W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Aug 9, 2020 – 09:24 AM BST

PDB ID	:	3T5O
Title	:	Crystal Structure of human Complement Component C6
Authors	:	Aleshin, A.E.; Stec, B.; Bankston, L.A.; DiScipio, R.G.; Liddington, R.C.
Deposited on	:	2011-07-27
Resolution	:	2.87 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 (i) symbol.

The following versions of software and data (see references (i)) 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	:	2.13.1
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)	:	2.13.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.87 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.
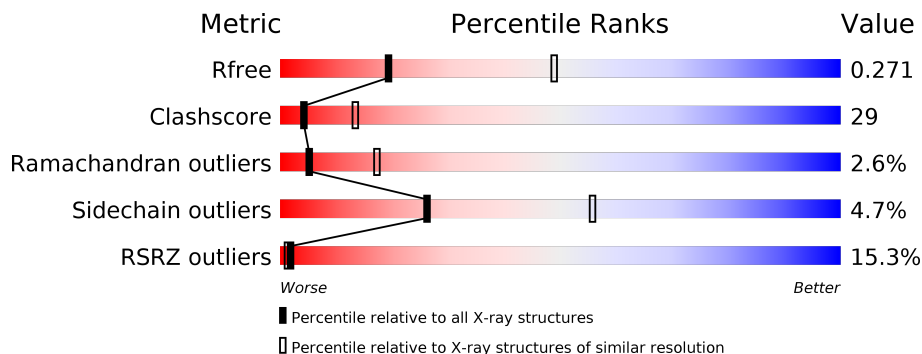
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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 on 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	913	15% / 52% / 38% / 5% / 5%
2	B	2	100%
3	C	2	50% / 50%

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	NAG	B	2	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 6934 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 Complement component C6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	871	Total 6830	C 4222	N 1203	O 1336	S 69	0	0	0

- Molecule 2 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a cetamido-2-deoxy-beta-D-glucopyranose.
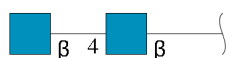
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 3 is an oligosaccharide called beta-D-glucopyranose-(1-3)-alpha-L-fucopyranose.
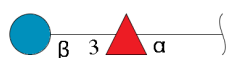
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	C	2	Total 21	C 12	O 9	0	0	0

- Molecule 4 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Cd 1	0	0

- Molecule 5 is beta-L-fucopyranose (three-letter code: FUL) (formula: $C_6H_{12}O_5$).

FUL

C6O5O1
C5(S)C1(S)
C4(S)C2(S)
C3(R)
HO
O4
OH
O3
OH
O2
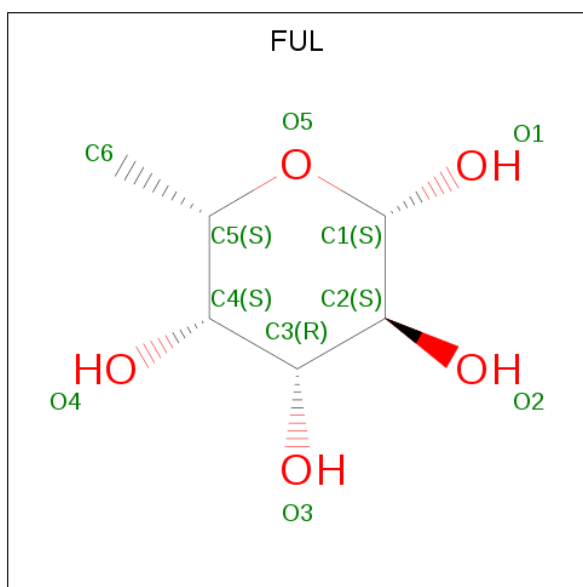
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total 10	C 6	O 4	0	0

- Molecule 6 is alpha-D-mannopyranose (three-letter code: MAN) (formula: $C_6H_{12}O_6$).
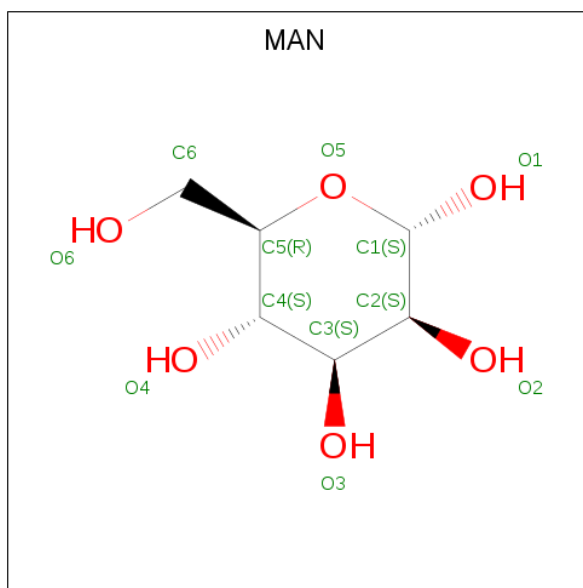
MAN

