Full wwPDB X-ray Structure Validation Report (i)

Jan 7, 2024 – 04:29 am GMT

PDB ID	:	6F6N
Title	:	CRYSTAL STRUCTURE OF EBOLAVIRUS GLYCOPROTEIN IN COMPLEX WITH SERTRALINE
Authors	:	Ren, J.; Zhao, Y.; Fry, E.E.; Stuart, D.I.
Deposited on	:	2017-12-05
Resolution	:	2.15 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 (i) 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 (i)) were used in the production of this report:

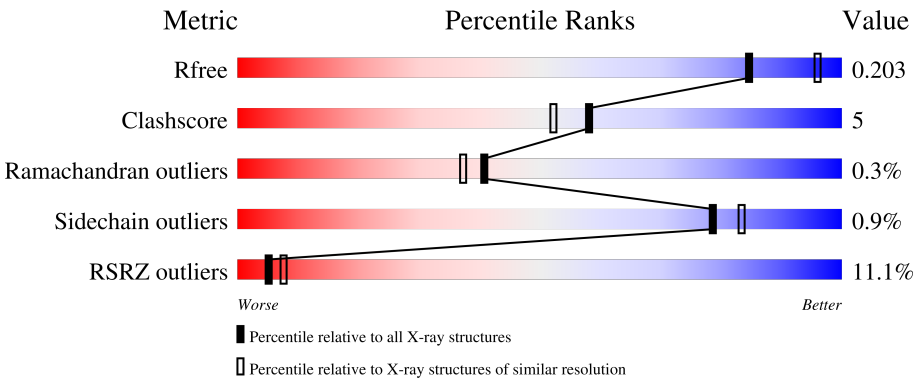
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	330	6% 68% 8% • 23%
2	B	168	12% 67% 11% • 22%
3	C	5	40% 60%

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
4	NAG	A	502	-	-	-	X
4	NAG	A	504	-	-	-	X

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 3308 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 Envelope glycoprotein,Envelope glycoprotein,GP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	253	1972	1259	331	377	5	0	1	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	28	GLU	-	expression tag	UNP Q05320
A	29	THR	-	expression tag	UNP Q05320
A	30	GLY	-	expression tag	UNP Q05320
A	31	ARG	-	expression tag	UNP Q05320
A	42	ALA	THR	engineered mutation	UNP Q05320

- Molecule 2 is a protein called Envelope glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	131	1035	660	177	192	6	0	0	0

There are 38 discrepancies between the modelled and reference sequences:

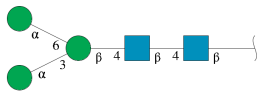
Chain	Residue	Modelled	Actual	Comment	Reference
B	613	ALA	HIS	engineered mutation	UNP A0A0U3BWW0
B	633	GLY	-	expression tag	UNP A0A0U3BWW0
B	634	SER	-	expression tag	UNP A0A0U3BWW0
B	635	GLY	-	expression tag	UNP A0A0U3BWW0
B	636	TYR	-	expression tag	UNP A0A0U3BWW0
B	637	ILE	-	expression tag	UNP A0A0U3BWW0
B	638	PRO	-	expression tag	UNP A0A0U3BWW0
B	639	GLU	-	expression tag	UNP A0A0U3BWW0
B	640	ALA	-	expression tag	UNP A0A0U3BWW0
B	641	PRO	-	expression tag	UNP A0A0U3BWW0
B	642	ARG	-	expression tag	UNP A0A0U3BWW0

Continued on next page...

Continued from previous page...

