



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 18, 2023 – 09:17 am GMT

PDB ID : 4B9Z
Title : Crystal Structure of Agd31B, alpha-transglucosylase, complexed with Acarbose
Authors : Larsbrink, J.; Izumi, A.; Hemsworth, G.R.; Davies, G.J.; Brumer, H.
Deposited on : 2012-09-09
Resolution : 2.00 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 : 4.02b-467
Mogul : 1.8.4, CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : **FAILED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

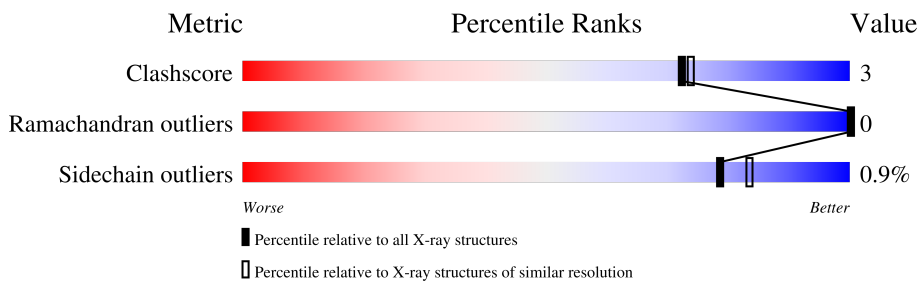
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 $\leq 5\%$.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	817	
2	B	3	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GLC	B	1	X	-	-	-
2	GLC	B	2	X	-	-	-

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 6804 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-GLUCOSIDASE, PUTATIVE, ADG31B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	780	6218	3978	1057	1161	22	4	4	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	817	LYS	-	expression tag	UNP B3PEE6

- 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.

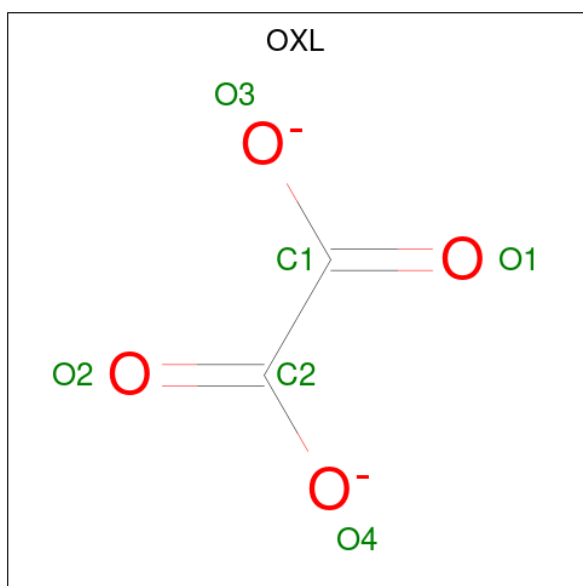
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	3	44	25	1	18	0	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0

- Molecule 4 is OXALATE ION (three-letter code: OXL) (formula: C_2O_4).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 2 4	0	0

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		
6	A	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0

- Molecule 7 is water.


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	444	Total O 444 444	0	0

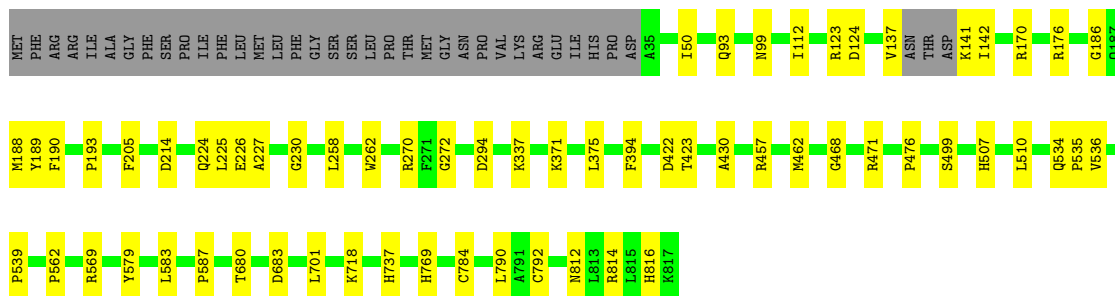
3 Residue-property plots

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. 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. 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.

Note EDS failed to run properly.

