



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 12, 2026 – 07:07 PM UTC

PDB ID : 7KMP / pdb_00007kmp
Title : Crystal structure of the GH31 alpha-xylosidase (Xac1773) from Xanthomonas citri
Authors : Vieira, P.S.; Murakami, M.T.
Deposited on : 2020-11-03
Resolution : 1.56 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

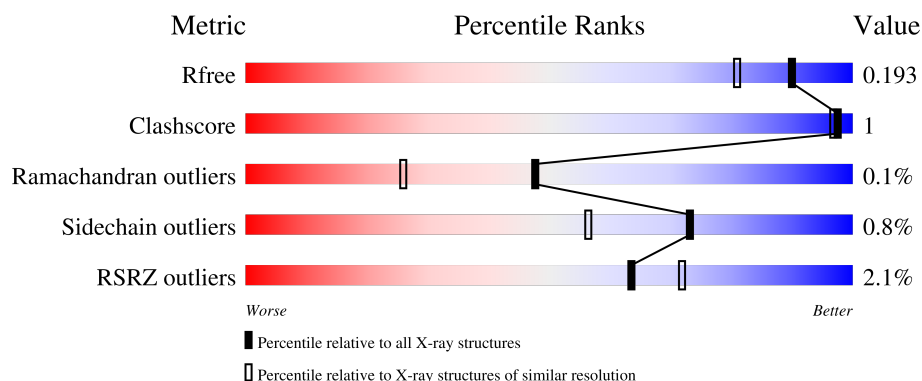
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

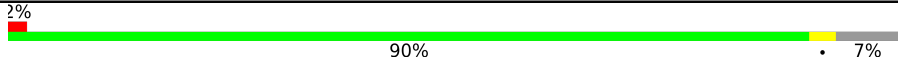
The reported resolution of this entry is 1.56 Å.

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	2145 (1.56-1.56)
Clashscore	190562	2189 (1.56-1.56)
Ramachandran outliers	187476	2153 (1.56-1.56)
Sidechain outliers	187428	2150 (1.56-1.56)
RSRZ outliers	180081	2146 (1.56-1.56)

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	988	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8625 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 Alpha-xylosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	920	Total	C	N	O	S	0	11	0
			7414	4710	1287	1397	20			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	initiating methionine	UNP Q8PLM2
A	-19	GLY	-	expression tag	UNP Q8PLM2
A	-18	SER	-	expression tag	UNP Q8PLM2
A	-17	SER	-	expression tag	UNP Q8PLM2
A	-16	HIS	-	expression tag	UNP Q8PLM2
A	-15	HIS	-	expression tag	UNP Q8PLM2
A	-14	HIS	-	expression tag	UNP Q8PLM2
A	-13	HIS	-	expression tag	UNP Q8PLM2
A	-12	HIS	-	expression tag	UNP Q8PLM2
A	-11	HIS	-	expression tag	UNP Q8PLM2
A	-10	SER	-	expression tag	UNP Q8PLM2
A	-9	SER	-	expression tag	UNP Q8PLM2
A	-8	GLY	-	expression tag	UNP Q8PLM2
A	-7	LEU	-	expression tag	UNP Q8PLM2
A	-6	VAL	-	expression tag	UNP Q8PLM2
A	-5	PRO	-	expression tag	UNP Q8PLM2
A	-4	ARG	-	expression tag	UNP Q8PLM2
A	-3	GLY	-	expression tag	UNP Q8PLM2
A	-2	SER	-	expression tag	UNP Q8PLM2
A	-1	HIS	-	expression tag	UNP Q8PLM2
A	295	VAL	ALA	engineered mutation	UNP Q8PLM2