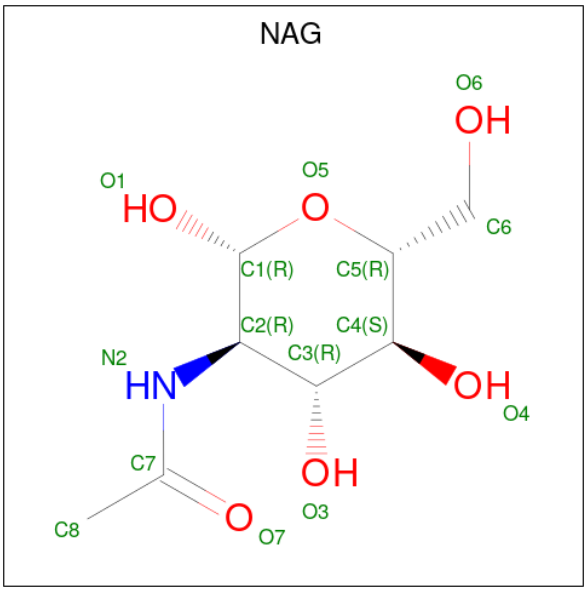
Chain	Residue	Modelled	Actual	Comment	Reference
B	643	ASP	-	expression tag	UNP A0A0U3BWW0
B	644	GLY	-	expression tag	UNP A0A0U3BWW0
B	645	GLN	-	expression tag	UNP A0A0U3BWW0
B	646	ALA	-	expression tag	UNP A0A0U3BWW0
B	647	TYR	-	expression tag	UNP A0A0U3BWW0
B	648	VAL	-	expression tag	UNP A0A0U3BWW0
B	649	ARG	-	expression tag	UNP A0A0U3BWW0
B	650	LYS	-	expression tag	UNP A0A0U3BWW0
B	651	ASP	-	expression tag	UNP A0A0U3BWW0
B	652	GLY	-	expression tag	UNP A0A0U3BWW0
B	653	GLU	-	expression tag	UNP A0A0U3BWW0
B	654	TRP	-	expression tag	UNP A0A0U3BWW0
B	655	VAL	-	expression tag	UNP A0A0U3BWW0
B	656	LEU	-	expression tag	UNP A0A0U3BWW0
B	657	LEU	-	expression tag	UNP A0A0U3BWW0
B	658	SER	-	expression tag	UNP A0A0U3BWW0
B	659	THR	-	expression tag	UNP A0A0U3BWW0
B	660	PHE	-	expression tag	UNP A0A0U3BWW0
B	661	LEU	-	expression tag	UNP A0A0U3BWW0
B	662	GLY	-	expression tag	UNP A0A0U3BWW0
B	663	THR	-	expression tag	UNP A0A0U3BWW0
B	664	HIS	-	expression tag	UNP A0A0U3BWW0
B	665	HIS	-	expression tag	UNP A0A0U3BWW0
B	666	HIS	-	expression tag	UNP A0A0U3BWW0
B	667	HIS	-	expression tag	UNP A0A0U3BWW0
B	668	HIS	-	expression tag	UNP A0A0U3BWW0
B	669	HIS	-	expression tag	UNP A0A0U3BWW0

- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



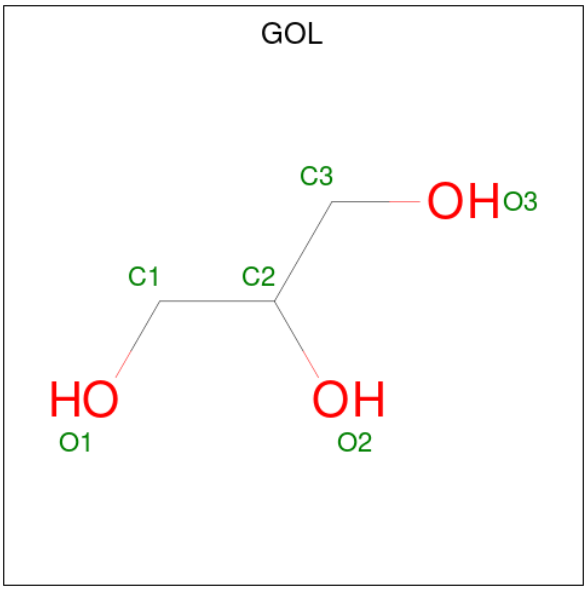
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	5	61	34	2	25	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



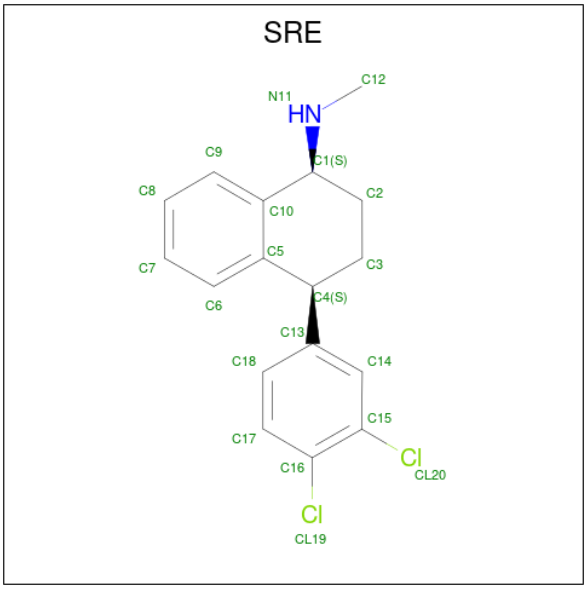
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	14	8	1	5	0	0
			Total	C	N	O		
4	A	1	14	8	1	5	0	0
			Total	C	N	O		
4	A	1	14	8	1	5	0	0
			Total	C	N	O		
4	A	1	14	8	1	5	0	0

