



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 02:09 PM UTC

PDB ID : 5Z5H / pdb_00005z5h
Title : Crystal structure of a thermostable glycoside hydrolase family 43 {beta}-1,4-xylosidase from *Geobacillus thermoleovorans* IT-08 in complex with D-xylose
Authors : Rohman, A.; van Oosterwijk, N.; Puspaningsih, N.N.T.; Dijkstra, B.W.
Deposited on : 2018-01-18
Resolution : 1.90 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

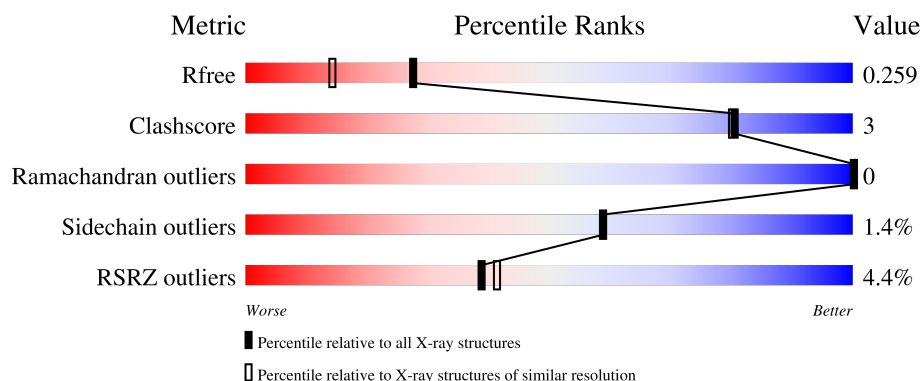
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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	7789 (1.90-1.90)
Clashscore	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)
RSRZ outliers	180081	7790 (1.90-1.90)

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	543	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4603 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-xylosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	504	4061	2629	669	750	13	0	2	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	512	LYS	-	expression tag	UNP Q2I2N4
A	513	GLY	-	expression tag	UNP Q2I2N4
A	514	GLU	-	expression tag	UNP Q2I2N4
A	515	LEU	-	expression tag	UNP Q2I2N4
A	516	ASN	-	expression tag	UNP Q2I2N4
A	517	SER	-	expression tag	UNP Q2I2N4
A	518	LYS	-	expression tag	UNP Q2I2N4
A	519	LEU	-	expression tag	UNP Q2I2N4
A	520	GLU	-	expression tag	UNP Q2I2N4
A	521	GLY	-	expression tag	UNP Q2I2N4
A	522	LYS	-	expression tag	UNP Q2I2N4
A	523	PRO	-	expression tag	UNP Q2I2N4
A	524	ILE	-	expression tag	UNP Q2I2N4
A	525	PRO	-	expression tag	UNP Q2I2N4
A	526	ASN	-	expression tag	UNP Q2I2N4
A	527	PRO	-	expression tag	UNP Q2I2N4
A	528	LEU	-	expression tag	UNP Q2I2N4
A	529	LEU	-	expression tag	UNP Q2I2N4
A	530	GLY	-	expression tag	UNP Q2I2N4
A	531	LEU	-	expression tag	UNP Q2I2N4
A	532	ASP	-	expression tag	UNP Q2I2N4
A	533	SER	-	expression tag	UNP Q2I2N4
A	534	THR	-	expression tag	UNP Q2I2N4
A	535	ARG	-	expression tag	UNP Q2I2N4
A	536	THR	-	expression tag	UNP Q2I2N4
A	537	GLY	-	expression tag	UNP Q2I2N4
A	538	HIS	-	expression tag	UNP Q2I2N4

Continued on next page...

