Full wwPDB X-ray Structure Validation Report ⓘ

Nov 5, 2023 – 01:46 AM EDT

PDB ID	:	2OWW
Title	:	Covalent intermediate in amylomaltase in complex with the acceptor analog 4-deoxyglucose
Authors	:	Barends, T.R.M.; Bultema, J.B.; Kaper, T.; van der Maarel, M.J.E.C.; Dijkhuizen, L.; Dijkstra, B.W.
Deposited on	:	2007-02-17
Resolution	:	2.20 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

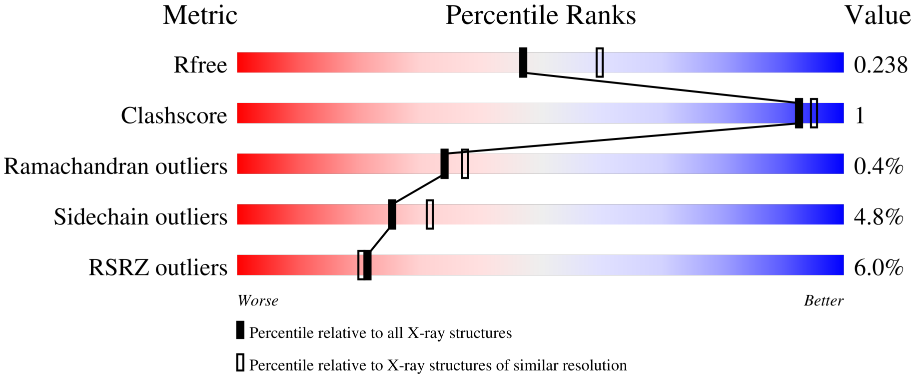
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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

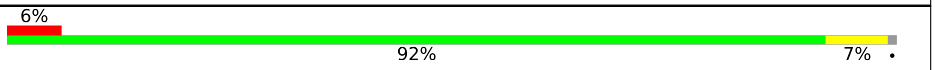
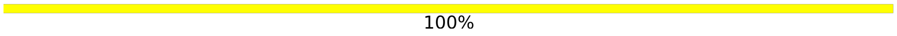
The reported resolution of this entry is 2.20 Å.

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}	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 <=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	502	
2	B	2	

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 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4476 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 4-alpha-glucanotransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	499	Total 4093	C 2659	N 712	O 711	S 11	0	4	0

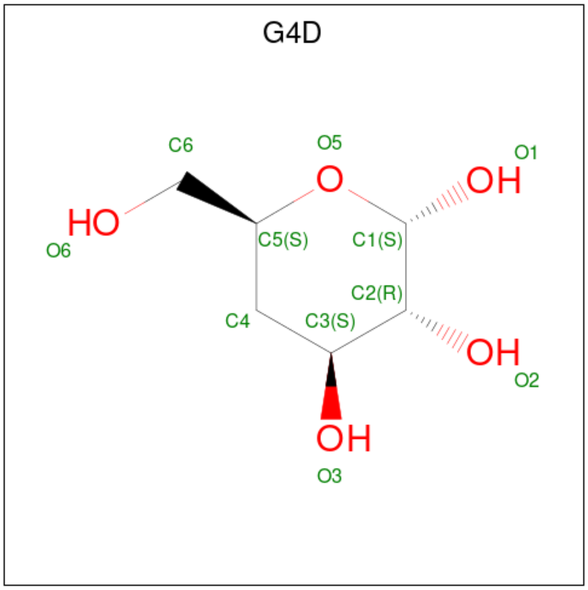
There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	cloning artifact	UNP Q72J82
A	-1	SER	-	cloning artifact	UNP Q72J82
A	0	HIS	-	cloning artifact	UNP Q72J82
A	370	SUI	ASP	modified residue	UNP Q72J82
A	370	SUI	GLY	modified residue	UNP Q72J82

- 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.

