



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 08:51 PM UTC

PDB ID : 5G56 / pdb_00005g56
Title : THE TETRA-MODULAR CELLULOSOMAL ARABINOXYLANASE
CtXyl5A STRUCTURE AS REVEALED BY X-RAY CRYSTALLOGRA-
PHY
Authors : Bras, J.L.A.; Gilbert, H.J.; Ferreira, L.M.A.; Fontes, C.M.G.A.; Najmudin, S.
Deposited on : 2016-05-21
Resolution : 2.64 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

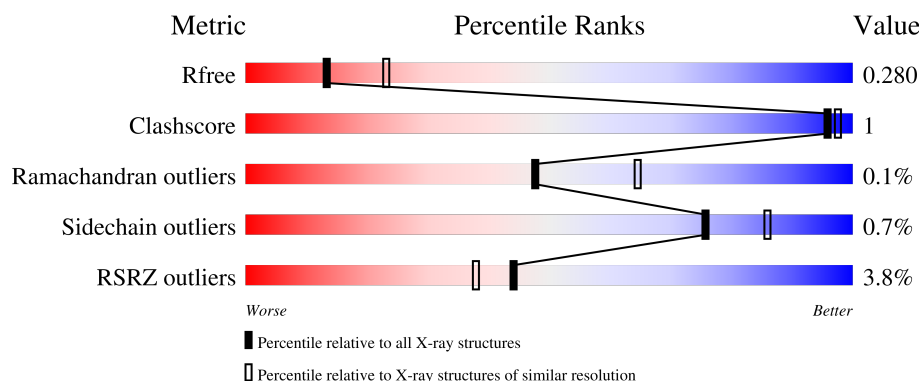
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.64 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	2053 (2.66-2.62)
Clashscore	190562	2097 (2.66-2.62)
Ramachandran outliers	187476	2066 (2.66-2.62)
Sidechain outliers	187428	2066 (2.66-2.62)
RSRZ outliers	180081	2052 (2.66-2.62)

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	862	3% 79% • 18%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5692 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 CARBOHYDRATE BINDING FAMILY 6.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	707	Total	C	N	O	S	Se	0	0	0
			5446	3464	928	1047	3	4			

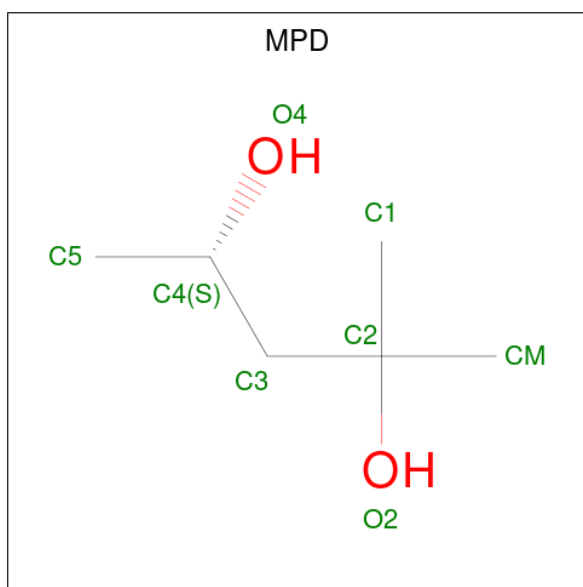
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	890	LEU	-	expression tag	UNP A3DHG6
A	891	GLU	-	expression tag	UNP A3DHG6
A	892	HIS	-	expression tag	UNP A3DHG6
A	893	HIS	-	expression tag	UNP A3DHG6
A	894	HIS	-	expression tag	UNP A3DHG6
A	895	HIS	-	expression tag	UNP A3DHG6
A	896	HIS	-	expression tag	UNP A3DHG6
A	897	HIS	-	expression tag	UNP A3DHG6

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total	Ca	0	0
			3	3		

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (CCD ID: MPD) (formula: C₆H₁₄O₂).

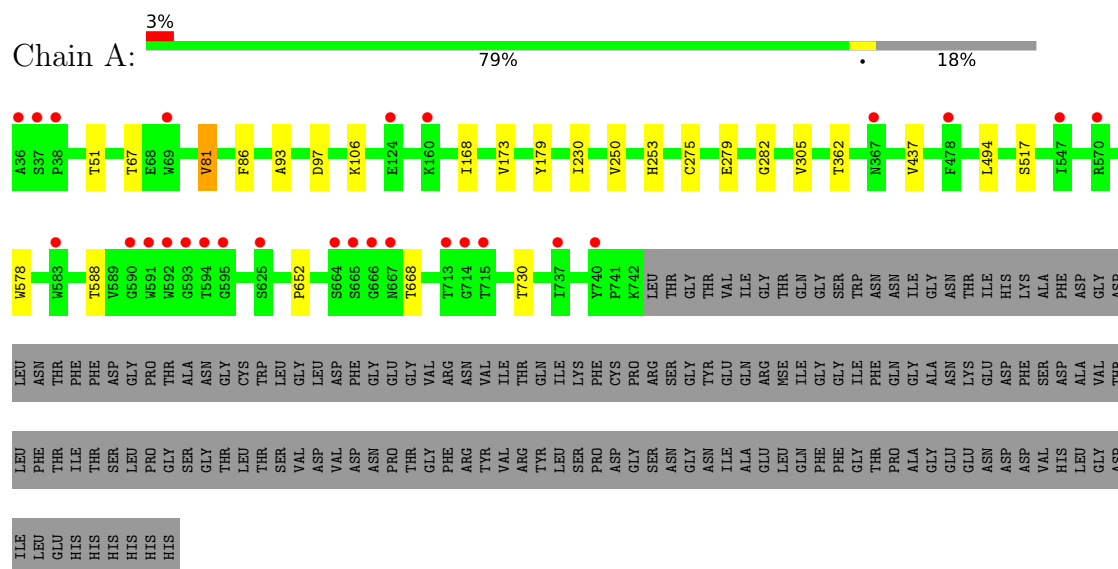


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			8	6	2		
3	A	1	Total	C	O	0	0
			8	6	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	227	Total	O	0	0
			227	227		

