



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 04:41 AM UTC

PDB ID : 1F1R / pdb_00001f1r
Title : CRYSTAL STRUCTURE OF HOMOPROTOCATECHUATE 2,3-DIOXYGENASE FROM ARTHROBACTER GLOBIFORMIS (NATIVE, NON-CRYO)
Authors : Vetting, M.W.; Lipscomb, J.D.; Wackett, L.P.; Que Jr., L.; Ohlendorf, D.H.
Deposited on : 2000-05-19
Resolution : 1.80 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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
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

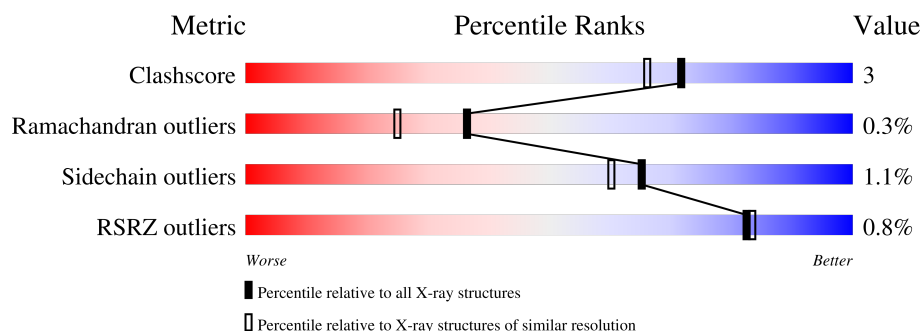
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	190562	8479 (1.80-1.80)
Ramachandran outliers	187476	8391 (1.80-1.80)
Sidechain outliers	187428	8390 (1.80-1.80)
RSRZ outliers	180081	7663 (1.80-1.80)

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	323	85% 13% ..
1	B	323	85% 13% ..

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5547 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 HOMOPROTOCATECHUATE 2,3-DIOXYGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	321	Total	C	N	O	S	0	0	0
			2597	1644	454	492	7			
1	B	321	Total	C	N	O	S	0	0	0
			2597	1644	454	492	7			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	233	GLY	-	SEE REMARK 999	UNP Q44048
A	234	ALA	ARG	SEE REMARK 999	UNP Q44048
A	235	LEU	PRO	SEE REMARK 999	UNP Q44048
A	236	ARG	CYS	SEE REMARK 999	UNP Q44048
A	237	ILE	ALA	SEE REMARK 999	UNP Q44048
A	239	ASP	ALA	SEE REMARK 999	UNP Q44048
A	240	ARG	THR	SEE REMARK 999	UNP Q44048
A	241	ILE	GLY	SEE REMARK 999	UNP Q44048
A	242	GLU	SER	SEE REMARK 999	UNP Q44048
A	243	ARG	ASN	SEE REMARK 999	UNP Q44048
A	246	GLY	ARG	SEE REMARK 999	UNP Q44048
A	247	ARG	PRO	SEE REMARK 999	UNP Q44048
A	248	HIS	ALA	SEE REMARK 999	UNP Q44048
A	249	GLY	PRO	SEE REMARK 999	UNP Q44048
B	233	GLY	-	SEE REMARK 999	UNP Q44048
B	234	ALA	ARG	SEE REMARK 999	UNP Q44048
B	235	LEU	PRO	SEE REMARK 999	UNP Q44048
B	236	ARG	CYS	SEE REMARK 999	UNP Q44048
B	237	ILE	ALA	SEE REMARK 999	UNP Q44048
B	239	ASP	ALA	SEE REMARK 999	UNP Q44048
B	240	ARG	THR	SEE REMARK 999	UNP Q44048
B	241	ILE	GLY	SEE REMARK 999	UNP Q44048
B	242	GLU	SER	SEE REMARK 999	UNP Q44048
B	243	ARG	ASN	SEE REMARK 999	UNP Q44048
B	246	GLY	ARG	SEE REMARK 999	UNP Q44048

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	247	ARG	PRO	SEE REMARK 999	UNP Q44048
B	248	HIS	ALA	SEE REMARK 999	UNP Q44048
B	249	GLY	PRO	SEE REMARK 999	UNP Q44048

