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 for Mogul analysis.

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
3	OGA	B	402	2	9,9,9	2.52	2 (22%)	8,11,11	1.57	2 (25%)

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
3	OGA	B	402	2	-	2/8/9/9	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	OGA	C2-N1	6.34	1.44	1.33
3	B	402	OGA	C2-C1	-2.91	1.50	1.54

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	402	OGA	O3-C5-C4	2.13	120.90	112.81
3	B	402	OGA	O2-C1-O1	-2.07	118.97	123.90

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	402	OGA	N1-C4-C5-O3
3	B	402	OGA	N1-C4-C5-O4

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	402	OGA	1	0

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(Å ²)	Q<0.9
1	A	289/353 (81%)	0.10	12 (4%) 40 37	25, 52, 85, 117	1 (0%)
1	B	292/353 (82%)	-0.10	10 (3%) 48 45	22, 45, 71, 103	2 (0%)
All	All	581/706 (82%)	0.00	22 (3%) 44 41	22, 48, 81, 117	3 (0%)

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	322	GLN	8.8
1	A	154	HIS	5.5
1	B	154	HIS	5.5
1	B	17	ASP	5.0
1	A	142	ASN	4.9

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 oligosaccharides 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($\text{\AA}^2$)	Q<0.9
3	OGA	B	402	10/10	0.77	0.20	87,93,95,97	0
2	FE2	B	401	1/1	0.99	0.03	55,55,55,55	0
2	FE2	A	401	1/1	0.99	0.04	66,66,66,66	0

6.5 Other polymers [i](#)

There are no such residues in this entry.