C6O5O1
HO
O6
C5(R)C1(S)
C4(S)C2(S)
C3(S)
HO
O4
OH
O2
OH
O3

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

Continued on next page...

Continued from previous page...

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

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: Complement component C6

E823 W824 G825 L826 E827 R828 T829 R830 L831 S832 S833 N834 S835 T836 K837 K838 E839 S840 C841 G842 Y843 D844 T845 C846 Y847 D848 W849 E850 K851 C852 S853 A854 S855 T856 S857 K858 C859 V860 C861 L862 L863 P864 P865 Q866 C867 F868 K869 G870 G871 N872 Q873 L874 Y875 C876 V877 K878 M879 G880 S881 S882

T883 S884 E885 K886 T887 L888 N889 I890 C891 E892 V893 G894 T895 I896 R897 C898 A899 N900 R901 K902 M903 E904 I905 L906 G909 K910 C911 L912 A913

• Molecule 2: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain B: 100%

NAG1 NAG2

• Molecule 3: beta-D-glucopyranose-(1-3)-alpha-L-fucopyranose

Chain C: 50%   50%

FUC1 BGC2

4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	146.81Å 180.15Å 60.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.06 – 2.87 47.53 – 2.87	Depositor EDS
% Data completeness (in resolution range)	94.9 (43.06-2.87) 94.9 (47.53-2.87)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$<I/\sigma(I)>$ [1]	1.66 (at 2.86Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.222 , 0.276 0.219 , 0.271	Depositor DCC
R_{free} test set	2179 reflections (6.10%)	wwPDB-VP
Wilson B-factor (Å^2)	82.3	Xtriage
Anisotropy	0.026	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 81.4	EDS
L-test for twinning[2]	$<|L|> = 0.50$, $<L^2> = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6934	wwPDB-VP
Average B, all atoms (Å^2)	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.45% 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: BGC, NAG, CD, FUC, FUL, MAN

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	0.50	1/6973 (0.0%)	0.67	3/9409 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	531	CYS	CB-SG	5.23	1.91	1.82

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	652	LEU	CA-CB-CG	-6.84	99.56	115.30
1	A	570	ASN	N-CA-C	-5.68	95.67	111.00
1	A	666	LEU	CA-CB-CG	-5.07	103.64	115.30

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	6830	0	6498	398	0
2	B	28	0	25	0	0
3	C	21	0	19	1	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	10	0	10	3	0
6	A	44	0	40	7	0
All	All	6934	0	6592	398	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 29.

The worst 5 of 398 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:570:ASN:HB3	1:A:571:PRO:CD	1.40	1.41
1:A:824:TRP:HE1	1:A:901:ARG:HD2	1.06	1.17
1:A:570:ASN:HB3	1:A:571:PRO:HD3	1.30	1.13
1:A:476:ILE:HB	1:A:477:PRO:HD2	1.29	1.11
1:A:570:ASN:HB3	1:A:571:PRO:HD2	1.22	1.08

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	861/913 (94%)	734 (85%)	105 (12%)	22 (3%)	5	17

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	150	ARG
1	A	540	SER
1	A	570	ASN
1	A	712	PRO
1	A	727	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	772/810 (95%)	736 (95%)	36 (5%)	26 56

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	530	ASN
1	A	569	ASN
1	A	776	SER
1	A	565	THR
1	A	604	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	312	HIS
1	A	408	HIS
1	A	817	GLN
1	A	329	ASN
1	A	414	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)

4 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	NAG	B	1	1,2	14,14,15	0.62	0	17,19,21	0.77	0
2	NAG	B	2	2	14,14,15	0.53	0	17,19,21	0.68	0
3	FUC	C	1	1,3	10,10,11	0.81	0	14,14,16	1.62	4 (28%)
3	BGC	C	2	3	11,11,12	1.18	1 (9%)	15,15,17	1.62	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
2	NAG	B	1	1,2	-	2/6/23/26	0/1/1/1
2	NAG	B	2	2	-	2/6/23/26	0/1/1/1
3	FUC	C	1	1,3	-	-	0/1/1/1
3	BGC	C	2	3	-	2/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2	BGC	O5-C1	-3.38	1.38	1.43

