



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 24, 2023 – 11:13 PM EST

PDB ID : 3S30
Title : Crystal structure of a putative glycoside hydrolase (BVU_0247) from *Bacteroides vulgatus* ATCC 8482 at 2.46 Å resolution
Authors : Joint Center for Structural Genomics (JCSG)
Deposited on : 2011-05-17
Resolution : 2.46 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 : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.31.2
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) : 2.31.2

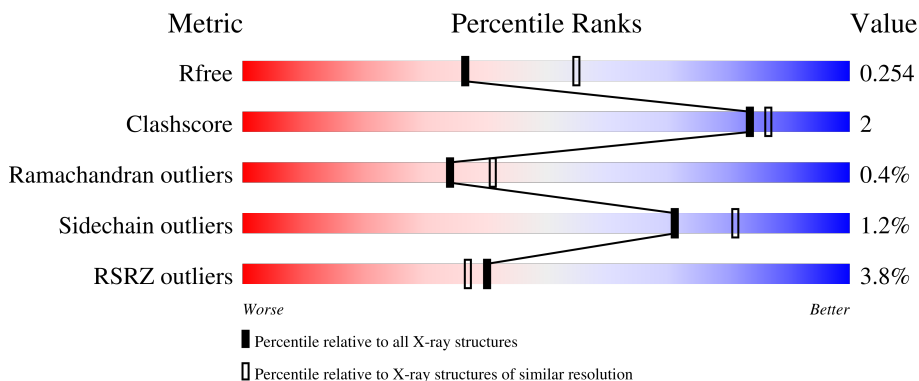
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.46 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	354	
1	B	354	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4145 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 Hypothetical glycoside hydrolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	254	1981	1266	315	394	2	4	0	1	0
1	B	256	1994	1274	317	397	2	4	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	expression tag	UNP A6KX05
B	0	GLY	-	expression tag	UNP A6KX05

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	92	Total	O	0	0
			92	92		
3	B	76	Total	O	0	0
			76	76		

4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	146.75Å 146.75Å 91.51Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.66 – 2.46 29.66 – 2.46	Depositor EDS
% Data completeness (in resolution range)	99.2 (29.66-2.46) 99.2 (29.66-2.46)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.26 (at 2.45Å)	Xtrriage
Refinement program	REFMAC 5.5.0110	Depositor
R, R_{free}	0.233 , 0.251 0.234 , 0.254	Depositor DCC
R_{free} test set	2033 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å ²)	46.4	Xtrriage
Anisotropy	0.105	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 39.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.033 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	4145	wwPDB-VP
Average B, all atoms (Å ²)	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% 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: NA

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	0.47	0/2029	0.57	0/2749
1	B	0.47	0/2040	0.56	0/2764
All	All	0.47	0/4069	0.56	0/5513

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

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	1981	0	1897	8	0
1	B	1994	0	1906	11	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	92	0	0	0	0
3	B	76	0	0	0	0
All	All	4145	0	3803	19	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 2.

All (19) 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:211:MSE:SE	1:A:229:MSE:HA	2.40	0.72
1:A:209:VAL:HG23	1:A:211:MSE:HG3	1.73	0.69
1:B:211:MSE:SE	1:B:229:MSE:HA	2.53	0.58
1:B:304:MSE:CE	1:B:307:ALA:HB2	2.34	0.57
1:A:249:LEU:HD11	1:A:373:LEU:HD11	1.89	0.53
1:B:171:LEU:HB3	1:B:229:MSE:HE3	1.94	0.49
1:B:304:MSE:HE2	1:B:307:ALA:HB2	1.95	0.47
1:A:209:VAL:HG23	1:A:211:MSE:CG	2.45	0.46
1:B:274:PHE:CZ	1:B:313:THR:HG22	2.50	0.46
1:A:304:MSE:HE1	1:A:306:ALA:O	2.16	0.45
1:A:211:MSE:SE	1:A:228:ALA:O	2.85	0.45
1:B:211:MSE:SE	1:B:229:MSE:HG2	2.67	0.44
1:B:162:TRP:CD2	1:B:218:LEU:HD21	2.53	0.43
1:B:306:ALA:HB2	1:B:328:TYR:CD2	2.53	0.43
1:A:211:MSE:HE2	1:A:360:PHE:CZ	2.53	0.43
1:B:274:PHE:CE1	1:B:313:THR:HG22	2.53	0.43
1:B:304:MSE:HE1	1:B:307:ALA:HB2	2.02	0.42
1:B:209:VAL:HG23	1:B:211:MSE:HG3	2.01	0.41
1:A:162:TRP:CD2	1:A:218:LEU:HD21	2.56	0.41

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	253/354 (72%)	245 (97%)	7 (3%)	1 (0%)	34	41
1	B	254/354 (72%)	243 (96%)	10 (4%)	1 (0%)	34	41
All	All	507/708 (72%)	488 (96%)	17 (3%)	2 (0%)	34	41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	204	SER
1	B	204	SER

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	214/296 (72%)	212 (99%)	2 (1%)	78 86
1	B	215/296 (73%)	212 (99%)	3 (1%)	67 77
All	All	429/592 (72%)	424 (99%)	5 (1%)	71 81

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	211	MSE
1	A	245	GLU
1	B	240	THR
1	B	245	GLU
1	B	338	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	179	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

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 [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	250/354 (70%)	0.05	11 (4%) 34 32	46, 60, 88, 114	0
1	B	252/354 (71%)	0.01	8 (3%) 47 44	46, 62, 92, 121	0
All	All	502/708 (70%)	0.03	19 (3%) 40 37	46, 61, 90, 121	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	125	THR	6.2
1	B	123	PRO	3.3
1	B	338	GLN	3.2
1	B	191	THR	3.0
1	A	351	LEU	2.5
1	A	338	GLN	2.4
1	A	218	LEU	2.4
1	B	302	GLU	2.4
1	A	217	ASN	2.3
1	B	203	GLY	2.3
1	B	285	LYS	2.3
1	B	349	ILE	2.3
1	A	215	ALA	2.2
1	B	342	ASN	2.2
1	A	163	SER	2.2
1	A	329	GLU	2.1
1	A	378	LYS	2.1
1	A	350	VAL	2.1
1	A	164	SER	2.0

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 monosaccharides 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(Å ²)	Q<0.9
2	NA	B	384	1/1	0.71	0.15	56,56,56,56	0
2	NA	A	383	1/1	0.94	0.30	47,47,47,47	0

6.5 Other polymers [i](#)

There are no such residues in this entry.