



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2026 – 04:45 PM UTC

PDB ID : 6AIJ / pdb_00006aij
Title : Cyclodextrin glycosyltransferase from Paenibacillus macerans mutant N603D
Authors : Li, C.M.; Ban, X.F.; Li, Z.F.; Li, Y.L.; Cheng, S.D.; Zhang, C.Y.; Jin, T.C.;
Gu, Z.B.
Deposited on : 2018-08-24
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
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

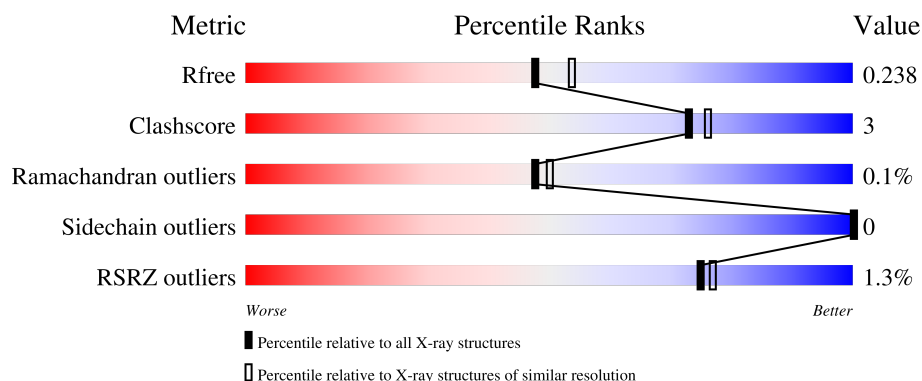
1 Overall quality at a glance

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	708	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5278 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 Cyclomaltodextrin glucanotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	687	Total	C	N	O	S	0	3	0
			5252	3311	879	1049	13			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP P04830
A	-19	LYS	-	expression tag	UNP P04830
A	-18	LYS	-	expression tag	UNP P04830
A	-17	THR	-	expression tag	UNP P04830
A	-16	ALA	-	expression tag	UNP P04830
A	-15	ILE	-	expression tag	UNP P04830
A	-14	ALA	-	expression tag	UNP P04830
A	-13	ILE	-	expression tag	UNP P04830
A	-12	ALA	-	expression tag	UNP P04830
A	-11	VAL	-	expression tag	UNP P04830
A	-10	ALA	-	expression tag	UNP P04830
A	-9	LEU	-	expression tag	UNP P04830
A	-8	ALA	-	expression tag	UNP P04830
A	-7	GLY	-	expression tag	UNP P04830
A	-6	PHE	-	expression tag	UNP P04830
A	-5	ALA	-	expression tag	UNP P04830
A	-4	THR	-	expression tag	UNP P04830
A	-3	VAL	-	expression tag	UNP P04830
A	-2	ALA	-	expression tag	UNP P04830
A	-1	GLN	-	expression tag	UNP P04830
A	156	ALA	GLY	see sequence details	UNP P04830
A	157	LEU	MET	see sequence details	UNP P04830
A	604	ASP	ASN	engineered mutation	UNP P04830

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total 4	Ca 4	0	0

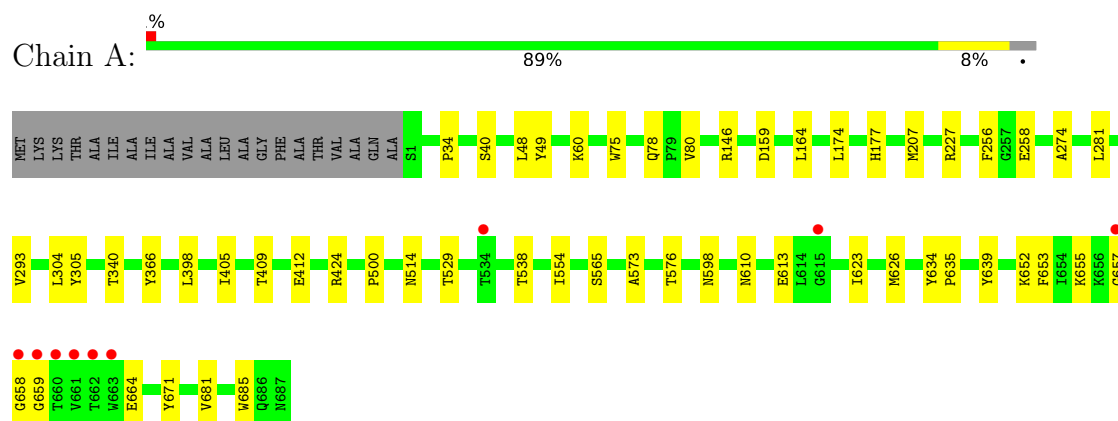
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	22	Total 22	O 22	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: Cyclomaltodextrin glucanotransferase