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	2	BGC	C1-C2-C3	-3.80	104.99	109.67
3	C	1	FUC	C6-C5-C4	-2.91	107.70	113.07
3	C	1	FUC	O5-C5-C6	2.71	113.17	107.33
3	C	1	FUC	C1-C2-C3	2.38	112.59	109.67
3	C	2	BGC	O5-C1-C2	-2.24	107.32	110.77

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	2	NAG	O5-C5-C6-O6
2	B	2	NAG	C4-C5-C6-O6
3	C	2	BGC	O5-C5-C6-O6
2	B	1	NAG	C8-C7-N2-C2
2	B	1	NAG	O7-C7-N2-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	1	FUC	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.

Oligosaccharide Chain B

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain C

Bond lengths

Bond angles

Torsions

Rings
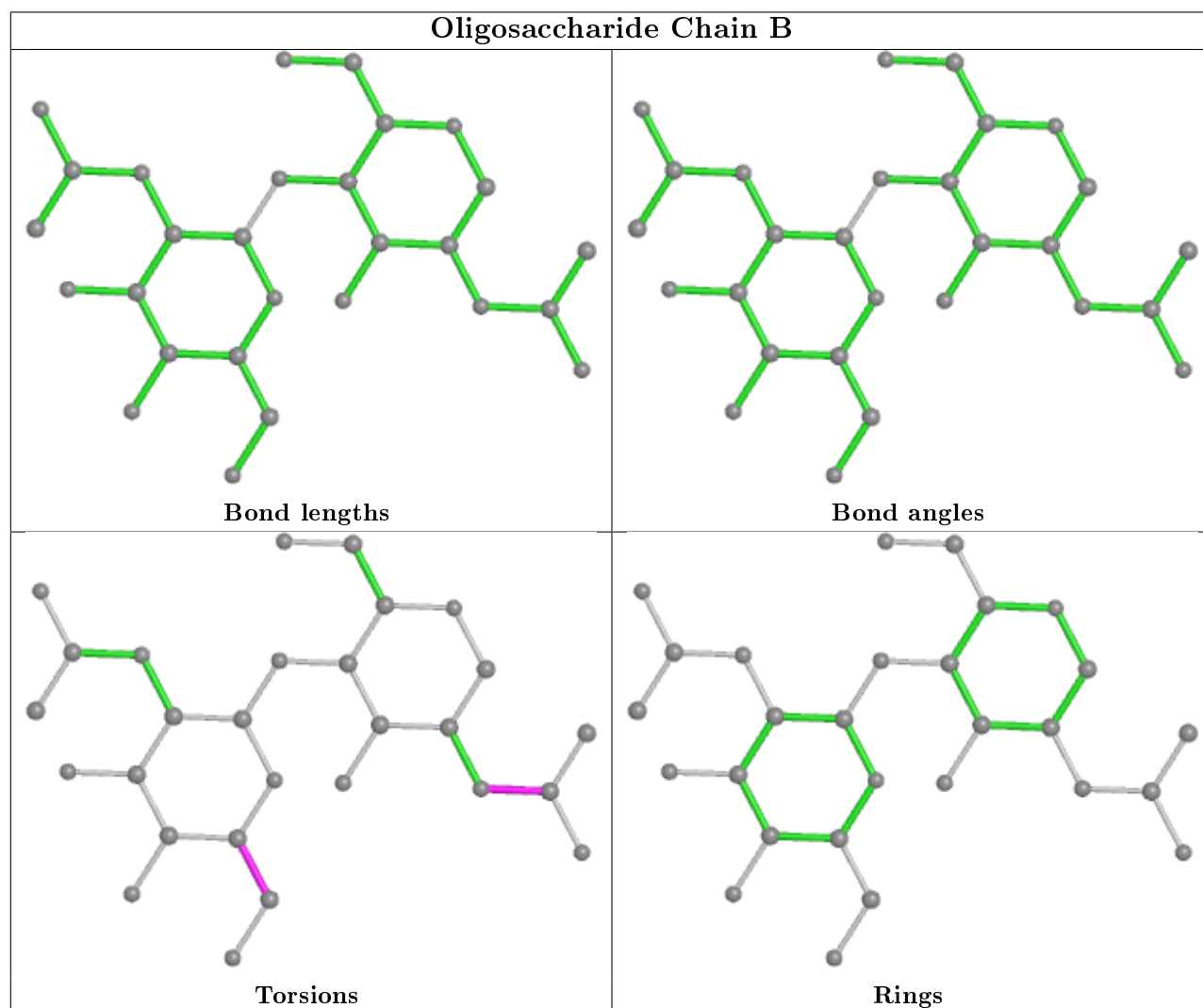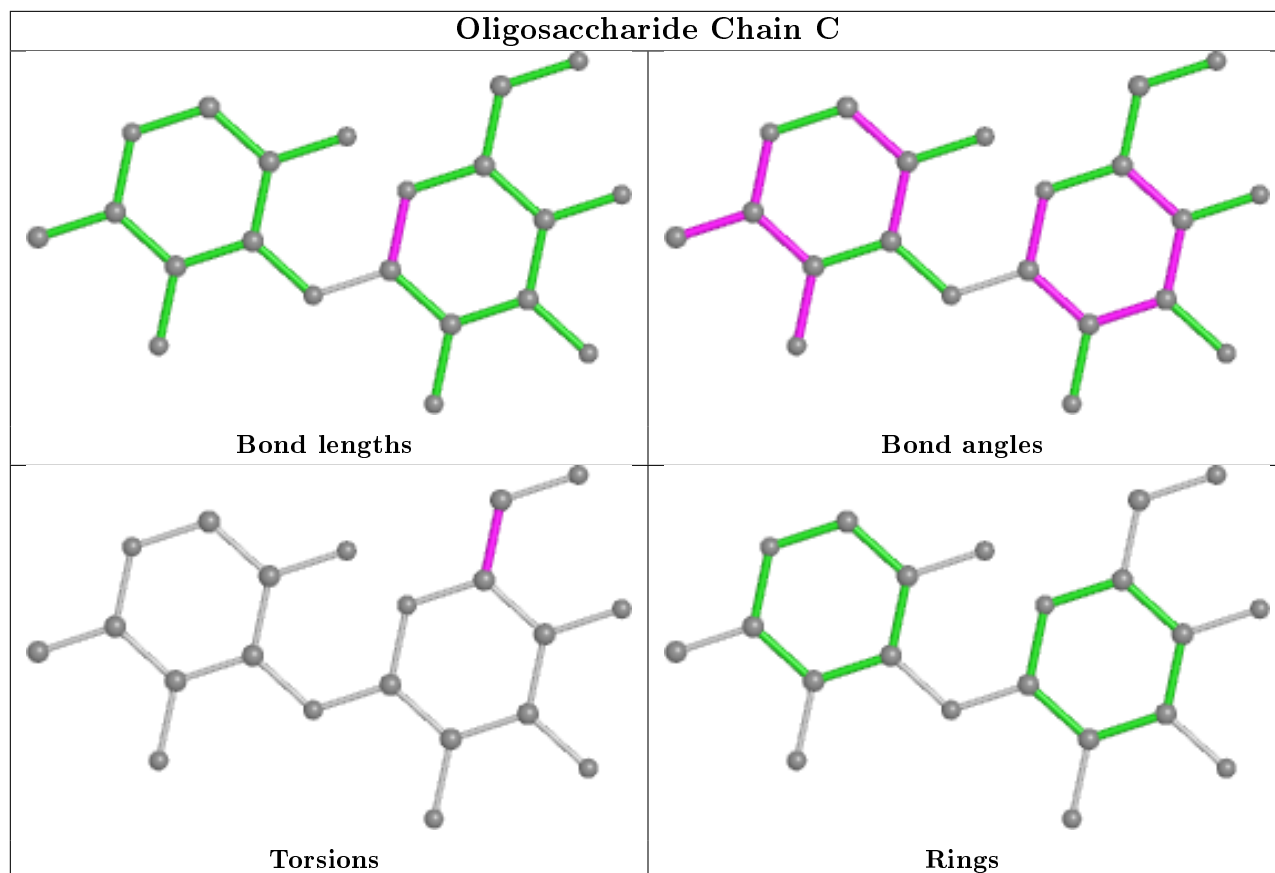
5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 1 is monoatomic - leaving 5 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	$\#|Z| > 2$
6	MAN	A	1010	1	11,11,12	0.47	0	15,15,17	1.35	2 (13%)
6	MAN	A	1007	1	11,11,12	0.90	0	15,15,17	1.35	3 (20%)
5	FUL	A	1004	1	10,10,11	0.66	0	14,14,16	1.03	1 (7%)
6	MAN	A	1008	1	11,11,12	1.18	0	15,15,17	1.77	4 (26%)
6	MAN	A	1009	1	11,11,12	0.65	0	15,15,17	1.51	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
6	MAN	A	1010	1	-	0/2/19/22	0/1/1/1
5	FUL	A	1004	1	-	-	0/1/1/1
6	MAN	A	1007	1	-	0/2/19/22	0/1/1/1
6	MAN	A	1008	1	-	2/2/19/22	1/1/1/1
6	MAN	A	1009	1	-	1/2/19/22	0/1/1/1

