



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 18, 2024 – 04:32 PM JST

PDB ID : 9J6M
Title : beta-D-galactofuranosidase ORF1110 from *Streptomyces* sp. JHA19, ligand-free form
Authors : Fujio, N.; Yamada, C.; Takegawa, K.; Fushinobu, S.
Deposited on : 2024-08-16
Resolution : 1.80 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 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 : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.003 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

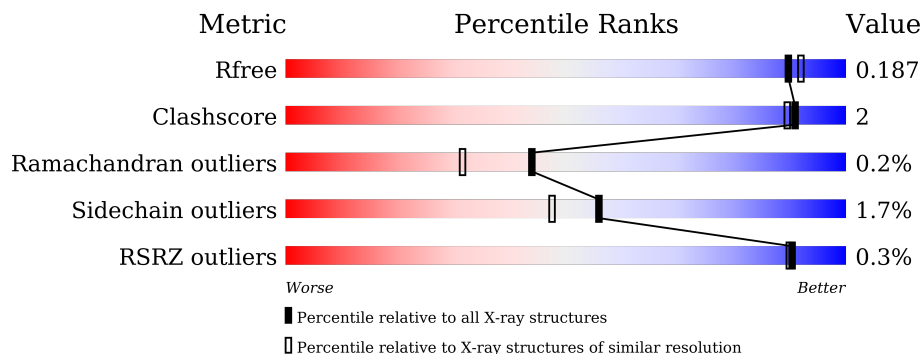
1 Overall quality at a glance

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))
R_{free}	164625	7108 (1.80-1.80)
Clashscore	180529	8162 (1.80-1.80)
Ramachandran outliers	177936	8077 (1.80-1.80)
Sidechain outliers	177891	8076 (1.80-1.80)
RSRZ outliers	164620	7108 (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	612	 89% 5% • 5%
1	B	612	 87% 7% • 6%

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 9832 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 Beta-D-galactofuranosidase 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	581	4507	2838	814	846	2	7	0	0	0
1	B	576	4482	2823	809	841	2	7	0	1	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	MSE	-	initiating methionine	UNP A0A0K2SRV3
A	25	GLY	-	expression tag	UNP A0A0K2SRV3
A	26	SER	-	expression tag	UNP A0A0K2SRV3
A	27	SER	-	expression tag	UNP A0A0K2SRV3
A	28	HIS	-	expression tag	UNP A0A0K2SRV3
A	29	HIS	-	expression tag	UNP A0A0K2SRV3
A	30	HIS	-	expression tag	UNP A0A0K2SRV3
A	31	HIS	-	expression tag	UNP A0A0K2SRV3
A	32	HIS	-	expression tag	UNP A0A0K2SRV3
A	33	HIS	-	expression tag	UNP A0A0K2SRV3
A	34	SER	-	expression tag	UNP A0A0K2SRV3
A	35	ALA	-	expression tag	UNP A0A0K2SRV3
A	36	ALA	-	expression tag	UNP A0A0K2SRV3
A	37	LEU	-	expression tag	UNP A0A0K2SRV3
A	38	GLU	-	expression tag	UNP A0A0K2SRV3
A	39	VAL	-	expression tag	UNP A0A0K2SRV3
A	40	LEU	-	expression tag	UNP A0A0K2SRV3
A	41	PHE	-	expression tag	UNP A0A0K2SRV3
A	42	GLN	-	expression tag	UNP A0A0K2SRV3
A	43	GLY	-	expression tag	UNP A0A0K2SRV3
A	44	PRO	-	expression tag	UNP A0A0K2SRV3
A	585	HIS	TYR	engineered mutation	UNP A0A0K2SRV3
B	24	MSE	-	initiating methionine	UNP A0A0K2SRV3
B	25	GLY	-	expression tag	UNP A0A0K2SRV3
B	26	SER	-	expression tag	UNP A0A0K2SRV3

Continued on next page...

Continued from previous page...

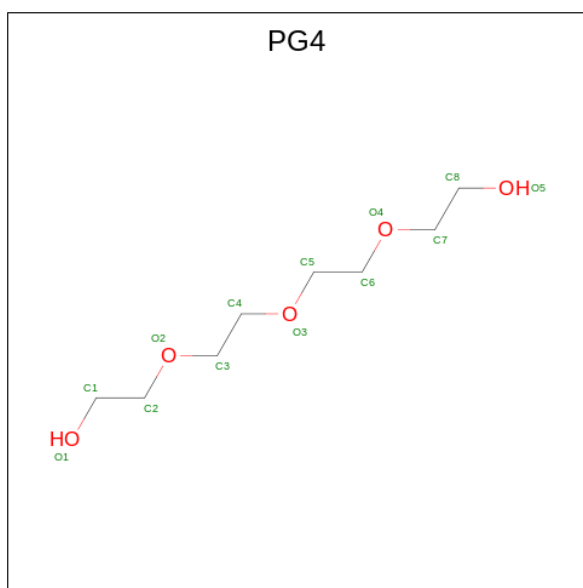
Chain	Residue	Modelled	Actual	Comment	Reference
B	27	SER	-	expression tag	UNP A0A0K2SRV3
B	28	HIS	-	expression tag	UNP A0A0K2SRV3
B	29	HIS	-	expression tag	UNP A0A0K2SRV3
B	30	HIS	-	expression tag	UNP A0A0K2SRV3
B	31	HIS	-	expression tag	UNP A0A0K2SRV3
B	32	HIS	-	expression tag	UNP A0A0K2SRV3
B	33	HIS	-	expression tag	UNP A0A0K2SRV3
B	34	SER	-	expression tag	UNP A0A0K2SRV3
B	35	ALA	-	expression tag	UNP A0A0K2SRV3
B	36	ALA	-	expression tag	UNP A0A0K2SRV3
B	37	LEU	-	expression tag	UNP A0A0K2SRV3
B	38	GLU	-	expression tag	UNP A0A0K2SRV3
B	39	VAL	-	expression tag	UNP A0A0K2SRV3
B	40	LEU	-	expression tag	UNP A0A0K2SRV3
B	41	PHE	-	expression tag	UNP A0A0K2SRV3
B	42	GLN	-	expression tag	UNP A0A0K2SRV3
B	43	GLY	-	expression tag	UNP A0A0K2SRV3
B	44	PRO	-	expression tag	UNP A0A0K2SRV3
B	585	HIS	TYR	engineered mutation	UNP A0A0K2SRV3