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



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

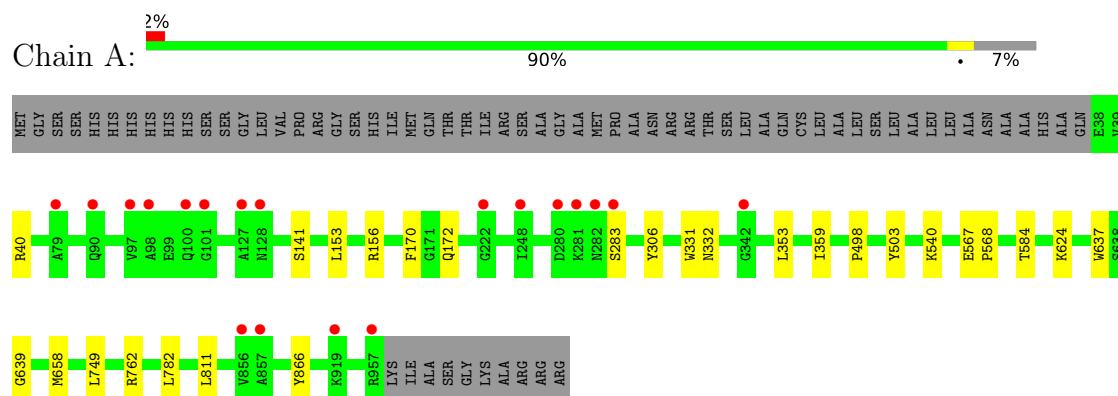
- Molecule 3 is POTASSIUM ION (CCD ID: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	K 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1132	Total 1132	O 1132	0	0

- Molecule 1: Alpha-xylosidase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	117.45Å 146.41Å 63.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.81 – 1.56 45.81 – 1.56	Depositor EDS
% Data completeness (in resolution range)	99.6 (45.81-1.56) 99.5 (45.81-1.56)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.52 (at 1.55Å)	Xtriage
Refinement program	PHENIX dev_3139	Depositor
R, R_{free}	0.162 , 0.185 (Not available) , 0.193	Depositor DCC
R_{free} test set	7779 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	16.9	Xtriage
Anisotropy	0.515	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 43.7	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	0.97	EDS
Total number of atoms	8625	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.71% 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: K, GOL

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	0.31	0/7625	0.51	0/10357

There are no bond length outliers.

There are no bond angle outliers.

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	7414	0	7126	12	0
2	A	78	0	104	2	0
3	A	1	0	0	0	0
4	A	1132	0	0	2	0
All	All	8625	0	7230	12	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 1.

All (12) 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:811:LEU:HD12	1:A:866:TYR:HB3	1.92	0.50
1:A:153:LEU:HG	2:A:1007:GOL:H2	1.93	0.50
1:A:40:ARG:NH2	4:A:1109:HOH:O	2.45	0.49
1:A:141:SER:HB2	1:A:156:ARG:HB3	1.95	0.48
1:A:498:PRO:HA	1:A:503:TYR:CG	2.50	0.46
1:A:749:LEU:HD23	2:A:1005:GOL:H2	1.97	0.46
1:A:331:TRP:CG	1:A:332:ASN:H	2.35	0.45
1:A:762:ARG:HB2	1:A:782:LEU:HB2	1.99	0.45
1:A:306:TYR:O	1:A:359:ILE:HA	2.18	0.43
1:A:567:GLU:N	1:A:568:PRO:HA	2.34	0.41
1:A:637:TRP:CD1	1:A:639:GLY:H	2.39	0.40
1:A:540:LYS:HE3	4:A:1251:HOH:O	2.22	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	930/988 (94%)	918 (99%)	11 (1%)	1 (0%)	48	26

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	283	SER

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	775/814 (95%)	769 (99%)	6 (1%)	73 56

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

Mol	Chain	Res	Type
1	A	170	PHE
1	A	172	GLN
1	A	353	LEU
1	A	584	THR
1	A	624	LYS
1	A	658	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	105	GLN
1	A	332	ASN
1	A	470	HIS
1	A	485	HIS
1	A	682	GLN
1	A	946	GLN

