



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 09:56 PM UTC

PDB ID : 4ZOC / pdb_00004zoc
Title : Crystal Structure of mutant (D270A) beta-glucosidase from *Listeria innocua* in complex with sophorotriose
Authors : Nakajima, M.; Yoshida, R.; Miyanaga, A.; Abe, K.; Takahashi, Y.; Sugimoto, N.; Toyozumi, H.; Nakai, H.; Kitaoka, M.; Taguchi, H.
Deposited on : 2015-05-06
Resolution : 1.79 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

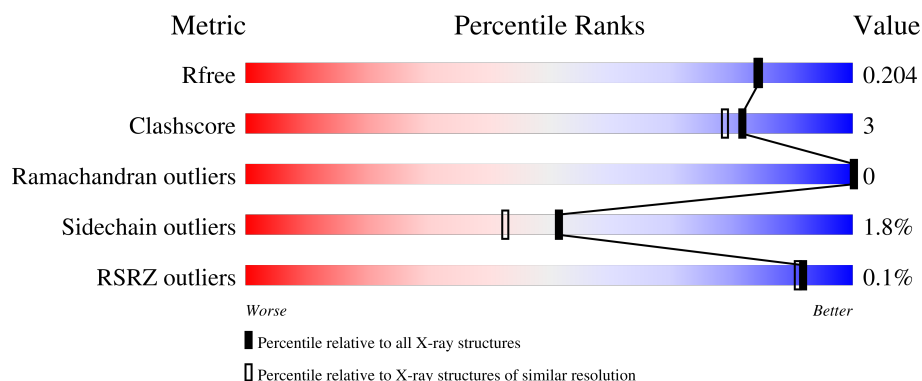
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.79 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}	180053	7662 (1.80-1.80)
Clashscore	190562	8479 (1.80-1.80)
Ramachandran outliers	187476	8391 (1.80-1.80)
Sidechain outliers	187428	8390 (1.80-1.80)
RSRZ outliers	180081	7663 (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	731	 91% 8% ..
1	B	731	 91% 7% .

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GOL	A	801[A]	-	X	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12626 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 Lin1840 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	724	Total	C	N	O	S	0	7	0
			5651	3575	936	1114	26			
1	B	724	Total	C	N	O	S	0	5	0
			5632	3564	933	1108	27			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLU	LYS	engineered mutation	UNP Q92AS9
A	270	ALA	ASP	engineered mutation	UNP Q92AS9
A	724	LEU	-	expression tag	UNP Q92AS9
A	725	GLU	-	expression tag	UNP Q92AS9
A	726	HIS	-	expression tag	UNP Q92AS9
A	727	HIS	-	expression tag	UNP Q92AS9
A	728	HIS	-	expression tag	UNP Q92AS9
A	729	HIS	-	expression tag	UNP Q92AS9
A	730	HIS	-	expression tag	UNP Q92AS9
A	731	HIS	-	expression tag	UNP Q92AS9
B	2	GLU	LYS	engineered mutation	UNP Q92AS9
B	270	ALA	ASP	engineered mutation	UNP Q92AS9
B	724	LEU	-	expression tag	UNP Q92AS9
B	725	GLU	-	expression tag	UNP Q92AS9
B	726	HIS	-	expression tag	UNP Q92AS9
B	727	HIS	-	expression tag	UNP Q92AS9
B	728	HIS	-	expression tag	UNP Q92AS9
B	729	HIS	-	expression tag	UNP Q92AS9
B	730	HIS	-	expression tag	UNP Q92AS9
B	731	HIS	-	expression tag	UNP Q92AS9

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	1
			12	6	6		
2	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

- Molecule 4 is beta-D-glucopyranose (CCD ID: BGC) (formula: C₆H₁₂O₆).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			12	6	6		
4	B	1	Total	C	O	0	0
			12	6	6		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	696	Total	O	0	0
			696	696		
5	B	603	Total	O	0	0
			603	603		

- Molecule 1: Lin1840 protein



4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	89.62Å 94.86Å 215.06Å 90.00° 96.33° 90.00°	Depositor
Resolution (Å)	32.50 – 1.79 32.50 – 1.79	Depositor EDS
% Data completeness (in resolution range)	98.9 (32.50-1.79) 98.9 (32.50-1.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.00 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.141 , 0.196 0.154 , 0.204	Depositor DCC
R_{free} test set	8390 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	13.3	Xtriage
Anisotropy	0.154	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 42.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	12626	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.84% 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: BGC, GOL, 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	1.30	11/5753 (0.2%)	1.12	13/7794 (0.2%)
1	B	1.27	9/5734 (0.2%)	1.12	9/7769 (0.1%)
All	All	1.29	20/11487 (0.2%)	1.12	22/15563 (0.1%)