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



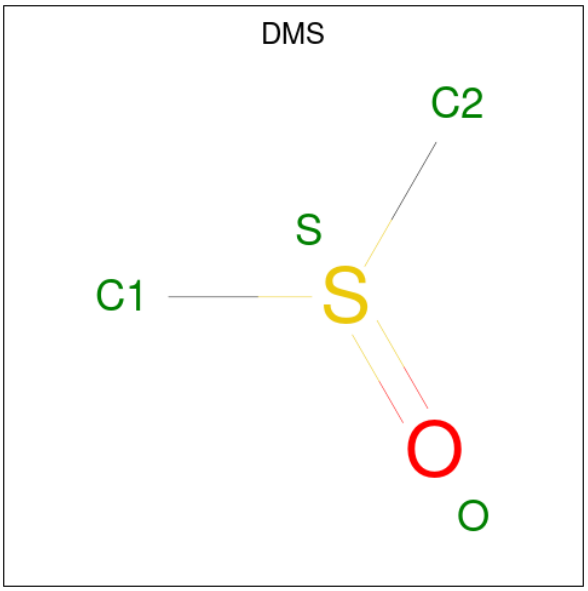
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O			
5	A	1	6	3	3		0	0
			Total	C	O			
5	A	1	6	3	3		0	0
			Total	C	O			
5	A	1	6	3	3		0	0
			Total	C	O			
5	B	1	6	3	3		0	0
			Total	C	O			
5	B	1	6	3	3		0	0

- Molecule 6 is (1S,4S)-4-(3,4-dichlorophenyl)-N-methyl-1,2,3,4-tetrahydronaphthalen-1-amin e (three-letter code: SRE) (formula: $C_{17}H_{17}Cl_2N$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	Cl	N		
6	A	1	20	17	2	1	0	0

- Molecule 7 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
7	B	1	4	2	1	1	0	0

- Molecule 8 is water.

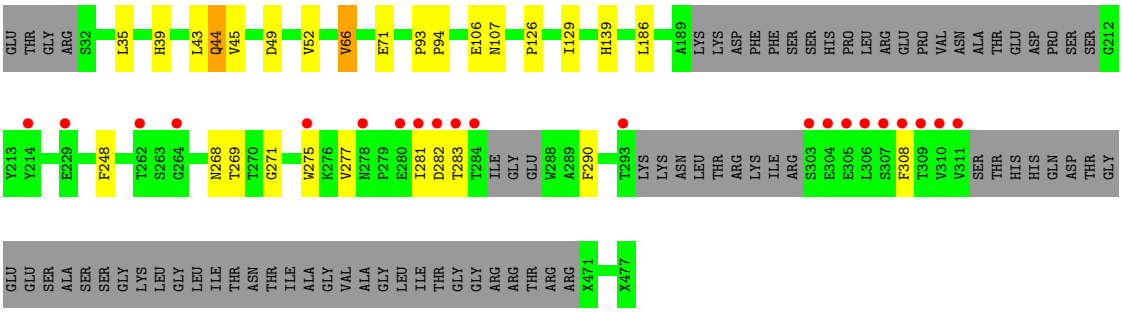
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
8	A	85	85	85	0	0
			Total	O		
8	B	45	45	45	0	0

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.

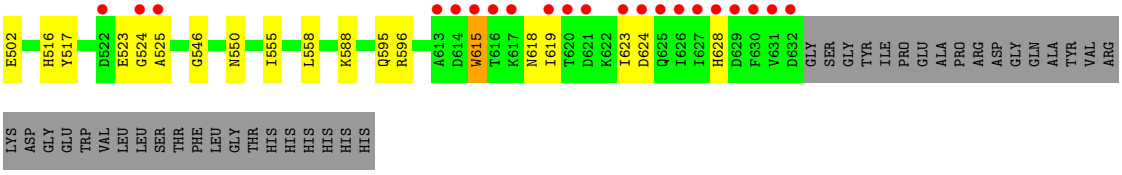
• Molecule 1: Envelope glycoprotein,Envelope glycoprotein,GP1

Chain A:



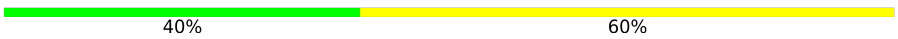
• Molecule 2: Envelope glycoprotein

Chain B:



• Molecule 3: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	114.25Å 114.25Å 306.19Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	60.54 – 2.15 60.54 – 2.15	Depositor EDS
% Data completeness (in resolution range)	100.0 (60.54-2.15) 100.0 (60.54-2.15)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.67 (at 2.14Å)	Xtriage
Refinement program	PHENIX (dev_2940: ???)	Depositor
R, R_{free}	0.184 , 0.203 0.184 , 0.203	Depositor DCC
R_{free} test set	1998 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å^2)	50.6	Xtriage
Anisotropy	0.325	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 59.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3308	wwPDB-VP
Average B, all atoms (Å^2)	75.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.25% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SRE, BMA, DMS, NAG, 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	0.28	0/1986	0.47	0/2699
2	B	0.28	0/1060	0.44	0/1445
All	All	0.28	0/3046	0.46	0/4144

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
2	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	615	TRP	Peptide

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	1972	0	1882	19	0
2	B	1035	0	1001	15	0
3	C	61	0	52	0	0
4	A	56	0	52	0	0
5	A	18	0	24	1	0
5	B	12	0	16	1	0
6	A	20	0	17	4	0
7	B	4	0	6	1	0
8	A	85	0	0	0	0
8	B	45	0	0	0	0
All	All	3308	0	3050	30	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 5.

All (30) 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:45:VAL:HG13	2:B:502:GLU:HG3	1.67	0.77
1:A:248:PHE:HA	1:A:283:THR:HG21	1.70	0.73
2:B:523:GLU:HG2	2:B:524:GLY:H	1.61	0.66
2:B:615:TRP:CD1	2:B:618:ASN:HB3	2.38	0.58
1:A:106:GLU:HG3	1:A:290:PHE:CD1	2.38	0.58
2:B:615:TRP:HD1	2:B:618:ASN:HB3	1.71	0.55
1:A:52:VAL:O	2:B:596:ARG:NH2	2.41	0.54
1:A:71:GLU:OE2	1:A:107:ASN:N	2.41	0.53
2:B:523:GLU:OE1	2:B:523:GLU:N	2.43	0.51
1:A:49:ASP:HA	2:B:595:GLN:HG2	1.92	0.51
1:A:66:VAL:HG22	6:A:508:SRE:H7	1.92	0.51
1:A:282:ASP:OD1	1:A:283:THR:N	2.44	0.50
1:A:35:LEU:HD22	2:B:588:LYS:HD3	1.95	0.49
1:A:43:LEU:H	5:B:708:GOL:C1	2.26	0.48
1:A:186:LEU:HD22	6:A:508:SRE:H17	1.94	0.48
2:B:619:ILE:O	2:B:623:ILE:HG23	2.15	0.47
2:B:517:TYR:CZ	2:B:546:GLY:HA3	2.50	0.46
1:A:39:HIS:HB3	1:A:44:GLN:HG3	1.98	0.46
1:A:277:VAL:HG13	1:A:281:ILE:HD12	1.97	0.46
1:A:126:PRO:HD2	1:A:129:ILE:HD12	1.98	0.44
1:A:107:ASN:OD1	1:A:139:HIS:HE1	2.01	0.43
2:B:550:ASN:HD22	2:B:555:ILE:H	1.67	0.43
2:B:624:ASP:C	2:B:628:HIS:HB2	2.40	0.43
6:A:508:SRE:H2A	2:B:517:TYR:CZ	2.54	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:93:PRO:HA	1:A:94:PRO:HD3	1.96	0.42
6:A:508:SRE:CL20	2:B:558:LEU:HD22	2.57	0.42
1:A:71:GLU:HG3	5:A:506:GOL:H11	2.01	0.42
2:B:516:HIS:HB2	7:B:706:DMS:H22	2.02	0.41
1:A:269:THR:HG22	1:A:271:GLY:H	1.85	0.41
1:A:275:TRP:CH2	1:A:308:PHE:HB3	2.56	0.40

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	239/330 (72%)	232 (97%)	7 (3%)	0	100	100
2	B	129/168 (77%)	118 (92%)	10 (8%)	1 (1%)	19	12
All	All	368/498 (74%)	350 (95%)	17 (5%)	1 (0%)	41	37

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	525	ALA

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	210/273 (77%)	207 (99%)	3 (1%)	67	72
2	B	109/139 (78%)	109 (100%)	0	100	100
All	All	319/412 (77%)	316 (99%)	3 (1%)	78	83

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	44	GLN
1	A	66	VAL
1	A	268	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	73	ASN
1	A	139	HIS
1	A	255	GLN
2	B	508	GLN
2	B	550	ASN
2	B	551	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)