- Molecule 1: ALPHA-GLUCOSIDASE, PUTATIVE, ADG31B

Chain A: 



- Molecule 2: 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

Chain B: 



4 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 6 2 2	Depositor
Cell constants a, b, c, α , β , γ	196.88Å 196.88Å 102.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.73 – 2.00	Depositor
% Data completeness (in resolution range)	100.0 (49.73-2.00)	Depositor
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.91 (at 2.00Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.161 , 0.187	Depositor
Wilson B-factor (Å ²)	24.6	Xtriage
Anisotropy	0.440	Xtriage
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
Total number of atoms	6804	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% 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: OXL, PEG, AC1, SO4, GLC, EDO

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	0.37	0/6396	0.53	0/8678

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 [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	6218	0	6023	43	0
2	B	44	0	30	0	0
3	A	25	0	0	0	0
4	A	6	0	0	0	0
5	A	7	0	10	1	0
6	A	60	0	90	7	0
7	A	444	0	0	2	0
All	All	6804	0	6153	43	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.

The worst 5 of 43 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:737:HIS:HD2	1:A:814:ARG:HG2	1.31	0.91
1:A:784:CYS:SG	1:A:792[B]:CYS:HB2	2.17	0.83
1:A:737:HIS:NE2	1:A:812:ASN:OD1	2.12	0.82
1:A:142:ILE:HD11	1:A:227:ALA:HB3	1.65	0.79
1:A:562:PRO:HB3	6:A:1835:EDO:H21	1.64	0.78

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	780/817 (96%)	755 (97%)	25 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	653/690 (95%)	647 (99%)	6 (1%)	78 83

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	394	PHE
1	A	680	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	701	LEU
1	A	294	ASP
1	A	190	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	735	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](#)

3 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	B	1	2	12,12,12	0.48	0	17,17,17	1.32	4 (23%)
2	GLC	B	2	2	11,11,12	0.32	0	15,15,17	1.98	4 (26%)
2	AC1	B	3	2	21,22,23	0.74	0	22,32,34	0.96	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
2	GLC	B	1	2	3/3/5/5	2/2/22/22	0/1/1/1
2	GLC	B	2	2	1/1/4/5	0/2/19/22	0/1/1/1
2	AC1	B	3	2	-	3/6/43/46	0/2/2/2

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2	GLC	C1-C2-C3	-4.81	103.75	109.67
2	B	2	GLC	C1-O5-C5	-3.34	107.67	112.19
2	B	2	GLC	O5-C1-C2	-3.25	105.76	110.77
2	B	1	GLC	C3-C4-C5	2.78	115.20	110.24
2	B	1	GLC	C4-C3-C2	2.28	114.81	110.82

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	1	GLC	C3
2	B	1	GLC	C1
2	B	1	GLC	C4
2	B	2	GLC	C1