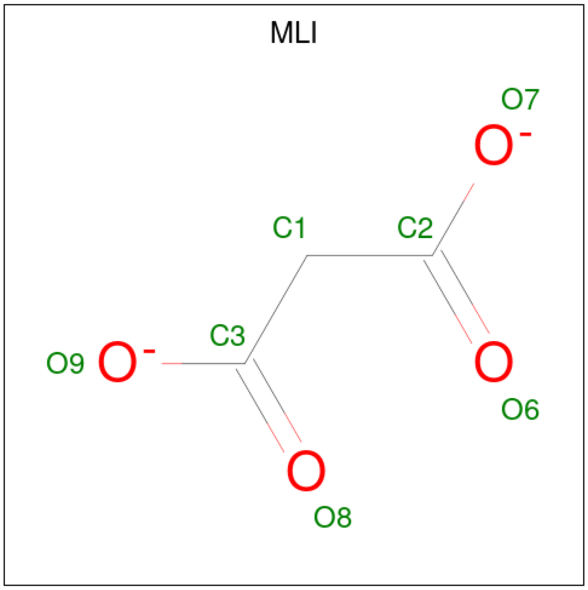
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	2	Total 32	C 19	N 1	O 12	0	0	0

- Molecule 3 is 4-deoxy-alpha-D-glucopyranose (three-letter code: G4D) (formula: $C_6H_{12}O_5$).



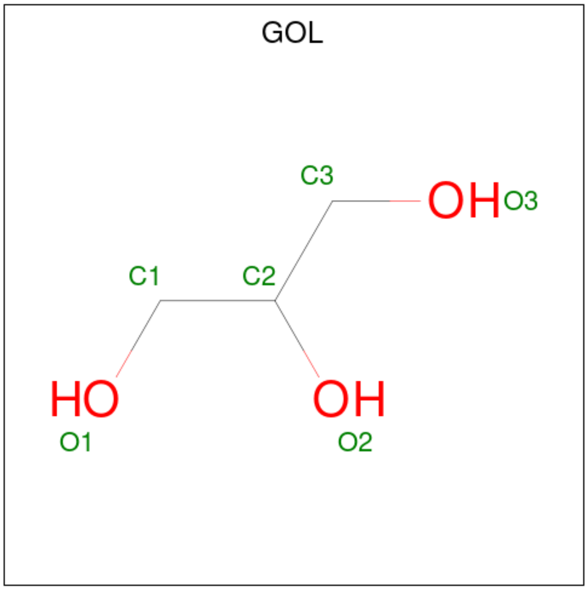
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 11 6 5	0	0

- Molecule 4 is MALONATE ION (three-letter code: MLI) (formula: $C_3H_2O_4$).



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

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 6 is water.

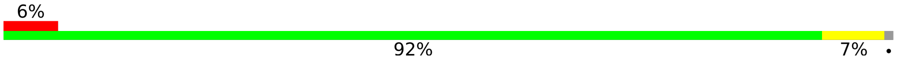
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	320	Total O 321 321	0	1

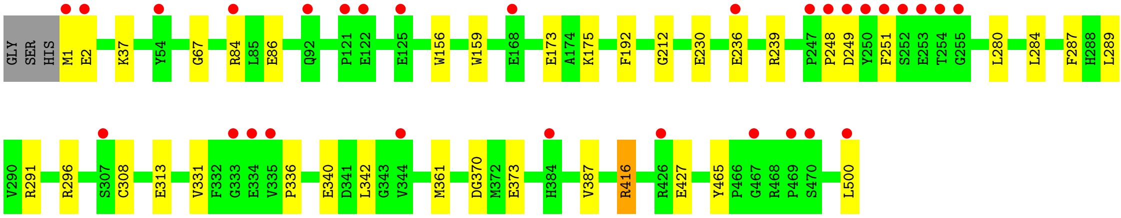
3 Residue-property plots

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 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: 4-alpha-glucanotransferase

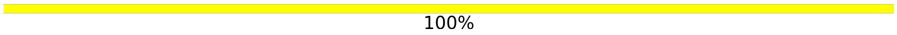
Chain A:





- 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

Chain B:



4 Data and refinement statistics

Property	**Value**	**Source**
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	92.46Å 92.46Å 154.17Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	14.94 – 2.20 14.94 – 2.20	Depositor EDS
% Data completeness (in resolution range)	98.7 (14.94-2.20) 98.6 (14.94-2.20)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.16 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.198 , 0.240 0.198 , 0.238	Depositor DCC
R_{free} test set	1948 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	25.9	Xtriage
Anisotropy	0.040	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 50.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.024 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4476	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 4.13% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: SUI, MLI, G4D, 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	0.41	0/4230	0.58	0/5750