5 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	$\#\|Z\| > 2$
3	NAG	C	1	2,3	14,14,15	0.24	0	17,19,21	0.55	0
3	NAG	C	2	3	14,14,15	0.20	0	17,19,21	0.50	0
3	BMA	C	3	3	11,11,12	0.84	0	15,15,17	1.15	1 (6%)
3	MAN	C	4	3	11,11,12	1.19	2 (18%)	15,15,17	1.09	1 (6%)
3	MAN	C	5	3	11,11,12	0.81	0	15,15,17	0.99	2 (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
3	NAG	C	1	2,3	-	0/6/23/26	0/1/1/1
3	NAG	C	2	3	-	0/6/23/26	0/1/1/1
3	BMA	C	3	3	-	0/2/19/22	0/1/1/1
3	MAN	C	4	3	-	0/2/19/22	0/1/1/1
3	MAN	C	5	3	-	0/2/19/22	1/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	4	MAN	C1-C2	2.26	1.57	1.52
3	C	4	MAN	O5-C5	2.06	1.47	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	5	MAN	C1-O5-C5	2.53	115.62	112.19
3	C	4	MAN	O5-C5-C6	2.27	110.77	107.20
3	C	5	MAN	O2-C2-C3	-2.20	105.73	110.14
3	C	3	BMA	O2-C2-C3	-2.01	106.11	110.14

There are no chirality outliers.

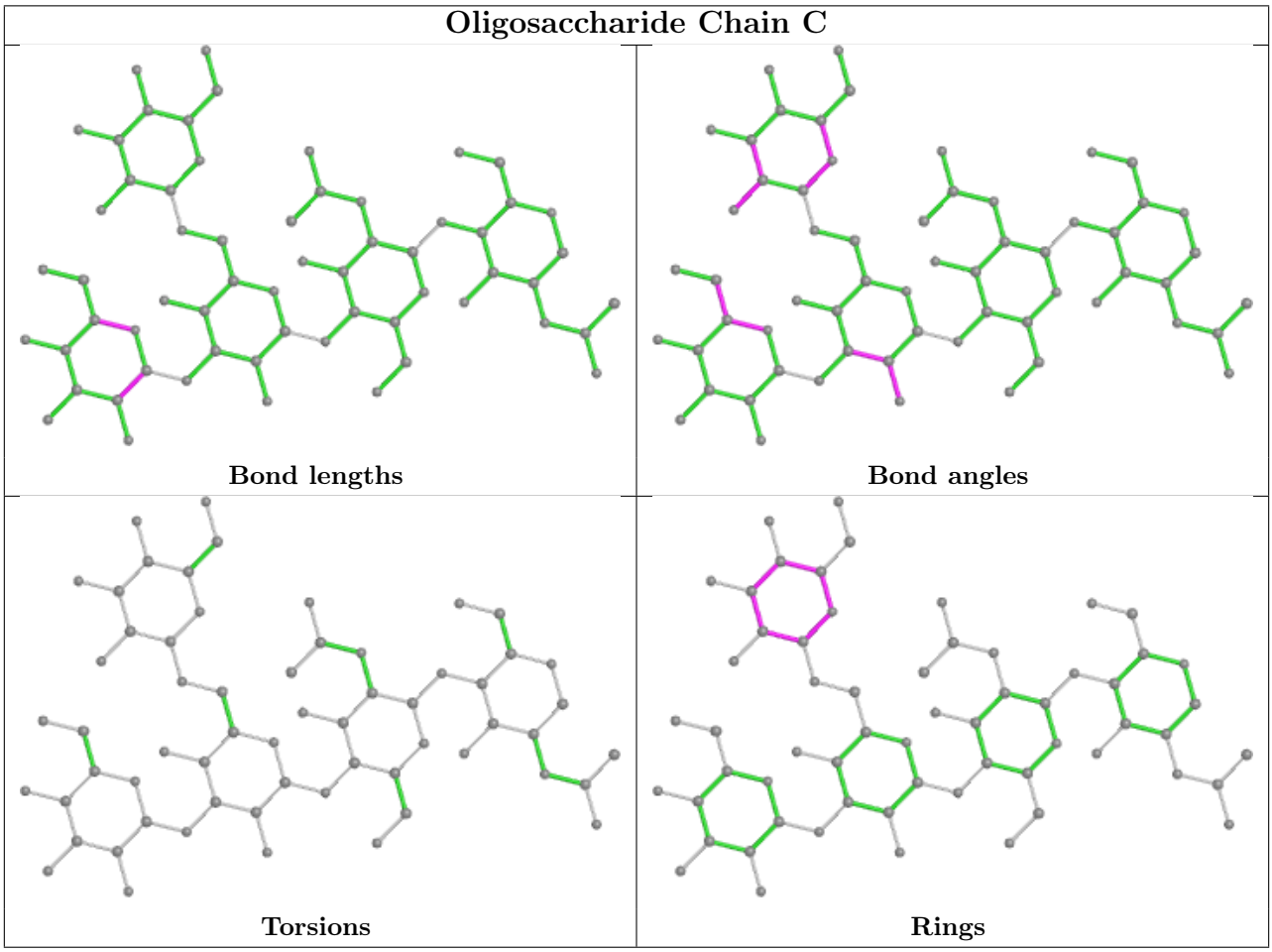
There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	5	MAN	C1-C2-C3-C4-C5-O5