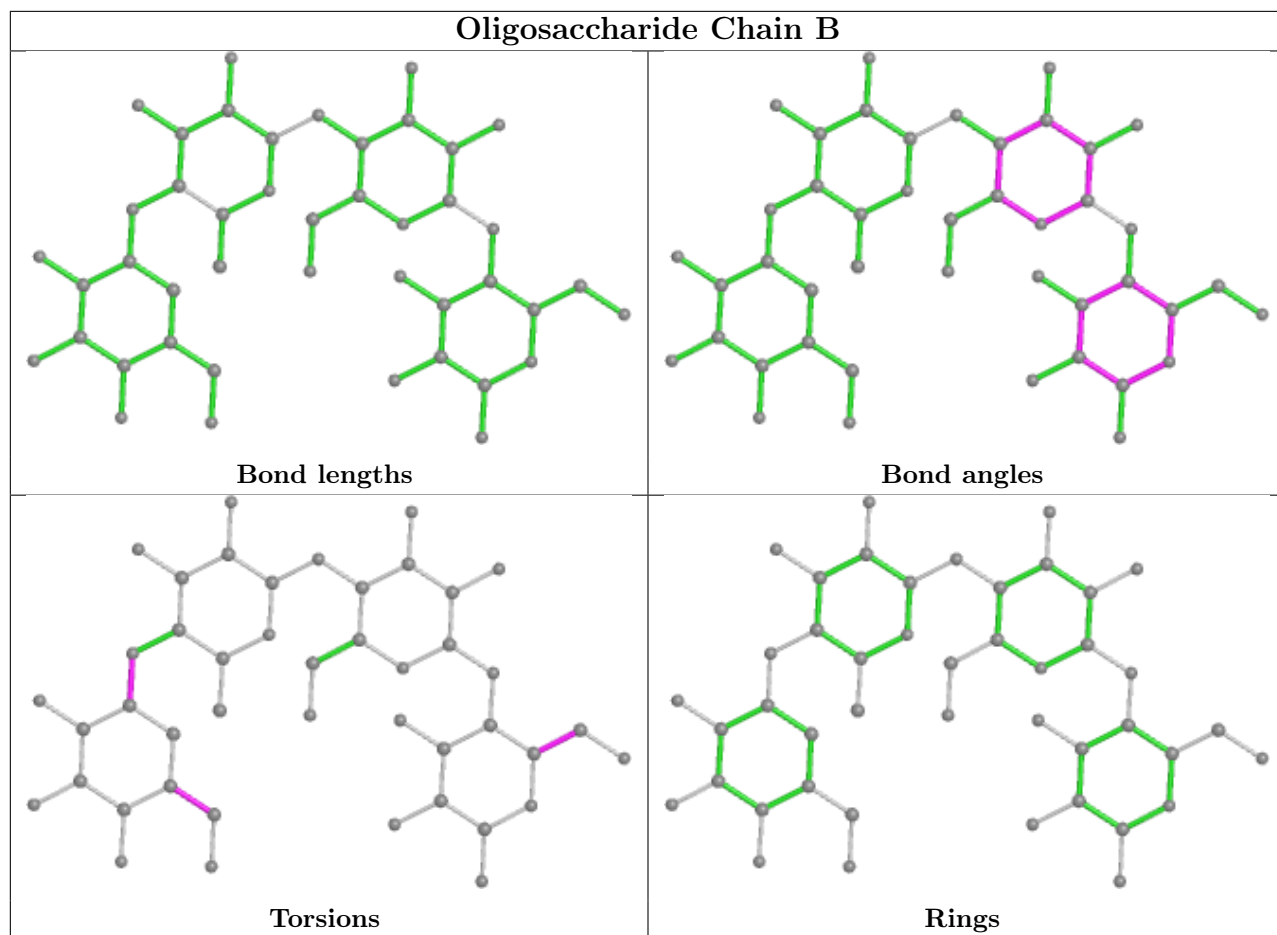
All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	3	AC1	C7B-C1B-N4A-C4
2	B	3	AC1	C4A-C5B-C6B-O6B
2	B	3	AC1	C7B-C5B-C6B-O6B
2	B	1	GLC	O5-C5-C6-O6
2	B	1	GLC	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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](#)

22 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	$\# Z > 2$
6	EDO	A	1827	-	3,3,3	0.49	0	2,2,2	0.31	0
6	EDO	A	1832	-	3,3,3	0.50	0	2,2,2	0.39	0
6	EDO	A	1834	-	3,3,3	0.47	0	2,2,2	0.34	0
3	SO4	A	1823	-	4,4,4	0.14	0	6,6,6	0.06	0
5	PEG	A	1825	-	6,6,6	0.44	0	5,5,5	0.30	0
6	EDO	A	1840	-	3,3,3	0.45	0	2,2,2	0.32	0
6	EDO	A	1829	-	3,3,3	0.44	0	2,2,2	0.39	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	SO4	A	1822	-	4,4,4	0.14	0	6,6,6	0.18	0
3	SO4	A	1821	-	4,4,4	0.14	0	6,6,6	0.08	0
6	EDO	A	1833	-	3,3,3	0.50	0	2,2,2	0.33	0
3	SO4	A	1820	-	4,4,4	0.14	0	6,6,6	0.17	0
4	OXL	A	1824	-	5,5,5	1.79	1 (20%)	6,6,6	1.29	1 (16%)
6	EDO	A	1835	-	3,3,3	0.51	0	2,2,2	0.31	0
6	EDO	A	1839	-	3,3,3	0.43	0	2,2,2	0.41	0
6	EDO	A	1830	-	3,3,3	0.52	0	2,2,2	0.14	0
6	EDO	A	1831	-	3,3,3	0.44	0	2,2,2	0.35	0
3	SO4	A	1819	-	4,4,4	0.14	0	6,6,6	0.16	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
6	EDO	A	1827	-	-	0/1/1/1	-
6	EDO	A	1832	-	-	1/1/1/1	-
6	EDO	A	1834	-	-	0/1/1/1	-
5	PEG	A	1825	-	-	3/4/4/4	-
6	EDO	A	1840	-	-	0/1/1/1	-
6	EDO	A	1829	-	-	1/1/1/1	-
6	EDO	A	1833	-	-	0/1/1/1	-
4	OXL	A	1824	-	-	3/4/4/4	-
6	EDO	A	1835	-	-	1/1/1/1	-
6	EDO	A	1839	-	-	0/1/1/1	-
6	EDO	A	1830	-	-	1/1/1/1	-
6	EDO	A	1831	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1824	OXL	C2-C1	-2.61	1.47	1.54

