



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2025 – 12:15 AM JST

PDB ID : 9IHT / pdb_00009iht
Title : Crystal structure of GH57 family amylopullulanase from Aquifex aeolicus in complex with acarbose
Authors : Zhu, Z.M.; Wang, W.W.; Yu, F.
Deposited on : 2024-06-18
Resolution : 1.69 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 Phenix2.0rc1
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.43.1

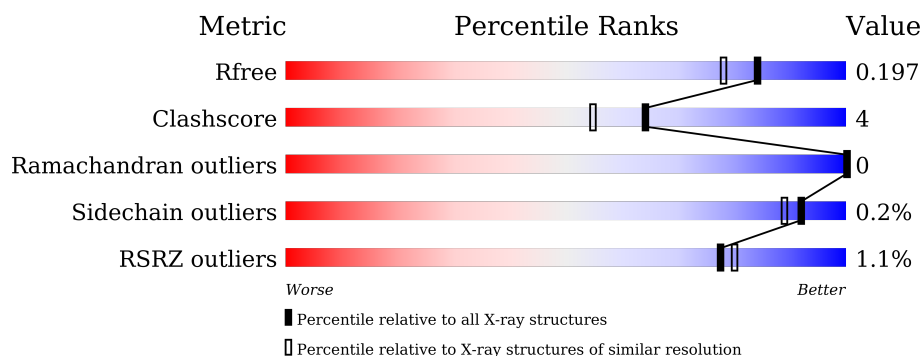
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.69 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}	164625	5161 (1.70-1.70)
Clashscore	180529	5671 (1.70-1.70)
Ramachandran outliers	177936	5594 (1.70-1.70)
Sidechain outliers	177891	5594 (1.70-1.70)
RSRZ outliers	164620	5159 (1.70-1.70)

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	477	<div> <div>90%</div> <div>9%</div> </div>
1	B	477	<div> <div>2%</div> <div>91%</div> <div>8%</div> </div>
2	C	3	<div> <div>67%</div> <div>33%</div> </div>
2	D	3	<div> <div>33%</div> <div>67%</div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 8717 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.

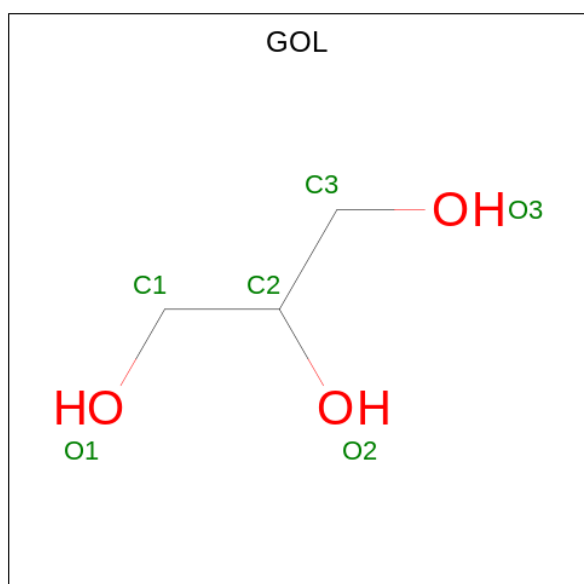
- Molecule 1 is a protein called Glycoside hydrolase family 57 N-terminal domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	476	Total	C	N	O	S	0	2	0
			4063	2660	649	744	10			
1	B	476	Total	C	N	O	S	0	3	0
			4069	2663	650	746	10			

- Molecule 2 is an oligosaccharide called 4,6-dideoxy-4-[(1S,4R,5S,6S)-4,5,6-trihydroxy-3-(hydroxymethyl)cyclohex-2-en-1-yl]amino}-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	3	Total	C	N	O	0	0	0
			44	25	1	18			
2	D	3	Total	C	N	O	0	0	0
			44	25	1	18			

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).