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)

11 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	$\#\|Z\| > 2$
4	NAG	A	501	1	14,14,15	0.21	0	17,19,21	0.51	0
4	NAG	A	502	1	14,14,15	1.50	1 (7%)	17,19,21	1.56	3 (17%)
4	NAG	A	504	1	14,14,15	0.75	1 (7%)	17,19,21	1.17	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	SRE	A	508	-	22,22,22	2.30	5 (22%)	31,31,31	2.34	6 (19%)
5	GOL	A	505	-	5,5,5	0.88	0	5,5,5	1.04	0
4	NAG	A	503	1	14,14,15	0.38	0	17,19,21	0.43	0
5	GOL	A	507	-	5,5,5	0.77	0	5,5,5	0.93	0
5	GOL	B	708	-	5,5,5	0.68	0	5,5,5	1.16	1 (20%)
7	DMS	B	706	-	3,3,3	0.64	0	3,3,3	0.46	0
5	GOL	B	707	-	5,5,5	0.83	0	5,5,5	1.07	0
5	GOL	A	506	-	5,5,5	0.91	0	5,5,5	0.98	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
4	NAG	A	501	1	-	2/6/23/26	0/1/1/1
4	NAG	A	502	1	-	2/6/23/26	0/1/1/1
4	NAG	A	504	1	-	2/6/23/26	0/1/1/1
6	SRE	A	508	-	-	0/5/19/19	0/3/3/3
5	GOL	A	505	-	-	2/4/4/4	-
4	NAG	A	503	1	-	2/6/23/26	0/1/1/1
5	GOL	A	507	-	-	2/4/4/4	-
5	GOL	B	708	-	-	4/4/4/4	-
5	GOL	B	707	-	-	2/4/4/4	-
5	GOL	A	506	-	-	1/4/4/4	-

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	508	SRE	C10-C1	7.87	1.59	1.52
6	A	508	SRE	C5-C4	5.69	1.61	1.51
4	A	502	NAG	O5-C1	-4.94	1.35	1.43
4	A	504	NAG	C1-C2	2.51	1.56	1.52
6	A	508	SRE	C3-C4	-2.44	1.50	1.54
6	A	508	SRE	C15-CL20	2.32	1.79	1.73
6	A	508	SRE	C16-CL19	2.12	1.78	1.73

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	508	SRE	C2-C3-C4	7.46	117.21	110.04
6	A	508	SRE	C2-C1-C10	6.50	120.09	111.34
6	A	508	SRE	C3-C4-C5	4.24	119.49	109.10
4	A	502	NAG	C1-O5-C5	-4.06	106.69	112.19
4	A	504	NAG	C1-O5-C5	4.04	117.67	112.19
4	A	502	NAG	C4-C3-C2	3.97	116.84	111.02
6	A	508	SRE	C13-C4-C5	-3.72	108.49	113.32
6	A	508	SRE	C3-C2-C1	3.61	118.31	110.68
6	A	508	SRE	C3-C4-C13	-3.20	104.66	111.29
4	A	502	NAG	C3-C4-C5	2.43	114.57	110.24
5	B	708	GOL	C3-C2-C1	-2.08	103.63	111.70

There are no chirality outliers.

All (19) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	507	GOL	C1-C2-C3-O3
5	B	707	GOL	O1-C1-C2-C3
5	B	708	GOL	O1-C1-C2-O2
5	B	708	GOL	O1-C1-C2-C3
4	A	501	NAG	C4-C5-C6-O6
4	A	502	NAG	C4-C5-C6-O6
4	A	502	NAG	O5-C5-C6-O6
5	B	707	GOL	O1-C1-C2-O2
4	A	501	NAG	O5-C5-C6-O6
5	A	505	GOL	C1-C2-C3-O3
5	B	708	GOL	C1-C2-C3-O3
5	A	505	GOL	O2-C2-C3-O3
5	A	507	GOL	O2-C2-C3-O3
4	A	504	NAG	C1-C2-N2-C7
4	A	503	NAG	C4-C5-C6-O6
5	B	708	GOL	O2-C2-C3-O3
4	A	503	NAG	O5-C5-C6-O6
4	A	504	NAG	C3-C2-N2-C7
5	A	506	GOL	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	508	SRE	4	0
5	B	708	GOL	1	0