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1824	OXL	O3-C1-C2	2.12	119.47	113.16

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1825	PEG	O2-C3-C4-O4
6	A	1835	EDO	O1-C1-C2-O2
6	A	1830	EDO	O1-C1-C2-O2
5	A	1825	PEG	C1-C2-O2-C3
6	A	1829	EDO	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1825	PEG	1	0
6	A	1829	EDO	2	0
6	A	1833	EDO	2	0
6	A	1835	EDO	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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

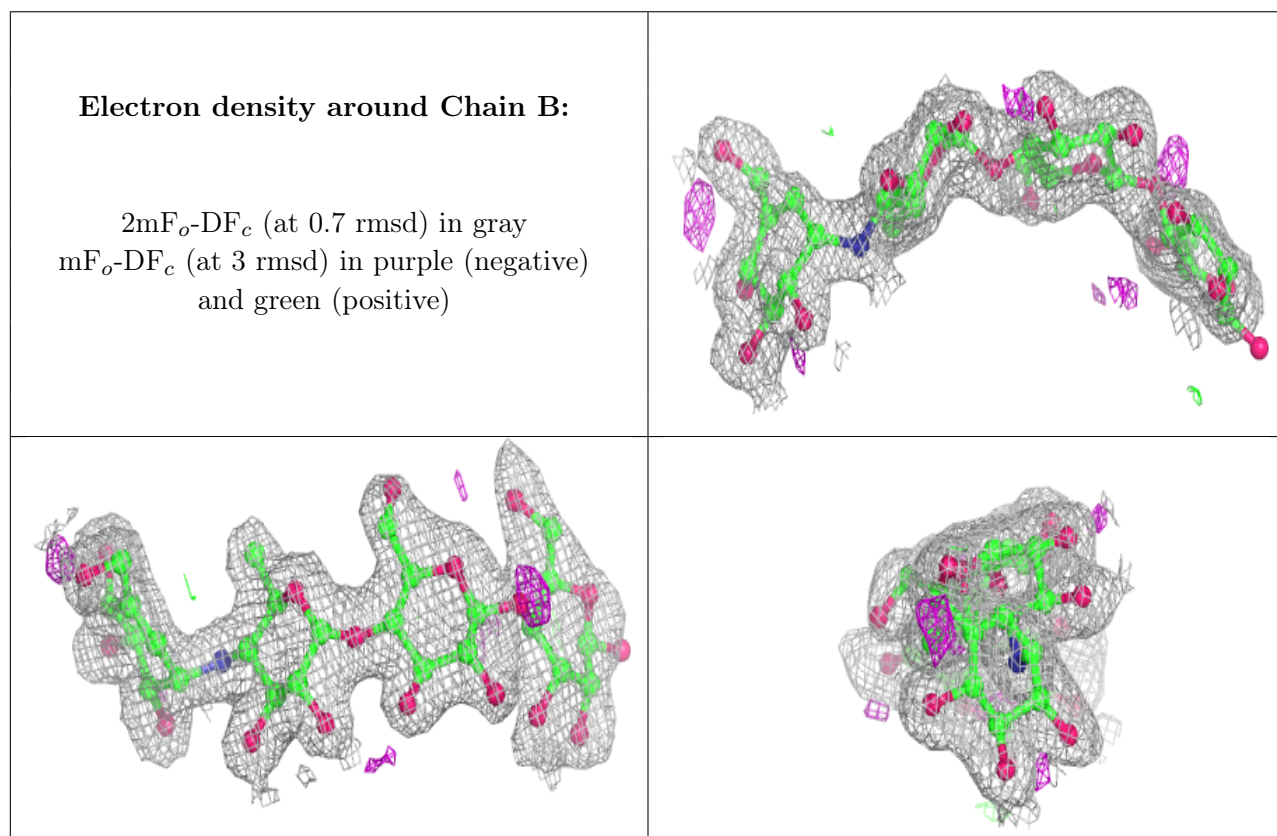
6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS failed to run properly - this section is therefore empty.

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](#)

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers [i](#)

EDS failed to run properly - this section is therefore empty.