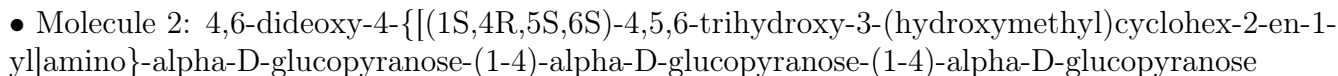
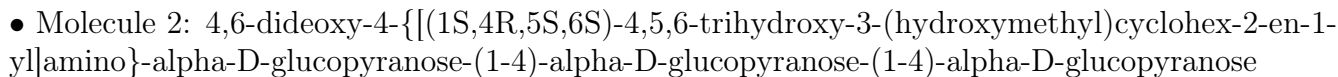
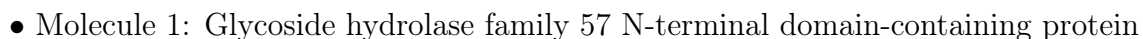


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total 6	C 3	O 3	0	0
3	B	1	Total 6	C 3	O 3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	268	Total 268	O 268	0	0
4	B	217	Total 217	O 217	0	0

- Molecule 1: Glycoside hydrolase family 57 N-terminal domain-containing protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	60.93Å 40.11Å 195.14Å 90.00° 95.80° 90.00°	Depositor
Resolution (Å)	60.62 – 1.69 60.62 – 1.69	Depositor EDS
% Data completeness (in resolution range)	99.3 (60.62-1.69) 99.3 (60.62-1.69)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 1.69Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.171 , 0.200 0.169 , 0.197	Depositor DCC
R_{free} test set	5153 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å ²)	22.1	Xtriage
Anisotropy	0.366	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 39.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	8717	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.85% 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: GLC, AC1, 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	0.42	0/4177	0.59	1/5649 (0.0%)
1	B	0.38	2/4183 (0.0%)	0.54	0/5657
All	All	0.40	2/8360 (0.0%)	0.57	1/11306 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	306	ARG	C-O	-5.48	1.16	1.23
1	B	241	TYR	C-O	-5.03	1.18	1.24

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	A	457	LYS	N-CA-C	7.31	121.54	112.47

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	195	PRO	Peptide
1	B	195	PRO	Peptide

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	4063	0	3991	36	0
1	B	4069	0	3995	30	0
2	C	44	0	30	1	0
2	D	44	0	30	0	0
3	A	6	0	8	0	0
3	B	6	0	8	0	0
4	A	268	0	0	4	1
4	B	217	0	0	7	0
All	All	8717	0	8062	62	1

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.

The worst 5 of 62 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:361:GLU:OE2	1:A:362:GLU:OE1	1.99	0.80
1:B:427:LYS:NZ	4:B:602:HOH:O	2.18	0.77
1:A:328:ARG:NH1	1:B:376:GLU:O	2.19	0.76
1:B:332:GLY:O	1:B:336:LYS:HD3	1.85	0.75
1:A:456:GLU:O	1:A:456:GLU:HG3	1.88	0.73

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:617:HOH:O	4:A:849:HOH:O[1_455]	2.13	0.07

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	476/477 (100%)	469 (98%)	7 (2%)	0	100	100
1	B	477/477 (100%)	473 (99%)	4 (1%)	0	100	100
All	All	953/954 (100%)	942 (99%)	11 (1%)	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	438/437 (100%)	437 (100%)	1 (0%)	92	89
1	B	439/437 (100%)	438 (100%)	1 (0%)	92	89
All	All	877/874 (100%)	875 (100%)	2 (0%)	92	89

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	456	GLU
1	B	121	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	134	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	152	ASN
1	B	355	ASN
1	A	248	HIS
1	A	134	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 ⓘ

6 monosaccharides 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	GLC	C	1	2	12,12,12	1.01	1 (8%)	17,17,17	1.38	4 (23%)
2	GLC	C	2	2	11,11,12	0.65	0	15,15,17	1.05	1 (6%)
2	AC1	C	3	2	21,22,23	6.74	10 (47%)	22,32,34	1.56	4 (18%)
2	GLC	D	1	2	12,12,12	0.99	1 (8%)	17,17,17	1.34	3 (17%)
2	GLC	D	2	2	11,11,12	0.50	0	15,15,17	0.83	0
2	AC1	D	3	2	21,22,23	6.67	10 (47%)	22,32,34	1.22	3 (13%)

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	GLC	C	1	2	-	0/2/22/22	0/1/1/1
2	GLC	C	2	2	-	0/2/19/22	0/1/1/1
2	AC1	C	3	2	-	3/6/43/46	0/2/2/2
2	GLC	D	1	2	-	0/2/22/22	0/1/1/1
2	GLC	D	2	2	-	0/2/19/22	0/1/1/1
2	AC1	D	3	2	-	1/6/43/46	0/2/2/2