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 14 ligands modelled in this entry, 1 is monoatomic - leaving 13 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	1008	-	5,5,5	0.85	0	5,5,5	1.08	0
2	GOL	A	1005	-	5,5,5	0.86	0	5,5,5	1.10	0
2	GOL	A	1001	-	5,5,5	1.08	0	5,5,5	0.65	0
2	GOL	A	1003	-	5,5,5	0.88	0	5,5,5	0.80	0
2	GOL	A	1004	-	5,5,5	0.75	0	5,5,5	1.04	0
2	GOL	A	1002	-	5,5,5	0.86	0	5,5,5	0.78	0
2	GOL	A	1013	-	5,5,5	1.03	0	5,5,5	1.12	0
2	GOL	A	1011	-	5,5,5	1.15	0	5,5,5	0.91	0
2	GOL	A	1006	-	5,5,5	0.86	0	5,5,5	1.11	0
2	GOL	A	1007	-	5,5,5	0.85	0	5,5,5	1.01	0
2	GOL	A	1010	-	5,5,5	1.05	0	5,5,5	1.12	0
2	GOL	A	1012	-	5,5,5	0.94	0	5,5,5	1.11	1 (20%)
2	GOL	A	1009	-	5,5,5	0.94	0	5,5,5	1.09	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	1008	-	-	0/4/4/4	-
2	GOL	A	1005	-	-	0/4/4/4	-
2	GOL	A	1001	-	-	0/4/4/4	-
2	GOL	A	1003	-	-	2/4/4/4	-
2	GOL	A	1004	-	-	0/4/4/4	-
2	GOL	A	1002	-	-	0/4/4/4	-
2	GOL	A	1013	-	-	1/4/4/4	-
2	GOL	A	1011	-	-	1/4/4/4	-
2	GOL	A	1006	-	-	2/4/4/4	-
2	GOL	A	1007	-	-	2/4/4/4	-
2	GOL	A	1010	-	-	2/4/4/4	-
2	GOL	A	1012	-	-	0/4/4/4	-
2	GOL	A	1009	-	-	2/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1012	GOL	C3-C2-C1	-2.08	104.17	111.80
2	A	1009	GOL	C3-C2-C1	-2.00	104.45	111.80

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1003	GOL	O1-C1-C2-C3
2	A	1007	GOL	C1-C2-C3-O3
2	A	1009	GOL	C1-C2-C3-O3
2	A	1006	GOL	O1-C1-C2-C3
2	A	1010	GOL	C1-C2-C3-O3
2	A	1011	GOL	O1-C1-C2-C3
2	A	1013	GOL	O1-C1-C2-C3
2	A	1003	GOL	O1-C1-C2-O2
2	A	1007	GOL	O2-C2-C3-O3
2	A	1010	GOL	O2-C2-C3-O3
2	A	1006	GOL	O1-C1-C2-O2
2	A	1009	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1005	GOL	1	0
2	A	1007	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 ⓘ

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	920/988 (93%)	0.07	19 (2%) 63 72	6, 18, 33, 70	11 (1%)

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	127	ALA	3.3
1	A	101	GLY	3.2
1	A	97	VAL	3.1
1	A	128	ASN	2.9
1	A	856	VAL	2.8
1	A	98	ALA	2.7
1	A	281	LYS	2.7
1	A	282	ASN	2.6
1	A	283	SER	2.5
1	A	100	GLN	2.5
1	A	280	ASP	2.3
1	A	248	ILE	2.3
1	A	919	LYS	2.2
1	A	79	ALA	2.2
1	A	222	GLY	2.1
1	A	342	GLY	2.1
1	A	957	ARG	2.0
1	A	90	GLN	2.0
1	A	857	ALA	2.0

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
2	GOL	A	1007	6/6	0.77	0.16	29,41,44,47	0
2	GOL	A	1003	6/6	0.89	0.12	28,31,33,39	0
2	GOL	A	1008	6/6	0.89	0.14	22,35,40,46	0
2	GOL	A	1012	6/6	0.89	0.14	21,35,37,45	0
2	GOL	A	1011	6/6	0.90	0.11	28,32,38,46	0
2	GOL	A	1010	6/6	0.90	0.11	28,36,38,39	0
2	GOL	A	1005	6/6	0.91	0.11	20,30,35,37	0
2	GOL	A	1004	6/6	0.91	0.11	20,26,33,37	0
2	GOL	A	1013	6/6	0.92	0.10	24,27,32,32	0
2	GOL	A	1006	6/6	0.93	0.10	27,34,38,42	0
2	GOL	A	1009	6/6	0.94	0.09	21,22,31,32	0
2	GOL	A	1002	6/6	0.97	0.06	14,15,18,22	0
2	GOL	A	1001	6/6	0.98	0.04	14,15,16,16	0
3	K	A	1014	1/1	0.98	0.11	25,25,25,25	0

6.5 Other polymers [i](#)

There are no such residues in this entry.