W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 18, 2020 – 07:47 am BST

PDB ID	:	1R8W
Title	:	Native structure of the B12-independent glycerol dehydratase from clostridium butyricum
Authors	:	Lanzilotta, W.N.; Soucaille, P.; O'Brien, J.R.; Raynaud, C.
Deposited on	:	2003-10-28
Resolution	:	2.50 Å(reported)

The following versions of software and data (see references (1)) were used in the production of this report:

MolProbity	:	4.02b-467
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 2.50 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	Percentile Ranks	Value
Rfree		0.204
Clashscore		9
Ramachandran outliers		0.5%
Sidechain outliers		2.7%
RSRZ outliers		0.3%

Worse ← → *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
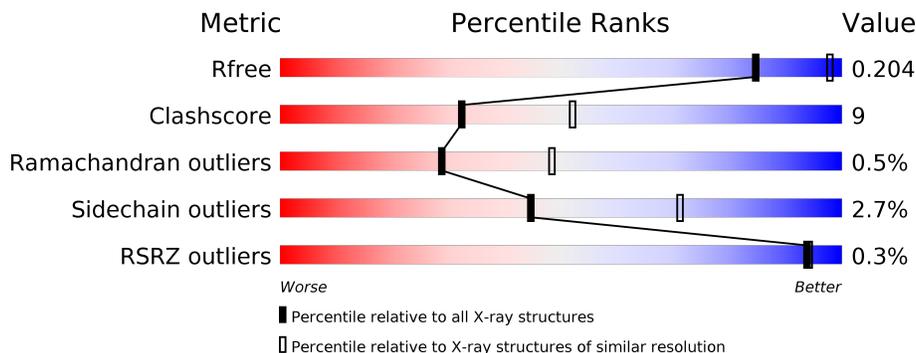
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 <=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	787	80% ... 18% ·
1	B	787	79% ... 20% ·

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12905 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 glycerol dehydratase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	786	Total 6197	C 3918	N 1059	O 1194	S 26	0	1	0
1	B	786	Total 6197	C 3918	N 1059	O 1194	S 26	0	1	0

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	255	Total 255	O 255	0	0
2	B	256	Total 256	O 256	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: glycerol dehydratase

Chain A:

- Molecule 1: glycerol dehydratase

Chain B:

F648 V669 R673 P680 N685 K689 L700 F701 N702 Q703 D713 M717 N718 L719 L722 H733 V734 Q735 I739 L744 E752 I758 V759 R760 V761 A762 G763 H785 V786 M787
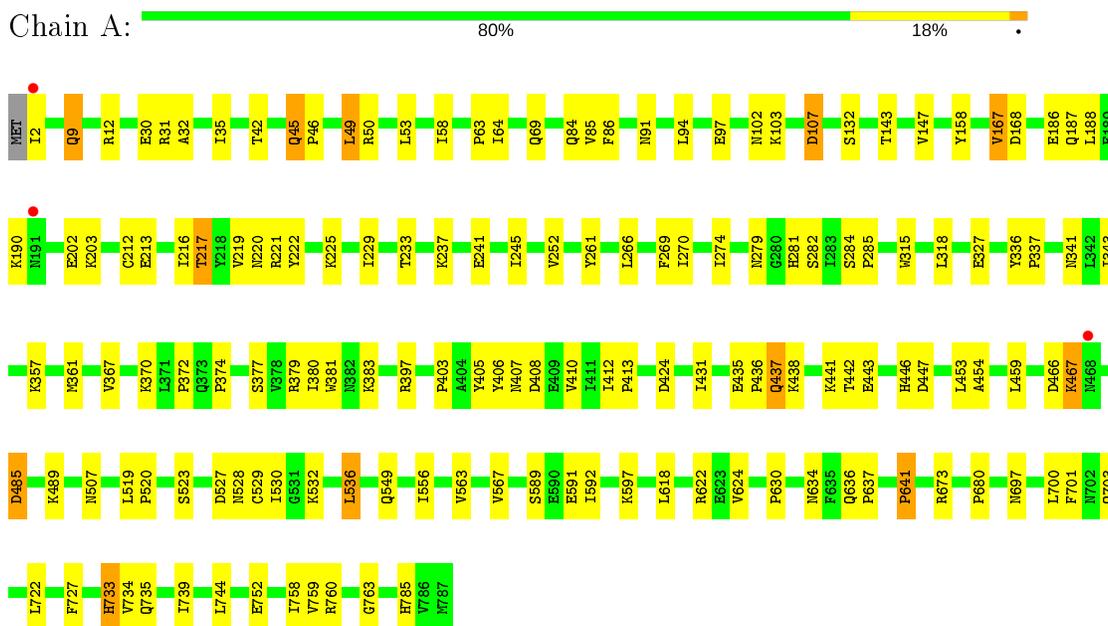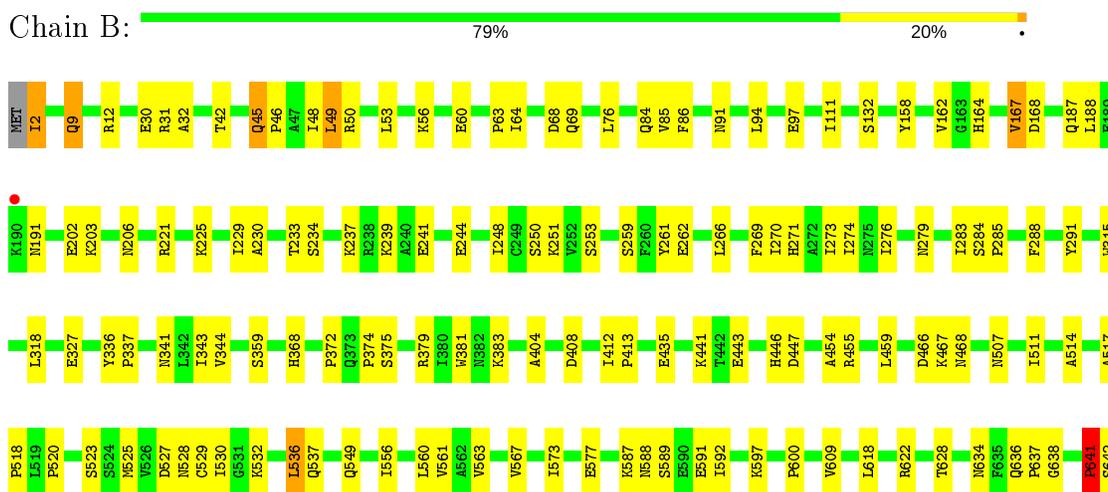
4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	103.43Å 212.91Å 199.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.16 – 2.50 42.16 – 2.05	Depositor EDS
% Data completeness (in resolution range)	98.4 (42.16-2.50) 68.3 (42.16-2.05)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.09	Depositor
$<I/\sigma(I)>$ [1]	1.37 (at 2.05Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.171 , 0.203 0.172 , 0.204	Depositor DCC
R_{free} test set	4810 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	26.9	Xtriage
Anisotropy	0.264	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 36.8	EDS
L-test for twinning[2]	$<\lvert L\rvert> = 0.51$, $<L^2> = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	12905	wwPDB-VP
Average B, all atoms (Å^2)	28.0	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 51.55 % 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 5.5242e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<\lvert L\rvert>$, $<L^2>$ 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)

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	0.33	0/6319	0.58	0/8530
1	B	0.33	0/6319	0.57	0/8530
All	All	0.33	0/12638	0.58	0/17060

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	6197	0	6130	104	0
1	B	6197	0	6130	112	0
2	A	255	0	0	5	0
2	B	256	0	0	3	0
All	All	12905	0	12260	216	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 9.

The worst 5 of 216 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:42:THR:O	1:A:50:ARG:HD3	1.78	0.83

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:31:ARG:HB3	1:B:84:GLN:HG3	1.65	0.78
1:A:252:VAL:HG13	1:A:266:LEU:HD13	1.66	0.77
1:A:31:ARG:HB3	1:A:84:GLN:HG3	1.67	0.76
1:B:739:ILE:HD12	1:B:744:LEU:HD11	1.66	0.76

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	785/787 (100%)	748 (95%)	33 (4%)	4 (0%)	29	48
1	B	785/787 (100%)	743 (95%)	38 (5%)	4 (0%)	29	48
All	All	1570/1574 (100%)	1491 (95%)	71 (4%)	8 (0%)	29	48

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	641	PRO
1	B	641	PRO
1	A	589	SER
1	A	91	ASN
1	A	467	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	676/676 (100%)	657 (97%)	19 (3%)	43	70
1	B	676/676 (100%)	659 (98%)	17 (2%)	47	73
All	All	1352/1352 (100%)	1316 (97%)	36 (3%)	44	71

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	641	PRO
1	B	9	GLN
1	B	648	PHE
1	A	759	VAL
1	B	45	GLN

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	735	GLN
1	B	206	ASN
1	B	733	HIS
1	B	45	GLN
1	B	164	HIS

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 (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	786/787 (99%)	-0.63	3 (0%)	92	93	14, 26, 47, 64	0
1	B	786/787 (99%)	-0.58	1 (0%)	95	96	15, 26, 47, 62	0
All	All	1572/1574 (99%)	-0.60	4 (0%)	94	94	14, 26, 47, 64	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	ILE	3.0
1	A	468	ASN	2.7
1	B	190	LYS	2.1
1	A	191	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 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.