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	3	AC1	C7B-C5B	20.50	1.61	1.32
2	D	3	AC1	C7B-C5B	20.45	1.61	1.32
2	D	3	AC1	C2B-C1B	12.22	1.69	1.52
2	C	3	AC1	C2B-C1B	12.14	1.68	1.52
2	D	3	AC1	C3B-C4A	11.90	1.70	1.53

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	3	AC1	C7B-C1B-N4A	-4.48	103.96	110.68
2	C	2	GLC	C1-O5-C5	2.87	116.08	112.19
2	C	1	GLC	C1-O5-C5	2.79	118.93	113.66
2	D	3	AC1	C7B-C1B-N4A	-2.71	106.61	110.68
2	C	1	GLC	O3-C3-C4	2.56	116.27	110.35

There are no chirality outliers.

All (4) torsion outliers are listed below:

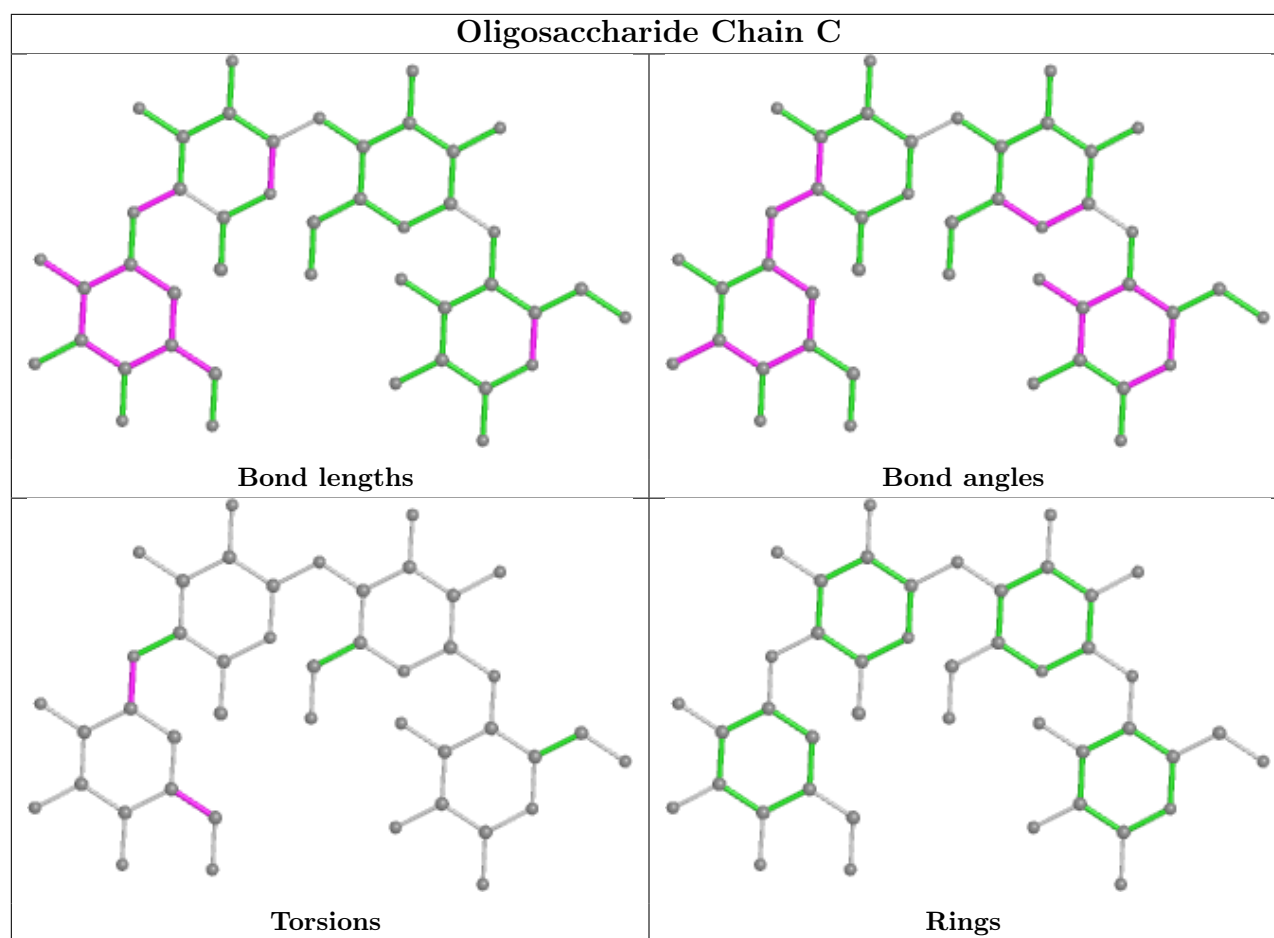
Mol	Chain	Res	Type	Atoms
2	C	3	AC1	C2B-C1B-N4A-C4
2	C	3	AC1	C7B-C5B-C6B-O6B
2	D	3	AC1	C2B-C1B-N4A-C4
2	C	3	AC1	C4A-C5B-C6B-O6B