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	1008	MAN	C1-C2-C3	-3.22	105.70	109.67
6	A	1010	MAN	C2-C3-C4	-3.18	105.40	110.89
6	A	1009	MAN	C3-C4-C5	3.18	115.90	110.24
6	A	1008	MAN	O4-C4-C3	-3.11	103.16	110.35
6	A	1009	MAN	C1-O5-C5	2.93	116.17	112.19

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	1008	MAN	O5-C5-C6-O6
6	A	1008	MAN	C4-C5-C6-O6
6	A	1009	MAN	C4-C5-C6-O6

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	1008	MAN	C1-C2-C3-C4-C5-O5

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1010	MAN	3	0
6	A	1007	MAN	1	0
5	A	1004	FUL	3	0
6	A	1008	MAN	1	0
6	A	1009	MAN	2	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)

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	871/913 (95%)	0.95	133 (15%)	2	1	41, 107, 238, 291	0

The worst 5 of 133 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	840	SER	14.8
1	A	861	CYS	11.7
1	A	892	GLU	10.2
1	A	913	ALA	10.2
1	A	879	MET	9.7

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

There are no non-standard protein/DNA/RNA residues in this entry.
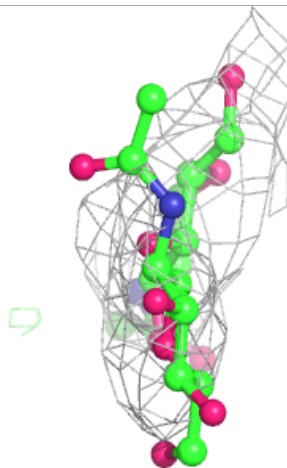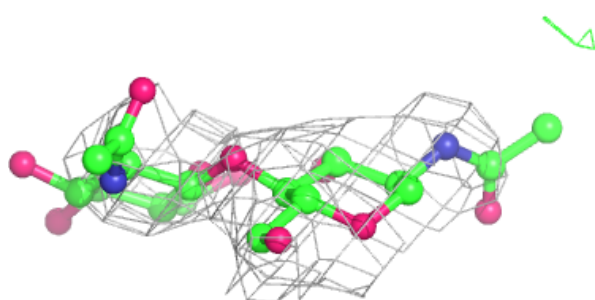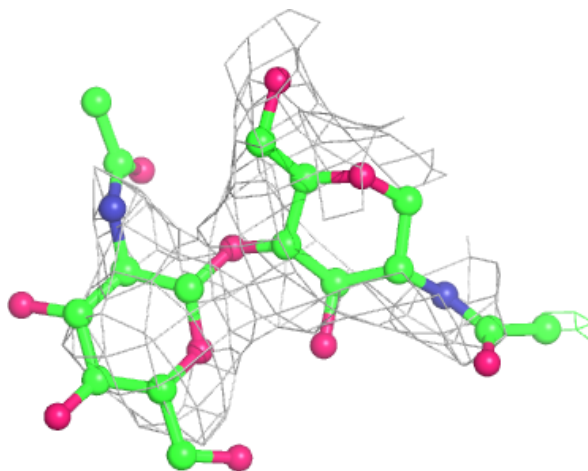
6.3 Carbohydrates (i)

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	NAG	B	2	14/15	0.77	0.54	148,174,183,187	0
2	NAG	B	1	14/15	0.90	0.25	120,151,163,169	0
3	BGC	C	2	11/12	0.97	0.20	51,62,72,80	0
3	FUC	C	1	10/11	0.98	0.19	47,56,68,78	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:

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

Electron density around Chain C:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
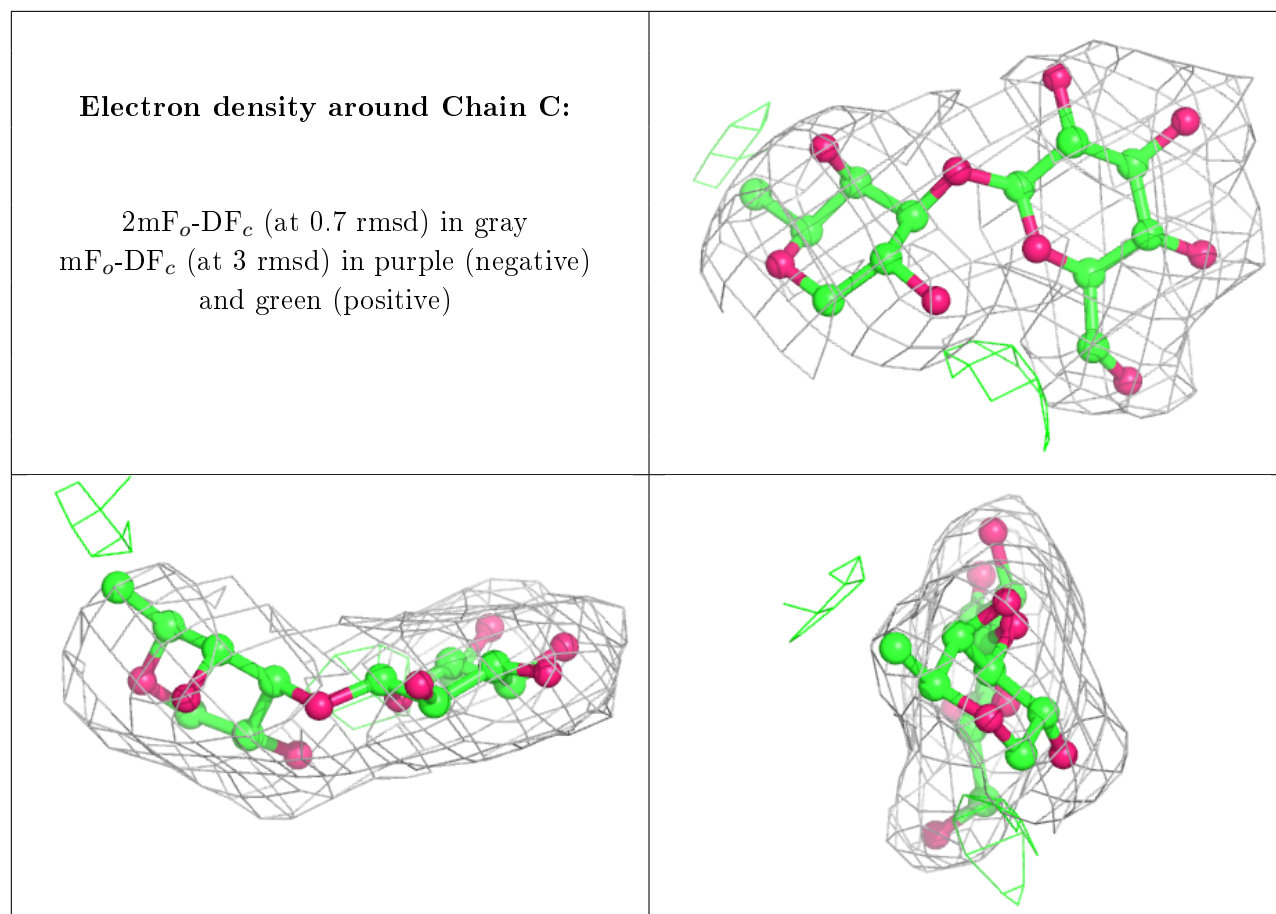
6.4 Ligands (i)

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(Å^2)	Q<0.9
5	FUL	A	1004	10/11	0.81	0.32	154,162,169,172	0
6	MAN	A	1010	11/12	0.88	0.23	127,137,142,143	0
6	MAN	A	1009	11/12	0.92	0.19	152,162,169,172	0
6	MAN	A	1007	11/12	0.95	0.22	73,87,95,102	0
6	MAN	A	1008	11/12	0.96	0.22	58,71,81,107	0
4	CD	A	1001	1/1	0.97	0.18	133,133,133,133	0

6.5 Other polymers (i)

There are no such residues in this entry.