- Molecule 2 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	190	Total 190	O 190	0	0
3	B	161	Total 161	O 161	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	138.44Å 59.15Å 102.80Å 90.00° 118.26° 90.00°	Depositor
Resolution (Å)	20.00 – 1.80 20.00 – 1.80	Depositor EDS
% Data completeness (in resolution range)	88.7 (20.00-1.80) 88.7 (20.00-1.80)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.25 (at 1.76Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.156 , 0.187 0.161 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	11.2	Xtriage
Anisotropy	0.224	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 40.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5547	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.39% 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

5.1 Standard geometry

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

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	1.05	1/2669 (0.0%)	1.27	19/3636 (0.5%)
1	B	1.07	0/2669	1.24	16/3636 (0.4%)
All	All	1.06	1/5338 (0.0%)	1.26	35/7272 (0.5%)

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	2
1	B	0	2
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	157	ASN	CA-C	5.33	1.58	1.52

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	82	LYS	N-CA-C	8.50	121.68	111.82
1	A	201	ASP	N-CA-C	-8.07	103.95	113.88
1	A	82	LYS	N-CA-C	8.07	121.18	111.82
1	A	254	PHE	N-CA-C	-7.53	96.99	108.96
1	B	195	ARG	N-CA-C	-7.42	103.06	112.93
1	A	195	ARG	N-CA-C	-7.38	103.12	112.93
1	A	5	VAL	N-CA-C	-7.34	102.19	109.02
1	A	276	GLY	N-CA-C	6.93	122.55	114.16

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	4	PHE	N-CA-C	6.84	120.41	110.28
1	B	254	PHE	N-CA-C	-6.70	98.31	108.96
1	B	201	ASP	N-CA-C	-6.62	105.74	113.88
1	B	87	VAL	N-CA-C	-6.36	104.30	110.72
1	A	215	VAL	N-CA-C	-6.28	99.44	108.48
1	B	152	ARG	CG-CD-NE	-6.27	98.20	112.00
1	A	197	GLN	N-CA-C	6.26	119.61	112.97
1	A	56	LEU	N-CA-C	6.15	117.98	111.28
1	B	77	PHE	N-CA-C	-6.10	98.57	108.52
1	A	5	VAL	CA-C-N	6.10	126.07	119.78
1	A	5	VAL	C-N-CA	6.10	126.07	119.78
1	A	163	VAL	CA-C-N	-5.98	112.70	119.28
1	A	163	VAL	C-N-CA	-5.98	112.70	119.28
1	A	61	HIS	N-CA-C	5.91	117.72	111.28
1	B	255	TYR	N-CA-C	5.77	118.18	109.41
1	A	119	GLU	N-CA-C	-5.74	97.08	107.75
1	B	49	ASN	N-CA-C	5.66	121.08	113.72
1	A	152	ARG	CG-CD-NE	-5.53	99.83	112.00
1	B	197	GLN	N-CA-C	5.52	118.82	112.97
1	B	276	GLY	N-CA-C	5.52	120.84	114.16
1	B	272	ASP	CB-CA-C	-5.47	101.04	111.48
1	B	56	LEU	N-CA-C	5.41	117.18	111.28
1	B	61	HIS	N-CA-C	5.36	117.13	111.28
1	A	194	HIS	N-CA-C	5.24	117.35	109.23
1	A	255	TYR	N-CA-C	5.22	117.35	109.41
1	A	268	ILE	N-CA-C	-5.20	100.83	108.11
1	B	215	VAL	N-CA-C	-5.20	100.85	108.85