• Molecule 1: CARBOHYDRATE BINDING FAMILY 6



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	147.41Å 191.73Å 50.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.70 – 2.64 50.70 – 2.64	Depositor EDS
% Data completeness (in resolution range)	97.7 (50.70-2.64) 97.7 (50.70-2.64)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.96 (at 2.65Å)	Xtriage
Refinement program	REFMAC 5.8.0151	Depositor
R, R_{free}	0.237 , 0.278 0.239 , 0.280	Depositor DCC
R_{free} test set	2125 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	37.8	Xtriage
Anisotropy	0.347	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 39.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	5692	wwPDB-VP
Average B, all atoms (Å ²)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.69% 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: CA, MPD

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	0.57	0/5601	0.77	1/7647 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	168	ILE	N-CA-C	5.21	115.83	110.36

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	5446	0	5163	12	0
2	A	3	0	0	0	0
3	A	16	0	28	0	0
4	A	227	0	0	0	0
All	All	5692	0	5191	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:51:THR:HG22	1:A:275:CYS:O	1.81	0.81
1:A:51:THR:HG21	1:A:305:VAL:HA	1.79	0.65
1:A:51:THR:CG2	1:A:305:VAL:HA	2.32	0.58
1:A:97:ASP:HB2	1:A:106:LYS:HG2	1.87	0.55
1:A:652:PRO:HG2	1:A:730:THR:HG22	1.90	0.53
1:A:578:TRP:HB3	1:A:588:THR:HG22	1.94	0.48
1:A:81:VAL:HG13	1:A:86:PHE:HB2	1.99	0.43
1:A:173:VAL:HB	1:A:179:TYR:CE1	2.54	0.43
1:A:67:THR:HG23	1:A:93:ALA:O	2.19	0.42
1:A:230:ILE:HD11	1:A:250:VAL:HG21	2.02	0.42
1:A:362:THR:HG21	1:A:437:VAL:HG22	2.03	0.41
1:A:253:HIS:CD2	1:A:279:GLU:HB3	2.56	0.41

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	705/862 (82%)	681 (97%)	23 (3%)	1 (0%)	48	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	282	GLY

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	554/676 (82%)	550 (99%)	4 (1%)	76	86

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

Mol	Chain	Res	Type
1	A	81	VAL
1	A	494	LEU
1	A	517	SER
1	A	668	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	39	GLN
1	A	138	ASN
1	A	163	HIS
1	A	240	ASN
1	A	464	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 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 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	MPD	A	1746	-	7,7,7	0.30	0	9,10,10	0.30	0
3	MPD	A	1747	-	7,7,7	0.25	0	9,10,10	0.52	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
3	MPD	A	1746	-	-	2/5/5/5	-
3	MPD	A	1747	-	-	3/5/5/5	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1746	MPD	C1-C2-C3-C4
3	A	1747	MPD	C1-C2-C3-C4
3	A	1747	MPD	C2-C3-C4-C5
3	A	1746	MPD	CM-C2-C3-C4
3	A	1747	MPD	CM-C2-C3-C4

There are no ring outliers.

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	703/862 (81%)	0.06	27 (3%)	44 38	24, 39, 64, 108	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	592	TRP	7.8
1	A	665	SER	6.2
1	A	593	GLY	5.6
1	A	591	TRP	5.0
1	A	666	GLY	4.8
1	A	713	THR	4.3
1	A	594	THR	4.2
1	A	36	ALA	4.2
1	A	667	ASN	3.8
1	A	664	SER	3.6
1	A	590	GLY	3.4
1	A	38	PRO	3.0
1	A	625	SER	2.8
1	A	478	PHE	2.8
1	A	547	ILE	2.6
1	A	715	THR	2.5
1	A	124	GLU	2.5
1	A	583	TRP	2.4
1	A	740	TYR	2.4
1	A	570	ARG	2.3
1	A	160	LYS	2.3
1	A	37	SER	2.3
1	A	595	GLY	2.3
1	A	714	GLY	2.2
1	A	69	TRP	2.1
1	A	367	ASN	2.1
1	A	737	ILE	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 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	MPD	A	1746	8/8	0.83	0.26	69,70,71,71	0
3	MPD	A	1747	8/8	0.86	0.24	68,70,70,71	0
2	CA	A	1743	1/1	0.97	0.04	38,38,38,38	0
2	CA	A	1744	1/1	0.97	0.04	38,38,38,38	0
2	CA	A	1745	1/1	0.98	0.03	34,34,34,34	0

6.5 Other polymers [i](#)

There are no such residues in this entry.