



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2026 – 08:44 AM UTC

PDB ID : 3ZYZ / pdb_00003zyz
Title : Crystal structure of a glycoside hydrolase family 3 beta-glucosidase, Bgl1 from Hypocrea jecorina at 2.1Å resolution.
Authors : Sandgren, M.; Kaper, T.; Mikkelsen, N.E.; Hansson, H.; Piens, K.; Gudmundsson, M.; Larenas, E.; Kelemen, B.; Karkehabadi, S.
Deposited on : 2011-08-30
Resolution : 2.10 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 : 4-5-2 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

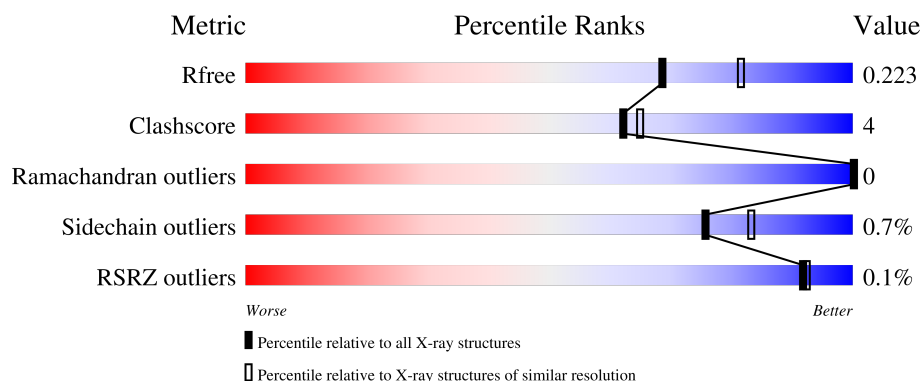
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


The reported resolution of this entry is 2.10 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	6658 (2.10-2.10)
Clashscore	190562	7164 (2.10-2.10)
Ramachandran outliers	187476	7099 (2.10-2.10)
Sidechain outliers	187428	7100 (2.10-2.10)
RSRZ outliers	180081	6662 (2.10-2.10)

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	713	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6040 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-GLUCOSIDE GLUCOHYDROLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	713	5370	3370	935	1052	13	0	12	0

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	14	8	1	5	0	0
2	A	1	14	8	1	5	0	0

- Molecule 3 is beta-D-glucopyranose (CCD ID: BGC) (formula: $C_6H_{12}O_6$).



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

- Molecule 4 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 5 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		
5	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	610	Total	O	0	0
			610	610		

- Molecule 1: BETA-D-GLUCOSIDE GLUCOHYDROLASE

V1	V2	C42	M45	V101	I106	G115	P116	P121	Q122	R125	N146	D177	D178	A198	K206	W211	Q216	Q226	T235	D236	Q240	A295	R328	L333	N336	L341	H361	G379	E441	E447	G448	N449	H450
N463	Q467	I477	V483	I487	A489	L505	P506	S510	V522	K527	P536	G576	V603	P604	V617	T618	V619	V632	Y639	P644	R645	T646	P647	S664	I671	D675	Q684	V697	L710	S711	V712	A713	

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	55.06Å 82.29Å 136.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.10 30.00 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-2.10) 99.6 (30.00-2.10)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.04 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.172 , 0.222 0.174 , 0.223	Depositor DCC
R_{free} test set	1844 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	15.5	Xtriage
Anisotropy	0.106	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 41.2	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	0.94	EDS
Total number of atoms	6040	wwPDB-VP
Average B, all atoms (Å ²)	14.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% 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, BGC, PEG, NAG

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	0.59	1/5523 (0.0%)	0.81	4/7547 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	647	PRO	CA-C	5.56	1.54	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	646[A]	THR	CA-C-N	-5.80	115.42	119.66
1	A	646[A]	THR	C-N-CA	-5.80	115.42	119.66
1	A	646[B]	THR	CA-C-N	-5.80	115.42	119.66
1	A	646[B]	THR	C-N-CA	-5.80	115.42	119.66

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	5370	0	5265	44	0
2	A	28	0	26	2	0
3	A	12	0	12	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	6	0	8	0	0
5	A	14	0	20	2	0
6	A	610	0	0	2	0
All	All	6040	0	5331	44	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 4.