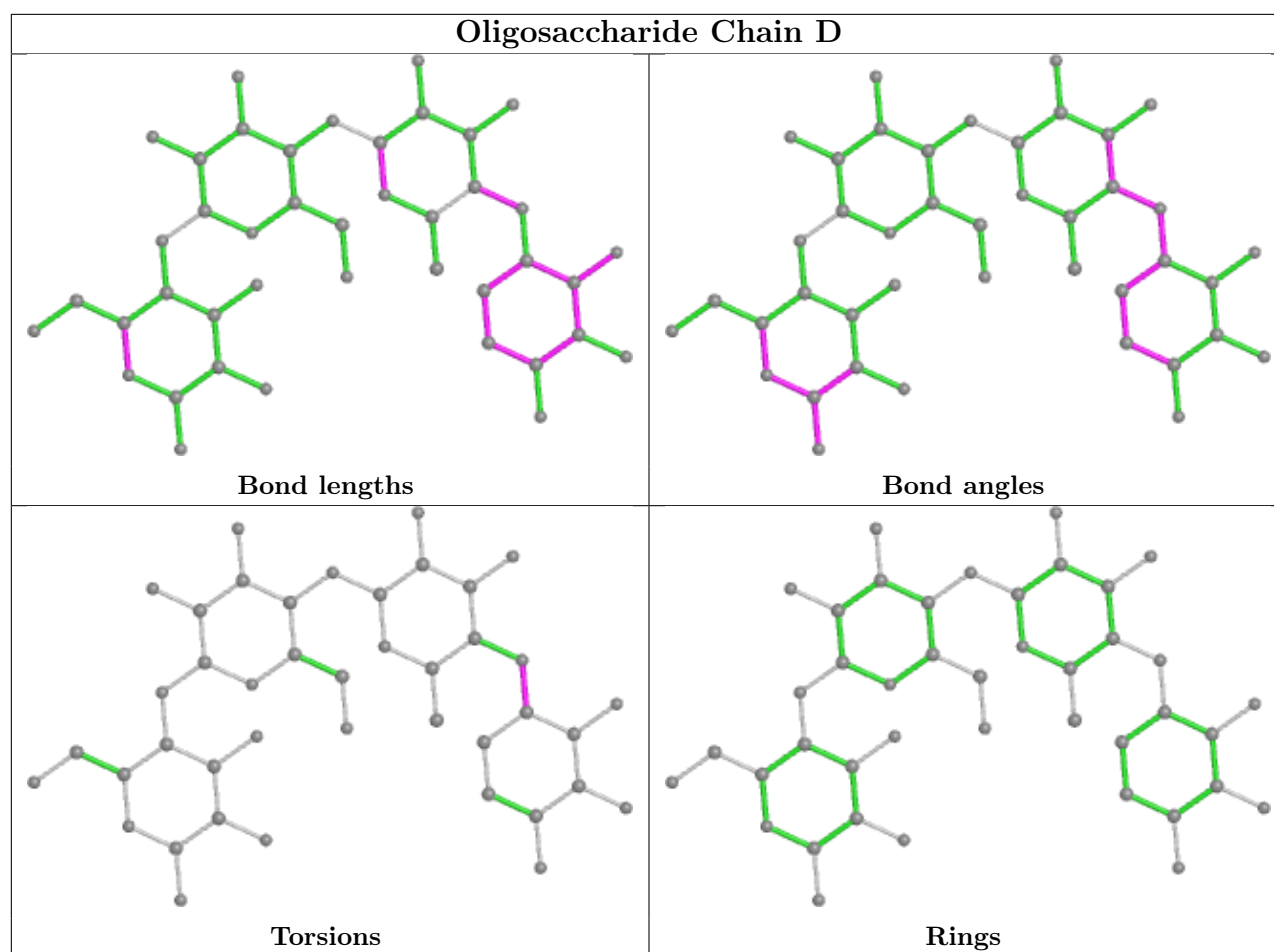
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	3	AC1	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.