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula: C₈H₁₈O₅).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 13 8 5	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Mg 1 1	0	0
4	B	1	Total Mg 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	439	Total O 439 439	0	0
5	B	377	Total O 377 377	0	0

4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	84.55Å 57.35Å 110.48Å 90.00° 92.05° 90.00°	Depositor
Resolution (Å)	47.50 – 1.80 47.50 – 1.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.50-1.80) 100.0 (47.50-1.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.75 (at 1.79Å)	Xtrriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.143 , 0.177 0.158 , 0.187	Depositor DCC
R_{free} test set	4889 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å ²)	14.5	Xtrriage
Anisotropy	0.039	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 39.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.013 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	9832	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.25% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PG4, MG

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	0/4627	0.93	12/6325 (0.2%)
1	B	0.47	1/4604 (0.0%)	0.92	13/6290 (0.2%)
All	All	0.47	1/9231 (0.0%)	0.92	25/12615 (0.2%)

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	4
1	B	0	3
All	All	0	7

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	610	GLU	CD-OE2	5.08	1.31	1.25

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	599	LEU	CB-CG-CD2	9.25	126.73	111.00
1	B	473	ARG	NE-CZ-NH1	8.62	124.61	120.30
1	A	614	ARG	NE-CZ-NH1	8.31	124.45	120.30
1	B	473	ARG	NE-CZ-NH2	-7.47	116.57	120.30
1	B	473	ARG	CD-NE-CZ	6.97	133.36	123.60

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	110	ARG	Sidechain
1	A	167	ARG	Sidechain
1	A	572	ARG	Sidechain
1	A	618	ARG	Sidechain
1	B	124	ARG	Sidechain

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	4507	0	4349	12	0
1	B	4482	0	4322	16	0
2	A	6	0	8	1	0
2	B	6	0	8	1	0
3	A	13	0	18	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	439	0	0	0	0
5	B	377	0	0	2	0
All	All	9832	0	8705	28	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.

The worst 5 of 28 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:487:VAL:CG2	5:B:983:HOH:O	2.42	0.67
1:B:495:CYS:HB3	1:B:496:CYS:SG	2.41	0.60
1:A:618:ARG:HH21	1:A:618:ARG:CG	2.19	0.55
1:B:576:MSE:HB3	1:B:577:PRO:HD3	1.93	0.50
1:A:564:PHE:HE1	1:A:585:HIS:ND1	2.10	0.50

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	579/612 (95%)	565 (98%)	13 (2%)	1 (0%)	44	31
1	B	573/612 (94%)	559 (98%)	13 (2%)	1 (0%)	44	31
All	All	1152/1224 (94%)	1124 (98%)	26 (2%)	2 (0%)	44	31

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	374	ILE
1	B	374	ILE

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	473/490 (96%)	465 (98%)	8 (2%)	56	47
1	B	471/490 (96%)	463 (98%)	8 (2%)	56	47
All	All	944/980 (96%)	928 (98%)	16 (2%)	56	47

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	443	TYR
1	B	375	TYR
1	B	46	PRO
1	B	339	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	599	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	63	GLN
1	B	98	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	GOL	B	701	-	5,5,5	0.17	0	5,5,5	0.94	0
3	PG4	A	702	-	12,12,12	0.26	0	11,11,11	0.38	0
2	GOL	A	701	-	5,5,5	0.27	0	5,5,5	1.08	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	B	701	-	-	4/4/4/4	-
3	PG4	A	702	-	-	2/10/10/10	-
2	GOL	A	701	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	701	GOL	O1-C1-C2-C3
2	A	701	GOL	C1-C2-C3-O3
2	B	701	GOL	O1-C1-C2-C3
2	B	701	GOL	C1-C2-C3-O3
2	B	701	GOL	O1-C1-C2-O2

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	701	GOL	1	0
3	A	702	PG4	1	0
2	A	701	GOL	1	0

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	574/612 (93%)	-0.64	1 (0%) 92 91	9, 13, 23, 34	0
1	B	569/612 (92%)	-0.51	3 (0%) 87 87	9, 15, 27, 54	1 (0%)
All	All	1143/1224 (93%)	-0.57	4 (0%) 90 90	9, 14, 25, 54	1 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	46	PRO	3.5
1	B	190	ALA	2.3
1	B	273	GLY	2.2
1	A	273	GLY	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
3	PG4	A	702	13/13	0.86	0.14	35,38,39,39	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	GOL	B	701	6/6	0.91	0.10	20,24,27,32	0
2	GOL	A	701	6/6	0.94	0.08	18,22,24,26	0
4	MG	A	703	1/1	0.99	0.02	10,10,10,10	0
4	MG	B	702	1/1	1.00	0.03	10,10,10,10	0

6.5 Other polymers [i](#)

There are no such residues in this entry.