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	720	GLY	C-O	-9.40	1.13	1.23
1	A	368	ARG	CD-NE	-6.42	1.37	1.46
1	B	655	ARG	CD-NE	-6.17	1.37	1.46
1	B	540	LEU	C-O	-5.72	1.17	1.24
1	B	480	ILE	C-O	-5.67	1.17	1.24

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	655	ARG	NE-CZ-NH2	12.13	130.12	119.20
1	B	655	ARG	NE-CZ-NH1	-8.99	112.51	121.50
1	A	40	THR	N-CA-C	8.74	120.50	110.97
1	A	723	VAL	CA-C-N	7.62	135.42	121.70
1	A	723	VAL	C-N-CA	7.62	135.42	121.70

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	5651	0	5597	28	0
1	B	5632	0	5583	27	0
2	A	18	0	24	9	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	12	0	12	0	0
4	B	12	0	12	0	0
5	A	696	0	0	8	0
5	B	603	0	0	2	0
All	All	12626	0	11228	57	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 3.

The worst 5 of 57 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)
2:A:801[A]:GOL:H32	1:B:207:THR:HG21	1.59	0.83
2:A:801[A]:GOL:C3	1:B:207:THR:HG21	2.12	0.78
1:B:75:ALA:HA	1:B:78:LYS:HD3	1.69	0.74
1:A:683[B]:GLU:OE2	1:A:694:LYS:NZ	2.21	0.73
1:A:658:LYS:HE2	1:A:709:SER:O	1.96	0.66

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	729/731 (100%)	712 (98%)	17 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	727/731 (100%)	709 (98%)	18 (2%)	0	100	100
All	All	1456/1462 (100%)	1421 (98%)	35 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	611/611 (100%)	602 (98%)	9 (2%)	57	49
1	B	609/611 (100%)	596 (98%)	13 (2%)	47	36
All	All	1220/1222 (100%)	1198 (98%)	22 (2%)	51	43

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	333	LEU
1	B	610	LYS
1	B	383	ARG
1	B	613	SER
1	A	616	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	182	ASN
1	B	567	HIS
1	A	185	GLN
1	A	351	ASN
1	A	455	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	A	801[A]	-	5,5,5	0.59	0	5,5,5	1.48	2 (40%)
2	GOL	A	804	-	5,5,5	0.29	0	5,5,5	1.17	0
4	BGC	B	801	-	12,12,12	1.30	2 (16%)	17,17,17	1.44	2 (11%)
4	BGC	A	803	-	12,12,12	0.89	0	17,17,17	1.09	2 (11%)
2	GOL	A	801[B]	-	5,5,5	0.44	0	5,5,5	1.71	1 (20%)

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	A	801[A]	-	-	4/4/4/4	-
2	GOL	A	804	-	-	0/4/4/4	-
4	BGC	B	801	-	-	0/2/22/22	0/1/1/1
4	BGC	A	803	-	-	0/2/22/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	A	801[B]	-	-	3/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	801	BGC	O4-C4	2.12	1.48	1.43
4	B	801	BGC	C4-C3	2.08	1.57	1.52

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	801	BGC	O5-C5-C6	3.73	115.69	106.44
2	A	801[B]	GOL	O2-C2-C1	-2.66	98.16	109.18
4	B	801	BGC	C3-C4-C5	-2.14	106.35	110.23
4	A	803	BGC	O2-C2-C3	2.08	115.27	110.38
2	A	801[A]	GOL	O2-C2-C1	2.06	117.72	109.18

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	801[A]	GOL	C1-C2-C3-O3
2	A	801[A]	GOL	O2-C2-C3-O3
2	A	801[B]	GOL	O1-C1-C2-C3
2	A	801[B]	GOL	O1-C1-C2-O2
2	A	801[A]	GOL	O1-C1-C2-C3

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801[A]	GOL	7	0
2	A	801[B]	GOL	2	0

5.7 Other polymers [i](#)

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	724/731 (99%)	-0.61	1 (0%)	92 91	4, 11, 25, 46	7 (0%)
1	B	724/731 (99%)	-0.51	1 (0%)	92 91	5, 14, 30, 51	5 (0%)
All	All	1448/1462 (99%)	-0.56	2 (0%)	92 91	4, 12, 28, 51	12 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	724	LEU	2.4
1	B	724	LEU	2.2

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 oligosaccharides 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	GOL	A	804	6/6	0.88	0.12	29,30,34,45	0
4	BGC	B	801	12/12	0.95	0.07	11,15,19,21	0
2	GOL	A	801[A]	6/6	0.97	0.09	6,8,8,9	6

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
4	BGC	A	803	12/12	0.97	0.05	14,19,21,24	0
2	GOL	A	801[B]	6/6	0.97	0.09	12,19,21,25	6
3	MG	A	802	1/1	0.99	0.01	11,11,11,11	0
3	MG	B	802	1/1	1.00	0.01	11,11,11,11	0

6.5 Other polymers [i](#)

There are no such residues in this entry.