All (44) 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:449:ASN:HD21	5:A:1719:PEG:H31	1.27	0.98
1:A:122:GLN:HE21	1:A:122:GLN:H	1.21	0.86
1:A:42:CYS:H	1:A:45:ASN:HD21	1.24	0.84
1:A:236:ASP:OD1	3:A:1717:BGC:H1	1.81	0.81
1:A:336:ASN:HD21	1:A:341:LEU:H	1.33	0.75
1:A:639:TYR:HB2	1:A:646[B]:THR:HG21	1.73	0.71
1:A:447:GLU:OE2	1:A:460:HIS:HE1	1.79	0.65
1:A:361:HIS:CE1	1:A:379:GLY:H	2.17	0.63
1:A:646[B]:THR:HG23	6:A:2550:HOH:O	1.99	0.62
1:A:527:LYS:HD2	1:A:632:VAL:HB	1.82	0.61
1:A:122:GLN:H	1:A:122:GLN:NE2	1.97	0.60
1:A:122:GLN:HE21	1:A:122:GLN:N	1.99	0.56
1:A:115:GLY:HA2	1:A:116:PRO:C	2.31	0.55
1:A:487:ILE:HG13	1:A:536:PRO:HD3	1.92	0.51
1:A:463:ASN:O	1:A:467:GLN:HG3	2.11	0.50
1:A:42:CYS:H	1:A:45:ASN:ND2	2.04	0.49
1:A:146[B]:ASN:ND2	5:A:1720:PEG:O2	2.45	0.48
1:A:206:LYS:HE3	1:A:211:TRP:CE2	2.48	0.48
1:A:527:LYS:HG2	1:A:576:GLY:HA3	1.96	0.48
1:A:617:VAL:HG23	1:A:671:ILE:HD11	1.95	0.47
1:A:483:VAL:HA	1:A:505:LEU:HD11	1.96	0.47
1:A:178:ASP:HB2	2:A:1716:NAG:H61	1.96	0.47
1:A:361:HIS:HE1	1:A:379:GLY:H	1.62	0.47
1:A:619[B]:VAL:HG21	1:A:697:VAL:HG11	1.97	0.46
1:A:447:GLU:OE2	1:A:460:HIS:CE1	2.66	0.46
1:A:235:THR:HG21	1:A:240:GLN:HB2	1.97	0.45
1:A:226:GLN:HE22	1:A:684:GLN:HE21	1.64	0.45
1:A:328:ARG:HD2	1:A:522:VAL:HG21	1.98	0.45
1:A:477:ILE:HG12	1:A:499:ALA:HB3	1.98	0.45
1:A:101:VAL:HG13	1:A:106:ILE:HB	2.00	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:506:PRO:HB2	1:A:510:SER:HA	2.00	0.44
1:A:226:GLN:HE22	1:A:684:GLN:NE2	2.16	0.44
1:A:198:ALA:HB1	1:A:295:ALA:HB1	1.99	0.43
1:A:603:VAL:HB	1:A:604:PRO:CD	2.48	0.43
1:A:177:ASP:HA	2:A:1716:NAG:H4	2.01	0.43
1:A:125:ARG:HD2	1:A:441:GLU:HG3	2.00	0.42
1:A:644:PRO:O	1:A:645:ARG:C	2.62	0.42
1:A:460:HIS:HD2	6:A:2397:HOH:O	2.03	0.41
1:A:121:PRO:HG2	1:A:122:GLN:NE2	2.36	0.41
1:A:333:LEU:HD21	1:A:336:ASN:HB2	2.02	0.40
1:A:216[B]:GLN:HE21	1:A:216[B]:GLN:HB3	1.58	0.40
1:A:603:VAL:HB	1:A:604:PRO:HD2	2.03	0.40
1:A:671:ILE:HG23	1:A:675:ASP:HB2	2.03	0.40
1:A:710:LEU:HD23	1:A:710:LEU:C	2.46	0.40

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	723/713 (101%)	696 (96%)	27 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	580/568 (102%)	576 (99%)	4 (1%)	76 83

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	VAL
1	A	122	GLN
1	A	664	SER
1	A	711	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	45	ASN
1	A	95	GLN
1	A	122	GLN
1	A	163	ASN
1	A	264	ASN
1	A	278	ASN
1	A	279	GLN
1	A	316	GLN
1	A	336	ASN
1	A	361	HIS
1	A	369	ASN
1	A	388	ASN
1	A	460	HIS
1	A	467	GLN
1	A	475	ASN
1	A	658	ASN
1	A	684	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 [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

6 ligands are modelled in this entry.

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	NAG	A	1715	1	14,14,15	0.52	0	17,19,21	1.00	1 (5%)
4	GOL	A	1718	-	5,5,5	0.38	0	5,5,5	0.59	0
2	NAG	A	1716	1	14,14,15	0.54	0	17,19,21	1.26	1 (5%)
3	BGC	A	1717	-	12,12,12	0.55	0	17,17,17	0.83	0
5	PEG	A	1720	-	6,6,6	0.44	0	5,5,5	0.28	0
5	PEG	A	1719	-	6,6,6	0.59	0	5,5,5	0.32	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	NAG	A	1715	1	-	0/6/23/26	0/1/1/1
4	GOL	A	1718	-	-	2/4/4/4	-
2	NAG	A	1716	1	-	2/6/23/26	0/1/1/1
3	BGC	A	1717	-	-	0/2/22/22	0/1/1/1
5	PEG	A	1720	-	-	3/4/4/4	-
5	PEG	A	1719	-	-	1/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1716	NAG	C1-O5-C5	3.33	116.65	112.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1715	NAG	C1-O5-C5	2.05	114.93	112.19

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1718	GOL	O1-C1-C2-O2
4	A	1718	GOL	O1-C1-C2-C3
5	A	1719	PEG	O1-C1-C2-O2
5	A	1720	PEG	O1-C1-C2-O2
5	A	1720	PEG	O2-C3-C4-O4
5	A	1720	PEG	C1-C2-O2-C3
2	A	1716	NAG	O5-C5-C6-O6
2	A	1716	NAG	C4-C5-C6-O6

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1716	NAG	2	0
3	A	1717	BGC	1	0
5	A	1720	PEG	1	0
5	A	1719	PEG	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	713/713 (100%)	-0.46	1 (0%) 92 92	5, 13, 20, 31	13 (1%)

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	VAL	3.5

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	NAG	A	1716	14/15	0.69	0.14	34,39,43,44	0
5	PEG	A	1719	7/7	0.75	0.17	27,29,30,31	0
5	PEG	A	1720	7/7	0.86	0.11	41,42,42,42	0
4	GOL	A	1718	6/6	0.93	0.07	14,15,15,17	0
2	NAG	A	1715	14/15	0.93	0.07	20,23,27,28	0
3	BGC	A	1717	12/12	0.93	0.07	10,11,14,17	0

6.5 Other polymers [i](#)

There are no such residues in this entry.