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	4093	0	3944	12	0
2	B	32	0	19	0	0
3	A	11	0	12	0	0
4	A	7	0	2	0	0
5	A	12	0	16	0	0
6	A	321	0	0	3	0
All	All	4476	0	3993	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:291[B]:ARG:NH2	6:A:3000[B]:HOH:O	1.82	0.97
1:A:291[B]:ARG:NH1	6:A:3000[B]:HOH:O	2.37	0.58
1:A:291[B]:ARG:CZ	6:A:3000[B]:HOH:O	2.37	0.57
1:A:1:MET:SD	1:A:289:LEU:HD21	2.49	0.52
1:A:156:TRP:HA	1:A:159:TRP:CD2	2.47	0.49
1:A:416:ARG:HD3	1:A:416:ARG:O	2.15	0.46
1:A:67:GLY:HA3	1:A:192:PHE:CD2	2.52	0.45
1:A:291[A]:ARG:HD3	1:A:340:GLU:HB2	1.99	0.44
1:A:284:LEU:HD21	1:A:336:PRO:HG2	2.00	0.43
1:A:308:CYS:SG	1:A:313:GLU:HB2	2.59	0.42
1:A:342:LEU:HD11	1:A:361:MET:HE2	2.01	0.42
1:A:212:GLY:HA3	1:A:287:PHE:CD1	2.55	0.42

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	500/502 (100%)	487 (97%)	11 (2%)	2 (0%)	34 37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	248	PRO
1	A	251	PHE

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	404/402 (100%)	384 (95%)	20 (5%)	24 30

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	GLU
1	A	37	LYS
1	A	84	ARG
1	A	86	GLU
1	A	173[A]	GLU
1	A	173[B]	GLU
1	A	175	LYS
1	A	230	GLU
1	A	236	GLU
1	A	239	ARG
1	A	249	ASP
1	A	280	LEU
1	A	296	ARG
1	A	331	VAL
1	A	373	GLU
1	A	387	VAL
1	A	416	ARG
1	A	427	GLU
1	A	465	TYR
1	A	500	LEU

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	409	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
1	SUI	A	370	1	10,11,12	2.95	3 (30%)	11,15,17	2.28	5 (45%)

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
1	SUI	A	370	1	-	1/2/19/20	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	370	SUI	C2A-N2	-8.32	1.30	1.47
1	A	370	SUI	C1-N2	-2.80	1.34	1.38
1	A	370	SUI	CA-N	-2.10	1.37	1.48

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	370	SUI	CB-CG-N2	3.85	110.40	108.13
1	A	370	SUI	O-C-C2A	-3.34	116.29	126.39
1	A	370	SUI	OD-CG-CB	-3.18	122.61	127.24
1	A	370	SUI	OD-CG-N2	2.77	126.98	123.92
1	A	370	SUI	O1-C1-N2	2.14	126.75	124.14

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	370	SUI	C-C2A-N2-CG

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

2 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	2,1	11,11,12	0.75	0	15,15,17	1.06	1 (6%)
2	AC1	B	2	2	21,22,23	0.66	0	22,32,34	1.05	1 (4%)

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	2,1	1/1/4/5	0/2/19/22	0/1/1/1
2	AC1	B	2	2	-	0/6/43/46	0/2/2/2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	2	AC1	C2-C3-C4	-2.53	108.39	110.63
2	B	1	GLC	C1-C2-C3	2.32	112.51	109.67

All (1) chirality outliers are listed below:

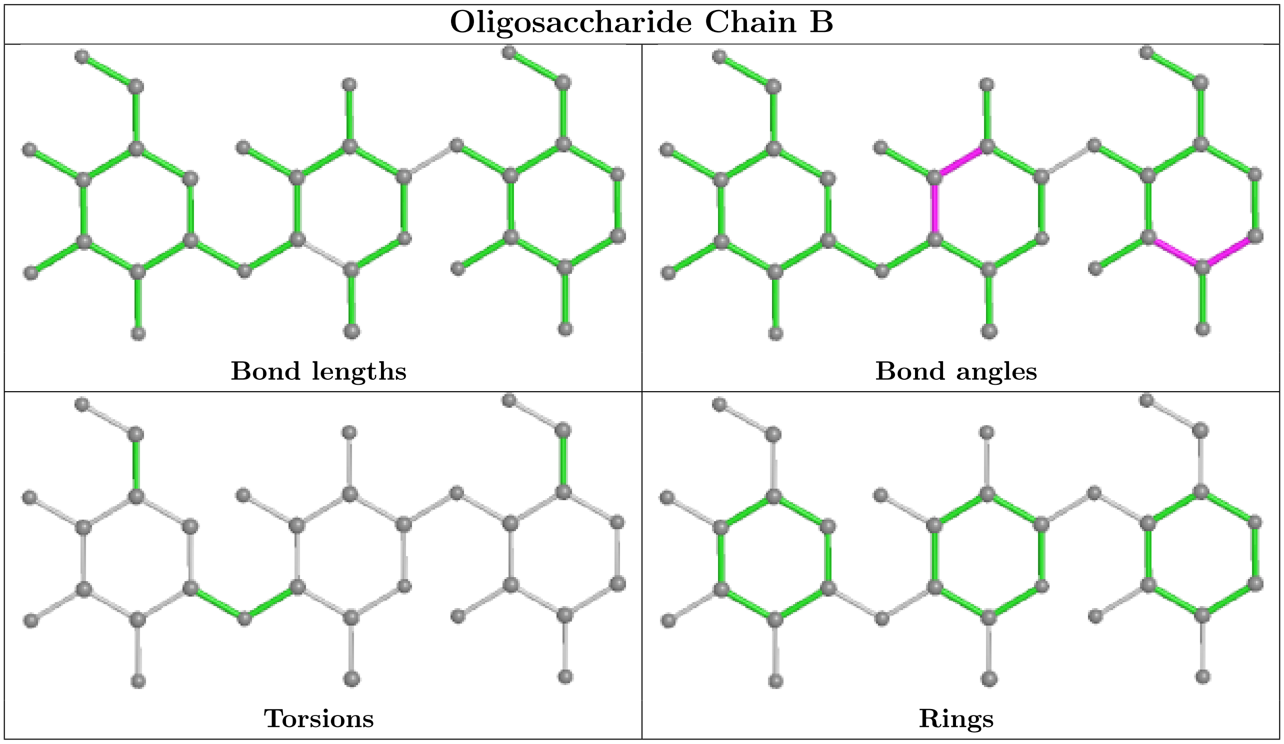
Mol	Chain	Res	Type	Atom
2	B	1	GLC	C1

There are no torsion outliers.

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.


Oligosaccharide Chain B

5.6 Ligand geometry

4 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	$\#\|Z\| > 2$
5	GOL	A	2303	-	5,5,5	0.39	0	5,5,5	0.46	0
5	GOL	A	2302	-	5,5,5	0.45	0	5,5,5	0.50	0
4	MLI	A	2301	-	6,6,6	1.11	0	7,7,7	0.98	0
3	G4D	A	2300	-	11,11,11	0.57	0	15,15,15	1.37	3 (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
5	GOL	A	2303	-	-	0/4/4/4	-
5	GOL	A	2302	-	-	4/4/4/4	-
4	MLI	A	2301	-	-	2/4/4/4	-
3	G4D	A	2300	-	-	0/2/18/18	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	2300	G4D	C1-C2-C3	2.35	115.19	110.31
3	A	2300	G4D	C4-C5-C6	-2.21	108.72	112.60
3	A	2300	G4D	O5-C5-C6	2.01	110.06	106.83

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	2302	GOL	O1-C1-C2-C3
5	A	2302	GOL	C1-C2-C3-O3
4	A	2301	MLI	C2-C1-C3-O9
4	A	2301	MLI	C2-C1-C3-O8
5	A	2302	GOL	O1-C1-C2-O2
5	A	2302	GOL	O2-C2-C3-O3

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	498/502 (99%)	-0.15	30 (6%)	21	20	17, 26, 47, 63	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	250	TYR	9.5
1	A	254	THR	8.2
1	A	251	PHE	7.8
1	A	253	GLU	7.0
1	A	249	ASP	6.7
1	A	2	GLU	5.8
1	A	248	PRO	5.6
1	A	335	VAL	5.5
1	A	252	SER	5.1
1	A	469	PRO	3.7
1	A	384	HIS	3.6
1	A	1	MET	3.4
1	A	255	GLY	3.1
1	A	426	ARG	2.9
1	A	333	GLY	2.9
1	A	84	ARG	2.9
1	A	467	GLY	2.8
1	A	307	SER	2.7
1	A	334	GLU	2.7
1	A	344	VAL	2.5
1	A	168	GLU	2.5
1	A	125	GLU	2.4
1	A	236	GLU	2.4
1	A	54	TYR	2.4
1	A	121	PRO	2.3
1	A	92	GLN	2.3
1	A	500	LEU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	247	PRO	2.2
1	A	470	SER	2.2
1	A	122	GLU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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, 95^{th} 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.

