



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 05:50 pm GMT

PDB ID : 3G1W
Title : Crystal structure of sugar ABC transporter (sugar-binding protein) from *Bacillus halodurans*
Authors : Zhang, Z.; Eswaramoorthy, S.; Swaminathan, S.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2009-01-30
Resolution : 2.02 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

<http://wwpdb.org/validation/2016/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)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

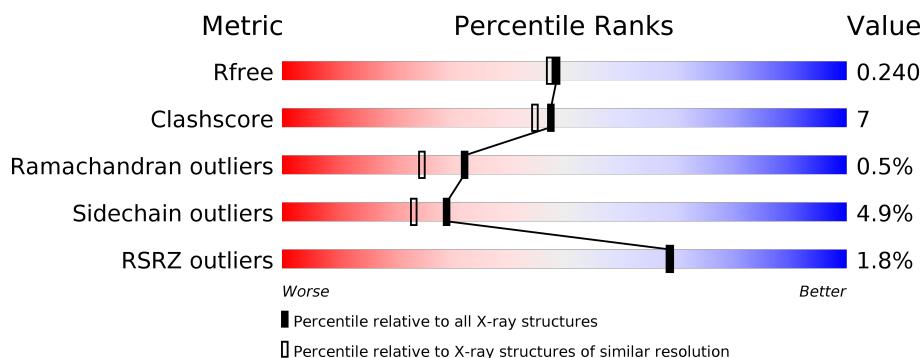
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.02 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}	100719	8396 (2.04-2.00)
Clashscore	112137	9678 (2.04-2.00)
Ramachandran outliers	110173	9566 (2.04-2.00)
Sidechain outliers	110143	9565 (2.04-2.00)
RSRZ outliers	101464	8490 (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	305	
1	B	305	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4699 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 Sugar ABC transporter.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	291	Total	C	N	O	S	Se	0	0	0
			2245	1417	370	451	1	6			
1	B	290	Total	C	N	O	S	Se	0	0	0
			2225	1403	366	449	1	6			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	MSE	-	EXPRESSION TAG	UNP Q9K7B8
A	40	SER	-	EXPRESSION TAG	UNP Q9K7B8
A	41	LEU	-	EXPRESSION TAG	UNP Q9K7B8
A	336	GLU	-	EXPRESSION TAG	UNP Q9K7B8
A	337	GLY	-	EXPRESSION TAG	UNP Q9K7B8
A	338	HIS	-	EXPRESSION TAG	UNP Q9K7B8
A	339	HIS	-	EXPRESSION TAG	UNP Q9K7B8
A	340	HIS	-	EXPRESSION TAG	UNP Q9K7B8
A	341	HIS	-	EXPRESSION TAG	UNP Q9K7B8
A	342	HIS	-	EXPRESSION TAG	UNP Q9K7B8
A	343	HIS	-	EXPRESSION TAG	UNP Q9K7B8
B	39	MSE	-	EXPRESSION TAG	UNP Q9K7B8
B	40	SER	-	EXPRESSION TAG	UNP Q9K7B8
B	41	LEU	-	EXPRESSION TAG	UNP Q9K7B8
B	336	GLU	-	EXPRESSION TAG	UNP Q9K7B8
B	337	GLY	-	EXPRESSION TAG	UNP Q9K7B8
B	338	HIS	-	EXPRESSION TAG	UNP Q9K7B8
B	339	HIS	-	EXPRESSION TAG	UNP Q9K7B8
B	340	HIS	-	EXPRESSION TAG	UNP Q9K7B8
B	341	HIS	-	EXPRESSION TAG	UNP Q9K7B8
B	342	HIS	-	EXPRESSION TAG	UNP Q9K7B8
B	343	HIS	-	EXPRESSION TAG	UNP Q9K7B8

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	130	Total 130	O 130	0	0
2	B	99	Total 99	O 99	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	60.46Å 87.86Å 113.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.92 – 2.02 33.92 – 2.02	Depositor EDS
% Data completeness (in resolution range)	100.0 (33.92-2.02) 87.2 (33.92-2.02)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.96 (at 2.03Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.193 , 0.237 0.196 , 0.240	Depositor DCC
R_{free} test set	1760 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	27.4	Xtriage
Anisotropy	0.240	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 49.1	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	0.94	EDS
Total number of atoms	4699	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.94% 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

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	1.01	1/2284 (0.0%)	0.92	3/3102 (0.1%)
1	B	0.93	1/2263 (0.0%)	0.87	5/3073 (0.2%)
All	All	0.97	2/4547 (0.0%)	0.90	8/6175 (0.1%)

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	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	45	TYR	CD1-CE1	6.76	1.49	1.39
1	A	64	PHE	CB-CG	5.84	1.61	1.51

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	113	LEU	CB-CG-CD2	-8.76	96.10	111.00
1	A	113	LEU	CA-CB-CG	-8.44	95.89	115.30
1	A	289	LEU	CA-CB-CG	-7.66	97.68	115.30
1	B	289	LEU	CA-CB-CG	-6.07	101.35	115.30
1	B	59	ARG	NE-CZ-NH2	5.98	123.29	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	250	ALA	Peptide

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	2245	0	2155	29	1
1	B	2225	0	2116	35	1
2	A	130	0	0	2	0
2	B	99	0	0	1	0
All	All	4699	0	4271	61	1

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 7.

The worst 5 of 61 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:304:ILE:HG22	1:B:305:LEU:H	1.47	0.79
1:A:176:GLN:O	1:A:180:GLN:HG3	1.86	0.74
1:B:211:HIS:O	1:B:215:VAL:HG23	1.89	0.72
1:B:302:PRO:HB2	1:B:310:GLU:HB3	1.72	0.71
1:B:153:ASN:HD22	1:B:321:ILE:HG22	1.55	0.70

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:327:GLU:OE2	1:B:262:LYS:NZ[2_665]	1.57	0.63

5.3 Torsion angles

5.3.1 Protein backbone

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	289/305 (95%)	280 (97%)	9 (3%)	0	100	100
1	B	286/305 (94%)	273 (96%)	10 (4%)	3 (1%)	18	10
All	All	575/610 (94%)	553 (96%)	19 (3%)	3 (0%)	32	25

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	250	ALA
1	B	251	GLY
1	B	141	SER

5.3.2 Protein sidechains ⓘ

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	235/240 (98%)	227 (97%)	8 (3%)	42	39
1	B	231/240 (96%)	216 (94%)	15 (6%)	20	13
All	All	466/480 (97%)	443 (95%)	23 (5%)	29	23

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	69	GLN
1	B	176	GLN
1	B	289	LEU
1	B	129	LEU
1	B	191	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	42	ASN
1	B	153	ASN
1	B	218	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	306	GLN
1	B	211	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	285/305 (93%)	-0.13	1 (0%) 92 92	18, 28, 39, 50	0
1	B	284/305 (93%)	0.25	9 (3%) 48 49	17, 31, 48, 64	0
All	All	569/610 (93%)	0.06	10 (1%) 69 68	17, 29, 46, 64	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	192	ALA	4.3
1	A	331	TYR	3.7
1	B	262	LYS	3.6
1	B	195	PRO	3.4
1	B	309	GLU	3.2

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.