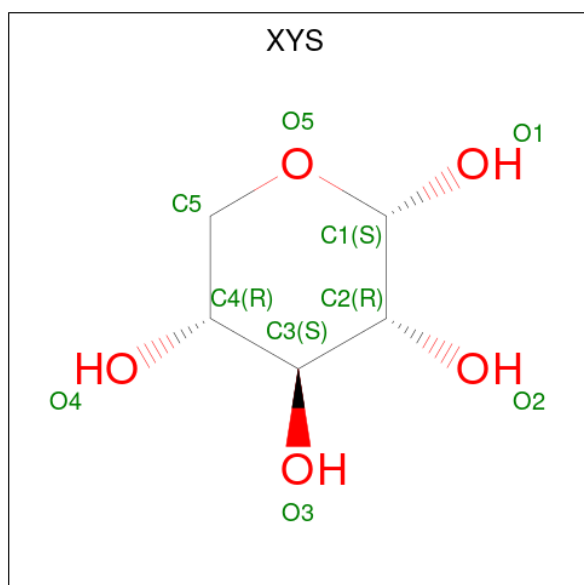
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	539	HIS	-	expression tag	UNP Q2I2N4
A	540	HIS	-	expression tag	UNP Q2I2N4
A	541	HIS	-	expression tag	UNP Q2I2N4
A	542	HIS	-	expression tag	UNP Q2I2N4
A	543	HIS	-	expression tag	UNP Q2I2N4

- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ca 1 1	0	0

- Molecule 3 is alpha-D-xylopyranose (CCD ID: YYS) (formula: C₅H₁₀O₅).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 10 5 5	0	0

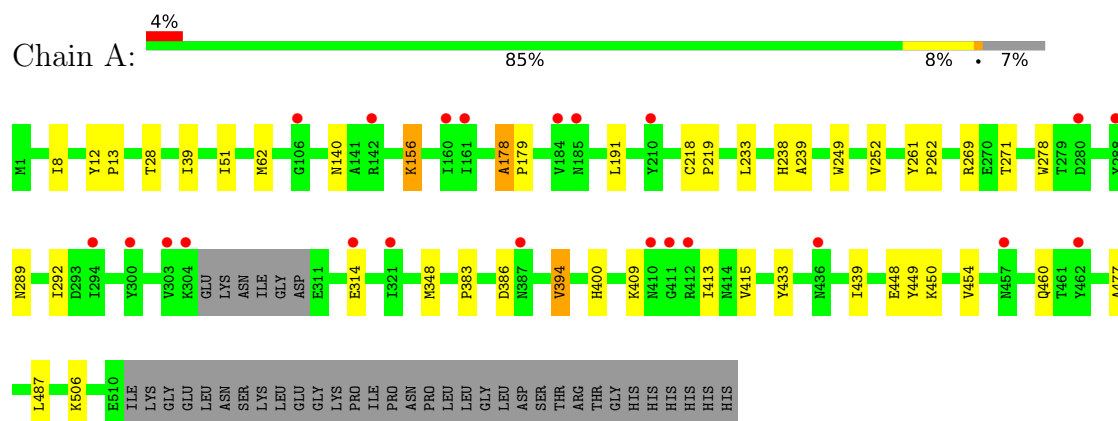
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	531	Total O 531 531	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

• Molecule 1: Beta-xylosidase



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	62.45 Å 62.45 Å 276.67 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.35 – 1.90 46.35 – 1.90	Depositor EDS
% Data completeness (in resolution range)	96.8 (46.35-1.90) 96.9 (46.35-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.98 (at 1.90 Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.216 , 0.259 0.216 , 0.259	Depositor DCC
R_{free} test set	2186 reflections (4.90%)	wwPDB-VP
Wilson B-factor (Å ²)	12.8	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 43.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	4603	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.18% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: XYS, CA

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	0.74	0/4193	0.85	4/5707 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	386	ASP	N-CA-C	6.03	118.65	111.71
1	A	178	ALA	CA-C-N	-5.50	114.13	119.90
1	A	178	ALA	C-N-CA	-5.50	114.13	119.90
1	A	140	ASN	N-CA-C	5.07	116.81	111.28

There are no chirality outliers.

There are no planarity outliers.

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	4061	0	3920	21	0
2	A	1	0	0	0	0
3	A	10	0	10	0	0
4	A	531	0	0	0	0
All	All	4603	0	3930	21	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.