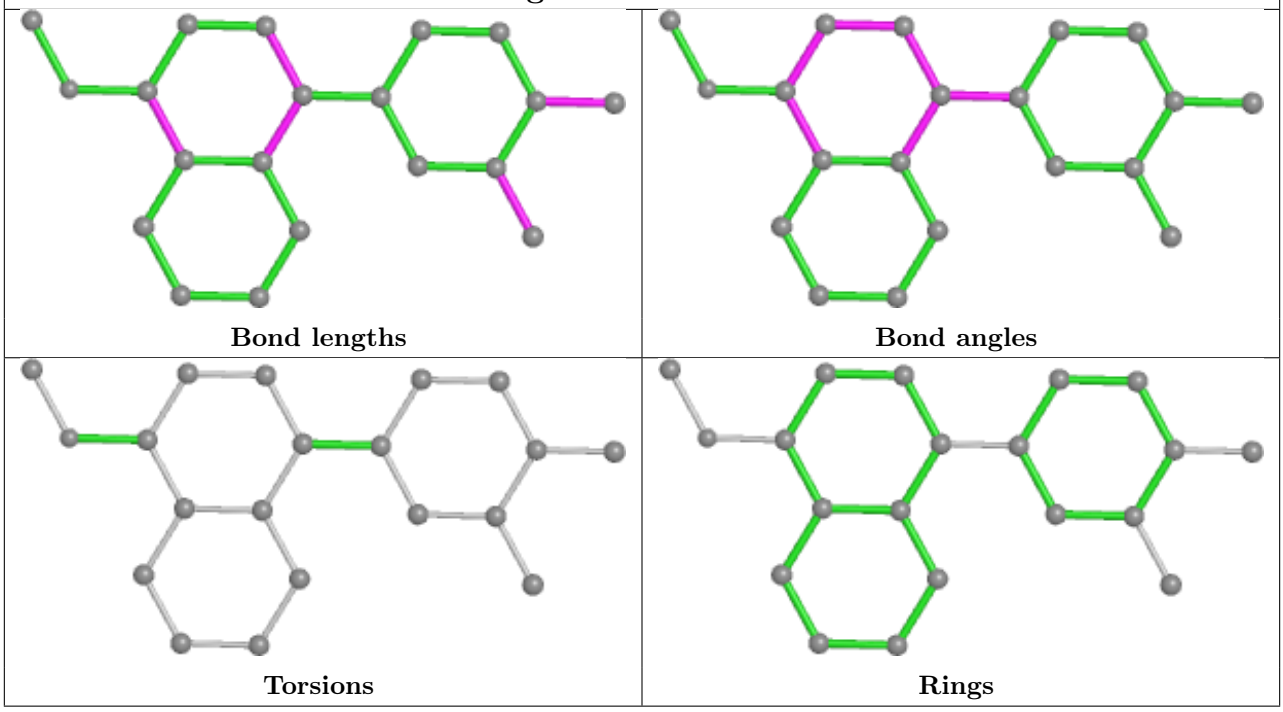
Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	706	DMS	1	0
5	A	506	GOL	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 all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand SRE A 508



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	246/330 (74%)	0.53	21 (8%)	10	15	40, 62, 130, 159	0
2	B	131/168 (77%)	1.30	21 (16%)	1	2	41, 56, 191, 204	0
All	All	377/498 (75%)	0.80	42 (11%)	5	7	40, 59, 159, 204	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	626	ILE	13.3
2	B	620	THR	12.2
2	B	630	PHE	10.4
2	B	627	ILE	9.3
2	B	631	VAL	9.1
2	B	632	ASP	8.3
2	B	621	ASP	8.3
2	B	629	ASP	8.2
2	B	623	ILE	8.1
2	B	616	THR	7.9
1	A	306	LEU	7.4
2	B	619	ILE	7.3
1	A	308	PHE	6.8
1	A	309	THR	6.7
1	A	293	THR	5.9
2	B	615	TRP	5.8
1	A	307	SER	5.7
1	A	310	VAL	5.6
1	A	311	VAL	5.3
2	B	617	LYS	5.3
2	B	522	ASP	5.2
2	B	628	HIS	5.2
2	B	625	GLN	5.1
1	A	303	SER	4.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	624	ASP	4.5
1	A	284	THR	3.8
2	B	613	ALA	3.8
2	B	525	ALA	3.7
1	A	278	ASN	3.6
1	A	262	THR	3.6
1	A	280	GLU	3.5
1	A	283	THR	3.4
1	A	214	TYR	3.2
2	B	614	ASP	3.2
1	A	305	GLU	2.9
1	A	304	GLU	2.8
2	B	524	GLY	2.8
1	A	264	GLY	2.4
1	A	229	GLU	2.3
1	A	281	ILE	2.3
1	A	275	TRP	2.2
1	A	282	ASP	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)

