



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 10:00 am GMT

PDB ID : 2RDY
Title : Crystal structure of a putative glycoside hydrolase family protein from *Bacillus halodurans*
Authors : Sugadev, R.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2007-09-25
Resolution : 2.03 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
Xtriage (Phenix)	:	1.13
EDS	:	trunk31020
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk31020

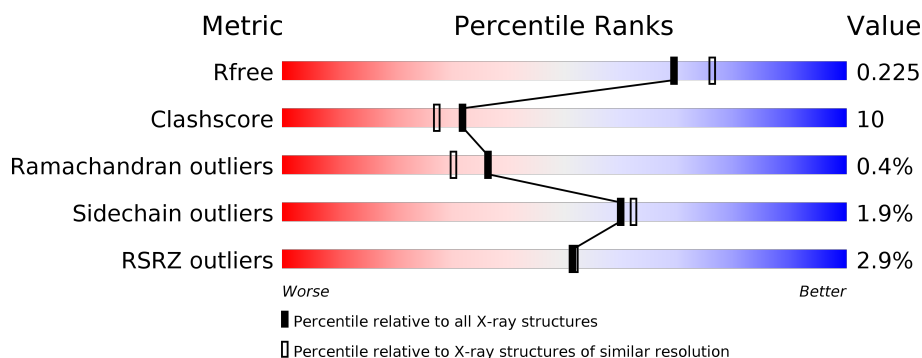
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.03 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}	111664	9172 (2.04-2.00)
Clashscore	122126	10355 (2.04-2.00)
Ramachandran outliers	120053	10237 (2.04-2.00)
Sidechain outliers	120020	10236 (2.04-2.00)
RSRZ outliers	108989	8961 (2.04-2.00)

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. 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	803	<div> <div>2%</div> <div> <div></div> <div>79%</div> <div>18%</div> <div>••</div> </div> </div>
1	B	803	<div> <div>3%</div> <div> <div></div> <div>80%</div> <div>17%</div> <div>••</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13037 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 BH0842 protein.

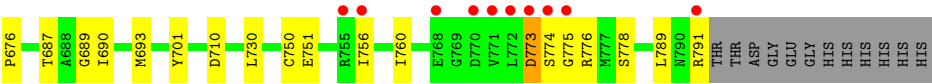
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	787	Total	C	N	O	S	Se	0	0	0
			6239	3955	1076	1181	9	18			
1	B	787	Total	C	N	O	S	Se	0	0	0
			6239	3955	1076	1181	9	18			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	EXPRESSION TAG	UNP Q9KEL0
A	2	SER	-	EXPRESSION TAG	UNP Q9KEL0
A	3	LEU	-	EXPRESSION TAG	UNP Q9KEL0
A	796	GLU	-	EXPRESSION TAG	UNP Q9KEL0
A	797	GLY	-	EXPRESSION TAG	UNP Q9KEL0
A	798	HIS	-	EXPRESSION TAG	UNP Q9KEL0
A	799	HIS	-	EXPRESSION TAG	UNP Q9KEL0
A	800	HIS	-	EXPRESSION TAG	UNP Q9KEL0
A	801	HIS	-	EXPRESSION TAG	UNP Q9KEL0
A	802	HIS	-	EXPRESSION TAG	UNP Q9KEL0
A	803	HIS	-	EXPRESSION TAG	UNP Q9KEL0
B	1	MSE	-	EXPRESSION TAG	UNP Q9KEL0
B	2	SER	-	EXPRESSION TAG	UNP Q9KEL0
B	3	LEU	-	EXPRESSION TAG	UNP Q9KEL0
B	796	GLU	-	EXPRESSION TAG	UNP Q9KEL0
B	797	GLY	-	EXPRESSION TAG	UNP Q9KEL0
B	798	HIS	-	EXPRESSION TAG	UNP Q9KEL0
B	799	HIS	-	EXPRESSION TAG	UNP Q9KEL0
B	800	HIS	-	EXPRESSION TAG	UNP Q9KEL0
B	801	HIS	-	EXPRESSION TAG	UNP Q9KEL0
B	802	HIS	-	EXPRESSION TAG	UNP Q9KEL0
B	803	HIS	-	EXPRESSION TAG	UNP Q9KEL0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	275	Total 275	O 275	0	0
2	B	284	Total 284	O 284	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	75.61Å 149.07Å 164.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.57 – 2.03 47.57 – 2.03	Depositor EDS
% Data completeness (in resolution range)	93.6 (47.57-2.03) 93.8 (47.57-2.03)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.55 (at 2.03Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.198 , 0.225 0.197 , 0.225	Depositor DCC
R_{free} test set	6536 reflections (5.63%)	wwPDB-VP
Wilson B-factor (Å ²)	14.9	Xtriage
Anisotropy	0.294	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 50.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	13037	wwPDB-VP
Average B, all atoms (Å ²)	16.0	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 26.94 % 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 2.4057e-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 [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	0/6378	0.60	2/8621 (0.0%)
1	B	0.34	0/6378	0.60	1/8621 (0.0%)
All	All	0.33	0/12756	0.60	3/17242 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	5	ILE	N-CA-C	-5.28	96.75	111.00
1	A	42	GLU	N-CA-C	-5.14	97.13	111.00
1	A	5	ILE	N-CA-C	-5.10	97.23	111.00

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	6239	0	6036	123	0
1	B	6239	0	6036	112	0
2	A	275	0	0	5	0
2	B	284	0	0	2	0
All	All	13037	0	12072	235	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 235 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:448:MSE:HE2	1:A:525:ASP:HA	1.22	1.17
1:B:690:ILE:HA	1:B:693:MSE:HE3	1.43	0.97
1:B:407:LYS:HE3	1:B:411:ILE:HD11	1.49	0.93
1:B:84:MSE:HE1	1:B:501:PRO:HG2	1.48	0.93
1:A:448:MSE:HE3	1:A:528:LEU:HB2	1.53	0.91

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	783/803 (98%)	749 (96%)	31 (4%)	3 (0%)	36	30
1	B	783/803 (98%)	750 (96%)	29 (4%)	4 (0%)	31	24
All	All	1566/1606 (98%)	1499 (96%)	60 (4%)	7 (0%)	36	30

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	373	ILE
1	B	373	ILE
1	A	505	PRO
1	B	505	PRO
1	B	773	ASP

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	661/654 (101%)	647 (98%)	14 (2%)	56	58
1	B	661/654 (101%)	650 (98%)	11 (2%)	63	66
All	All	1322/1308 (101%)	1297 (98%)	25 (2%)	60	62

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	730	LEU
1	A	766	GLU
1	B	558	GLN
1	A	759	ARG
1	B	91	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	24	ASN
1	B	179	HIS
1	B	662	GLN
1	B	133	GLN
1	B	189	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry

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	769/803 (95%)	0.08	19 (2%) 57 57	7, 14, 28, 42	0
1	B	769/803 (95%)	0.15	26 (3%) 45 45	6, 14, 30, 49	0
All	All	1538/1606 (95%)	0.12	45 (2%) 51 52	6, 14, 28, 49	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	775	GLY	6.4
1	B	773	ASP	5.0
1	B	261	LEU	5.0
1	B	774	SER	4.8
1	B	756	ILE	4.4

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.