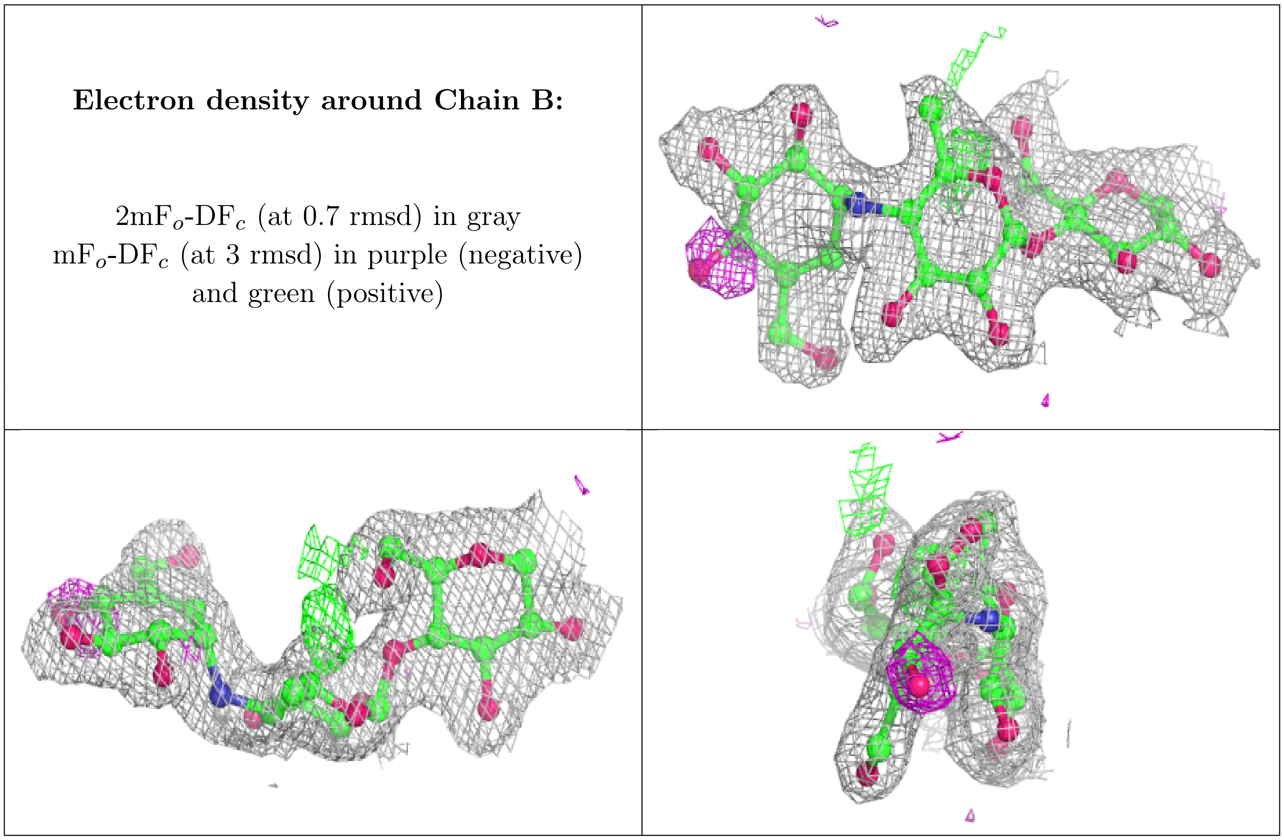
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
1	SUI	A	370	11/12	0.97	0.06	22,23,25,25	0

6.3 Carbohydrates ⓘ

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, 95^{th} 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	AC1	B	2	21/22	0.75	0.20	32,36,42,44	0
2	GLC	B	1	11/12	0.94	0.12	28,30,32,32	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.



Electron density around Chain B:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

6.4 Ligands ⓘ

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
5	GOL	A	2302	6/6	0.79	0.20	45,47,48,48	0
3	G4D	A	2300	11/11	0.82	0.15	57,57,58,60	0
4	MLI	A	2301	7/7	0.94	0.12	45,46,46,46	0
5	GOL	A	2303	6/6	0.96	0.08	29,31,31,32	0

6.5 Other polymers ⓘ

There are no such residues in this entry.