4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	102.35Å 102.35Å 115.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.39 – 2.10 48.39 – 2.10	Depositor EDS
% Data completeness (in resolution range)	98.9 (48.39-2.10) 98.9 (48.39-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.26 (at 2.10Å)	Xtriage
Refinement program	PHENIX (1.12_2829: ???)	Depositor
R, R_{free}	0.181 , 0.237 0.183 , 0.238	Depositor DCC
R_{free} test set	1996 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	35.9	Xtriage
Anisotropy	0.353	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 40.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.055 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	5278	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% 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: 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	0.33	0/5378	0.51	0/7331

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	5252	0	4969	33	0
2	A	4	0	0	0	0
3	A	22	0	0	0	0
All	All	5278	0	4969	33	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 (33) 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:565:SER:HB3	1:A:576:THR:HG22	1.75	0.68

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:613:GLU:O	1:A:623:ILE:HD11	1.96	0.66
1:A:227:ARG:HG2	1:A:256:PHE:CE2	2.35	0.62
1:A:174:LEU:HD22	1:A:207:MET:HE3	1.81	0.62
1:A:653:PHE:O	1:A:664:GLU:HB2	2.04	0.58
1:A:227:ARG:HH11	1:A:258:GLU:HB3	1.68	0.57
1:A:657:GLY:O	1:A:659:GLY:N	2.37	0.57
1:A:293:VAL:HG21	1:A:304:LEU:HB2	1.85	0.57
1:A:258:GLU:HG3	1:A:258:GLU:O	2.05	0.56
1:A:227:ARG:NH1	1:A:258:GLU:HB3	2.21	0.55
1:A:305:TYR:CE1	1:A:412:GLU:HB3	2.42	0.54
1:A:78:GLN:HG2	1:A:80:VAL:HG22	1.90	0.53
1:A:40:SER:HB2	1:A:48:LEU:HD12	1.90	0.53
1:A:671:TYR:CG	1:A:681:VAL:HG21	2.44	0.52
1:A:75:TRP:CZ2	1:A:227:ARG:HG3	2.46	0.51
1:A:398[B]:LEU:HB3	1:A:405[B]:ILE:HD12	1.91	0.51
1:A:340:THR:HB	1:A:366:TYR:CE1	2.46	0.50
1:A:409:THR:O	1:A:424:ARG:HA	2.14	0.48
1:A:598:ASN:O	1:A:655:LYS:NZ	2.37	0.47
1:A:60:LYS:HD3	1:A:60:LYS:HA	1.60	0.46
1:A:514:ASN:HB2	1:A:554:ILE:HD12	1.98	0.45
1:A:634:TYR:CD1	1:A:635:PRO:HA	2.53	0.43
1:A:146:ARG:NH1	1:A:177:HIS:HB3	2.34	0.42
1:A:610:ASN:OD1	1:A:610:ASN:N	2.48	0.42
1:A:626:MET:HG2	1:A:639:TYR:HB2	2.01	0.42
1:A:75:TRP:CE2	1:A:227:ARG:HG3	2.55	0.42
1:A:34:PRO:HG2	1:A:49:TYR:CG	2.55	0.42
1:A:500:PRO:HG2	1:A:573:ALA:O	2.20	0.41
1:A:655:LYS:HE3	1:A:685:TRP:CZ2	2.56	0.41
1:A:652:LYS:NZ	1:A:664:GLU:O	2.50	0.41
1:A:159:ASP:HB2	1:A:164:LEU:HD11	2.02	0.40
1:A:529:THR:HG23	1:A:538:THR:HG23	2.03	0.40
1:A:274:ALA:HB2	1:A:281:LEU:HD12	2.03	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	688/708 (97%)	661 (96%)	26 (4%)	1 (0%)	48	50

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	658	GLY

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	552/562 (98%)	552 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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	120	ASN
1	A	160	ASN
1	A	239	GLN
1	A	251	HIS
1	A	393	GLN
1	A	581	ASN
1	A	686	GLN

5.3.3 RNA ⓘ

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 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	687/708 (97%)	-0.10	9 (1%) 75 77	16, 39, 68, 129	3 (0%)

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	661	VAL	4.1
1	A	659	GLY	3.8
1	A	658	GLY	3.8
1	A	534	THR	2.8
1	A	615	GLY	2.7
1	A	662	THR	2.4
1	A	663	TRP	2.4
1	A	660	THR	2.3
1	A	657	GLY	2.1

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
2	CA	A	704	1/1	0.71	0.13	95,95,95,95	0
2	CA	A	702	1/1	0.73	0.25	100,100,100,100	0
2	CA	A	703	1/1	0.99	0.02	32,32,32,32	0
2	CA	A	701	1/1	0.99	0.03	41,41,41,41	0

6.5 Other polymers [i](#)

There are no such residues in this entry.