5.6 Ligand geometry [i](#)

2 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	$\# Z > 2$
3	GOL	B	501	-	5,5,5	1.02	0	5,5,5	1.02	0
3	GOL	A	501	-	5,5,5	0.97	0	5,5,5	0.90	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	GOL	B	501	-	-	0/4/4/4	-
3	GOL	A	501	-	-	0/4/4/4	-

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	476/477 (99%)	-0.36	2 (0%) 89 90	9, 20, 40, 75	2 (0%)
1	B	476/477 (99%)	-0.16	8 (1%) 69 71	8, 24, 46, 80	3 (0%)
All	All	952/954 (99%)	-0.26	10 (1%) 77 80	8, 21, 44, 80	5 (0%)

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	103	TYR	3.3
1	B	289	ALA	3.1
1	B	76[A]	PHE	3.1
1	A	458	ALA	3.0
1	B	288	ARG	2.9

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](#)

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.

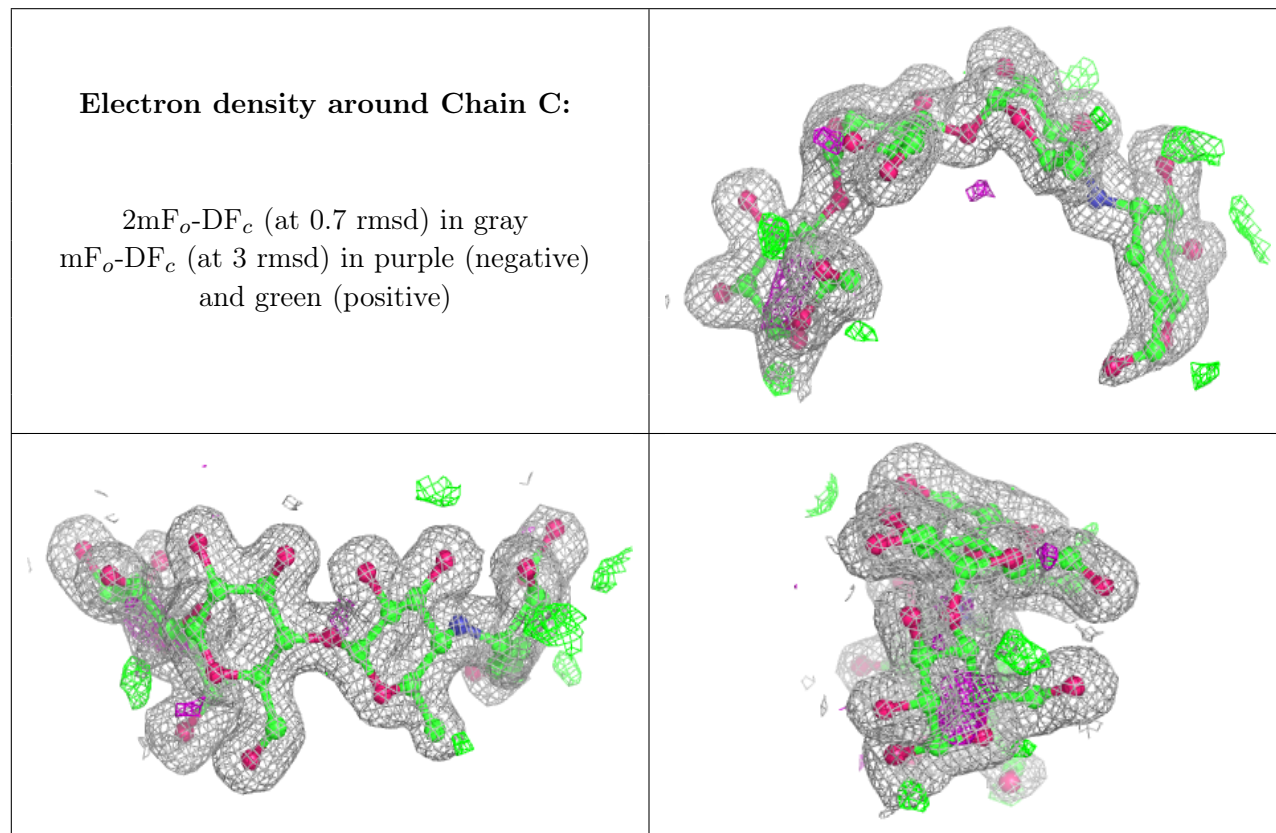
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	GLC	C	1	12/12	0.95	0.08	10,17,27,37	0
2	AC1	C	3	21/22	0.96	0.07	9,18,43,48	0
2	GLC	D	1	12/12	0.96	0.07	11,16,24,42	0
2	AC1	D	3	21/22	0.96	0.07	10,16,39,53	0
2	GLC	D	2	11/12	0.98	0.04	9,11,12,14	0