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	126	TYR	Sidechain
1	A	305	TYR	Sidechain
1	B	126	TYR	Sidechain
1	B	305	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	2597	0	2471	17	0
1	B	2597	0	2471	18	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	190	0	0	1	0
3	B	161	0	0	1	0
All	All	5547	0	4942	33	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 (33) 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:30:LYS:HD3	3:A:768:HOH:O	1.92	0.69
1:B:119:GLU:HG3	1:B:314:LEU:HD22	1.78	0.65
1:B:104:ARG:HD3	1:B:108:PHE:CE2	2.33	0.63
1:B:176:ARG:NH2	1:B:277:ASP:O	2.33	0.58
1:A:26:THR:O	1:A:68:GLN:NE2	2.38	0.56
1:B:194:HIS:HB3	1:B:200:HIS:HA	1.88	0.55
1:A:119:GLU:HG3	1:A:314:LEU:HG	1.87	0.55
1:A:255:TYR:C	1:A:255:TYR:CD1	2.85	0.54
1:A:221:GLU:H	1:A:224:ASN:HD22	1.54	0.54
1:A:194:HIS:HB3	1:A:200:HIS:HA	1.90	0.53
1:B:221:GLU:H	1:B:224:ASN:HD22	1.58	0.50
1:A:104:ARG:HD3	1:A:108:PHE:CE2	2.46	0.50
1:B:255:TYR:C	1:B:255:TYR:CD1	2.89	0.50
1:A:147:ALA:CB	1:A:227:GLN:HG2	2.42	0.50
1:B:28:LEU:H	1:B:68:GLN:NE2	2.10	0.49
1:A:140:GLN:HE21	1:B:252:ASN:HD22	1.59	0.49
1:B:106:GLU:HB2	3:B:783:HOH:O	2.13	0.48
1:A:147:ALA:HB3	1:A:227:GLN:HG2	1.97	0.46
1:A:140:GLN:HE21	1:B:252:ASN:ND2	2.15	0.45
1:A:23:ILE:HG12	1:A:66:LEU:HD23	1.99	0.44
1:B:162:ASP:OD1	1:B:164:PRO:HD2	2.18	0.44
1:A:49:ASN:C	1:A:50:THR:CG2	2.91	0.44
1:A:300:VAL:HG12	1:A:301:VAL:N	2.32	0.43
1:A:218:ALA:HA	1:A:269:TYR:O	2.19	0.43
1:B:279:ASP:O	1:B:280:ASN:C	2.62	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:55:SER:OG	1:B:152:ARG:HA	2.18	0.42
1:B:147:ALA:HB3	1:B:227:GLN:HG2	2.01	0.42
1:B:49:ASN:C	1:B:50:THR:CG2	2.92	0.41
1:B:147:ALA:CB	1:B:227:GLN:HG2	2.49	0.41
1:A:279:ASP:O	1:A:280:ASN:C	2.63	0.41
1:A:195:ARG:HD2	1:A:195:ARG:HA	1.99	0.41
1:B:245:PRO:HA	1:B:255:TYR:O	2.21	0.41
1:B:246:GLY:HA2	1:B:296:TRP:CZ3	2.57	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	319/323 (99%)	312 (98%)	6 (2%)	1 (0%)	36	25
1	B	319/323 (99%)	313 (98%)	5 (2%)	1 (0%)	36	25
All	All	638/646 (99%)	625 (98%)	11 (2%)	2 (0%)	36	25

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	271	GLN
1	B	271	GLN

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	277/279 (99%)	274 (99%)	3 (1%)	65	60
1	B	277/279 (99%)	274 (99%)	3 (1%)	65	60
All	All	554/558 (99%)	548 (99%)	6 (1%)	65	60

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	120	ASP
1	A	184	SER
1	A	236	ARG
1	B	120	ASP
1	B	184	SER
1	B	314	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	134	HIS
1	A	158	GLN
1	A	224	ASN
1	B	68	GLN
1	B	140	GLN
1	B	158	GLN
1	B	224	ASN
1	B	252	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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

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	321/323 (99%)	-0.56	2 (0%) 85 86	6, 12, 26, 39	0
1	B	321/323 (99%)	-0.61	3 (0%) 81 81	6, 12, 25, 40	0
All	All	642/646 (99%)	-0.58	5 (0%) 82 83	6, 12, 26, 40	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	290	ASN	2.8
1	A	184	SER	2.3
1	A	290	ASN	2.2
1	B	184	SER	2.2
1	B	3	ASN	2.1

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
2	MN	B	500	1/1	0.98	0.02	15,15,15,15	1
2	MN	A	500	1/1	0.99	0.02	14,14,14,14	1

6.5 Other polymers [i](#)

There are no such residues in this entry.