All (21) 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:249:TRP:CG	1:A:278:TRP:HE1	2.27	0.52
1:A:400:HIS:CE1	1:A:477:ALA:HA	2.46	0.51
1:A:314:GLU:HG2	1:A:506:LYS:HG2	1.95	0.48
1:A:448:GLU:OE1	1:A:450:LYS:HE2	2.14	0.48
1:A:415:VAL:HG23	1:A:433:TYR:HB2	1.96	0.47
1:A:233:LEU:HD13	1:A:292:ILE:HD12	1.99	0.44
1:A:239:ALA:HA	1:A:252:VAL:O	2.18	0.44
1:A:439:ILE:HB	1:A:454:VAL:HB	1.99	0.44
1:A:179:PRO:HA	1:A:191:LEU:O	2.18	0.43
1:A:8:ILE:HB	1:A:271:THR:HB	2.00	0.43
1:A:39:ILE:HG23	1:A:51:ILE:HG13	2.00	0.43
1:A:178:ALA:HB2	1:A:238:HIS:CE1	2.54	0.43
1:A:12:TYR:HB3	1:A:28:THR:HG21	2.02	0.42
1:A:13:PRO:HB3	1:A:269:ARG:HB2	2.02	0.42
1:A:383:PRO:HG2	1:A:413:ILE:HD12	2.02	0.42
1:A:218:CYS:HA	1:A:219:PRO:HD3	1.95	0.42
1:A:178:ALA:HB2	1:A:238:HIS:ND1	2.35	0.41
1:A:348:MET:HE1	1:A:487:LEU:HB2	2.01	0.41
1:A:261:TYR:HA	1:A:262:PRO:C	2.45	0.41
1:A:156:LYS:H	1:A:156:LYS:HG2	1.69	0.40
1:A:394:VAL:HG11	1:A:449:TYR:CZ	2.57	0.40

There are no symmetry-related clashes.

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	502/543 (92%)	485 (97%)	17 (3%)	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	440/472 (93%)	434 (99%)	6 (1%)	59 59

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	62	MET
1	A	156	LYS
1	A	289	ASN
1	A	394	VAL
1	A	409	LYS
1	A	460	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	49	ASN
1	A	128	ASN
1	A	368	GLN
1	A	384	ASN
1	A	436	ASN
1	A	457	ASN
1	A	492	ASN

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 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	$\# Z > 2$
3	XYS	A	602	-	10,10,10	0.87	0	14,14,14	1.70	3 (21%)

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
3	XYS	A	602	-	-	-	1/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	602	XYS	C4-C3-C2	-3.69	104.36	110.86
3	A	602	XYS	O2-C2-C1	3.39	117.06	109.25
3	A	602	XYS	O4-C4-C3	-2.26	105.47	110.15

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	602	XYS	C1-C2-C3-C4-C5-O5

No monomer is involved in short contacts.

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

6.1 Protein, DNA and RNA chains

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	504/543 (92%)	0.47	22 (4%)	39 41	10, 17, 32, 47	2 (0%)

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	462	TYR	3.5
1	A	288	TYR	3.1
1	A	411	GLY	3.0
1	A	457	ASN	2.9
1	A	142	ARG	2.7
1	A	300	TYR	2.5
1	A	106	GLY	2.5
1	A	412	ARG	2.4
1	A	410	ASN	2.3
1	A	185	ASN	2.3
1	A	436	ASN	2.3
1	A	303	VAL	2.3
1	A	160	ILE	2.2
1	A	161	ILE	2.2
1	A	314	GLU	2.2
1	A	210	TYR	2.1
1	A	304	LYS	2.1
1	A	321	ILE	2.1
1	A	184	VAL	2.1
1	A	387	ASN	2.1
1	A	294	ILE	2.1
1	A	280	ASP	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

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($\text{\AA}^2$)	Q<0.9
3	XYS	A	602	10/10	0.79	0.15	22,25,25,31	0
2	CA	A	601	1/1	0.98	0.04	18,18,18,18	0

6.5 Other polymers [i](#)

There are no such residues in this entry.