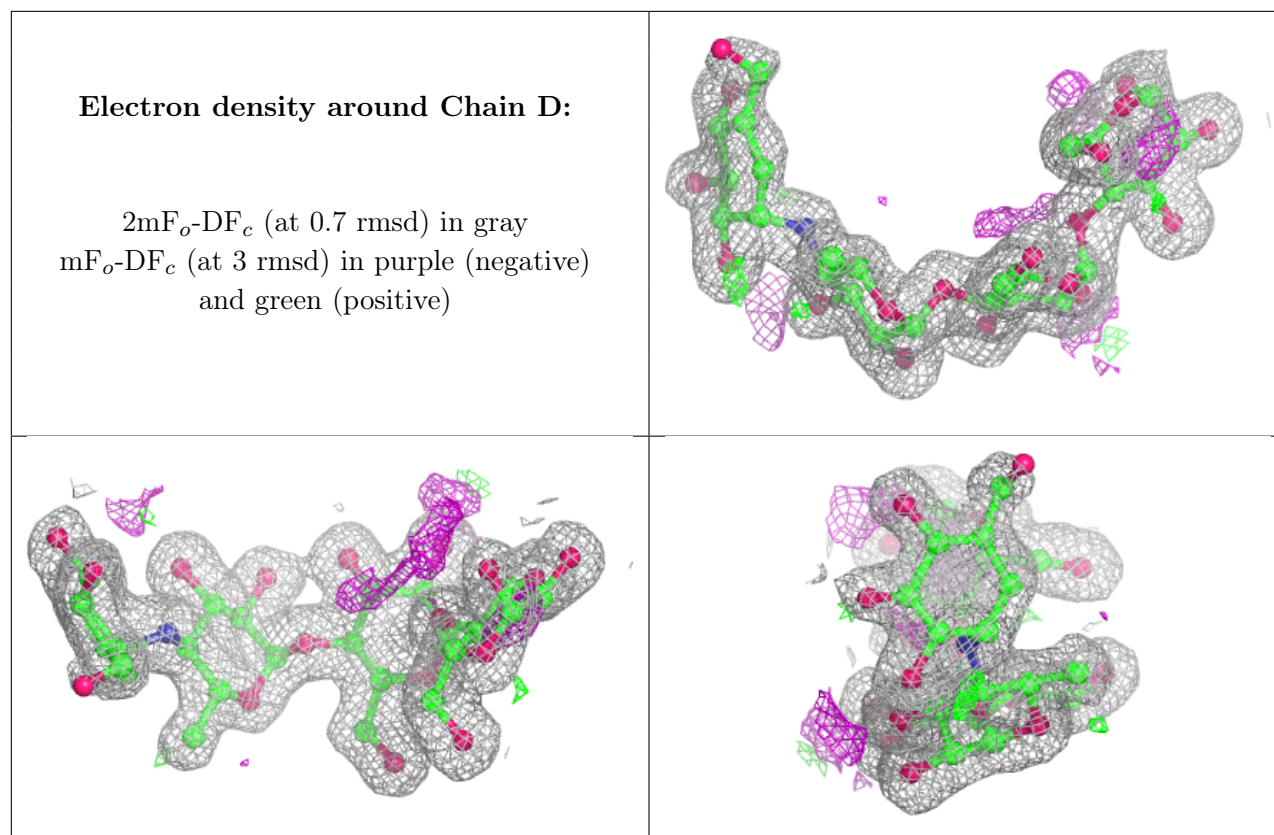
Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	GLC	C	2	11/12	0.98	0.03	9,11,12,12	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.





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(\AA^2)	Q<0.9
3	GOL	A	501	6/6	0.91	0.09	30,35,36,38	0
3	GOL	B	501	6/6	0.97	0.13	11,24,27,29	0

6.5 Other polymers [i](#)

There are no such residues in this entry.