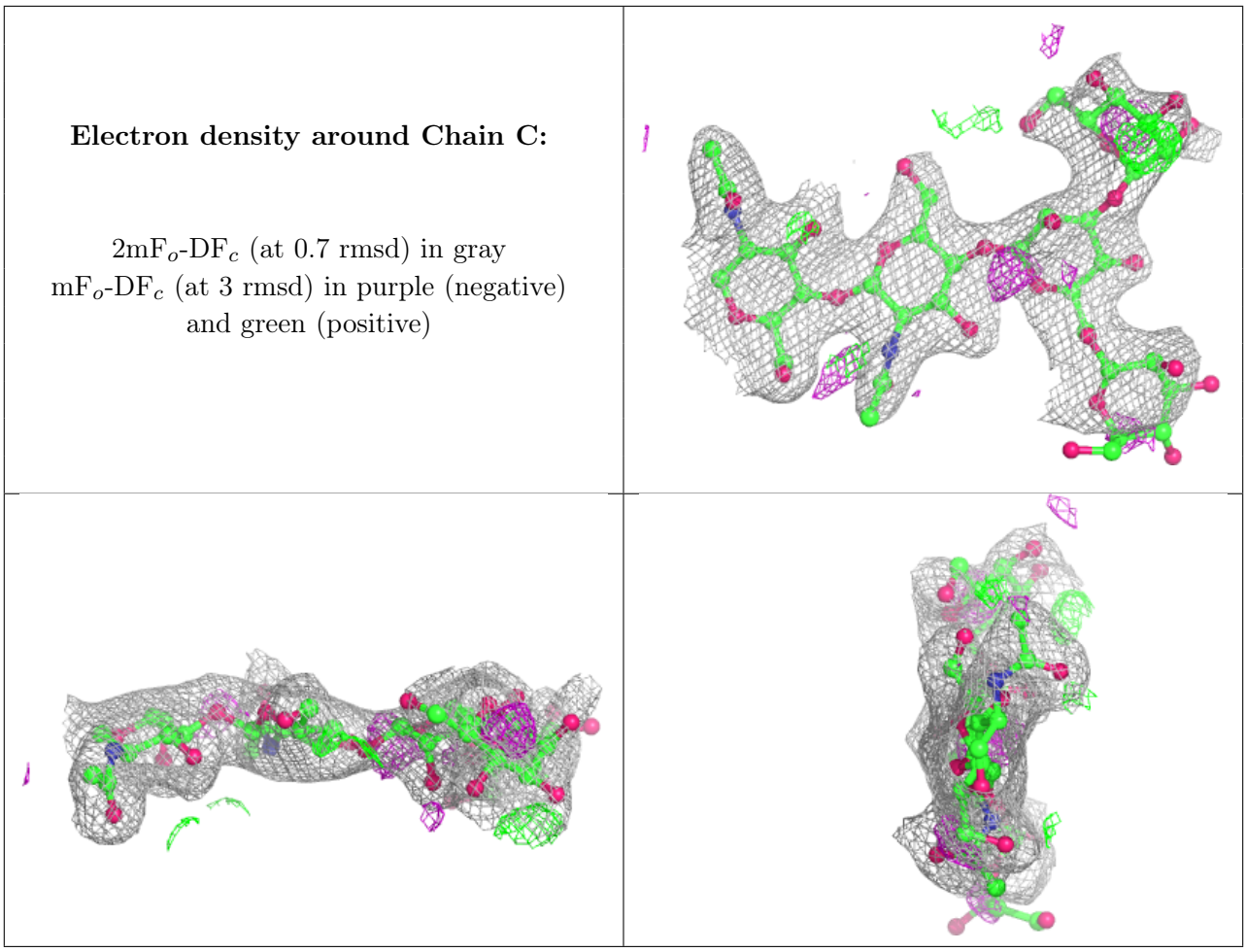
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	MAN	C	4	11/12	0.69	0.31	117,135,140,142	0
3	MAN	C	5	11/12	0.78	0.38	140,145,147,147	0
3	BMA	C	3	11/12	0.83	0.21	106,111,125,135	0
3	NAG	C	2	14/15	0.92	0.14	72,81,91,101	0
3	NAG	C	1	14/15	0.96	0.11	45,52,59,65	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 C:

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



6.4 Ligands (i)

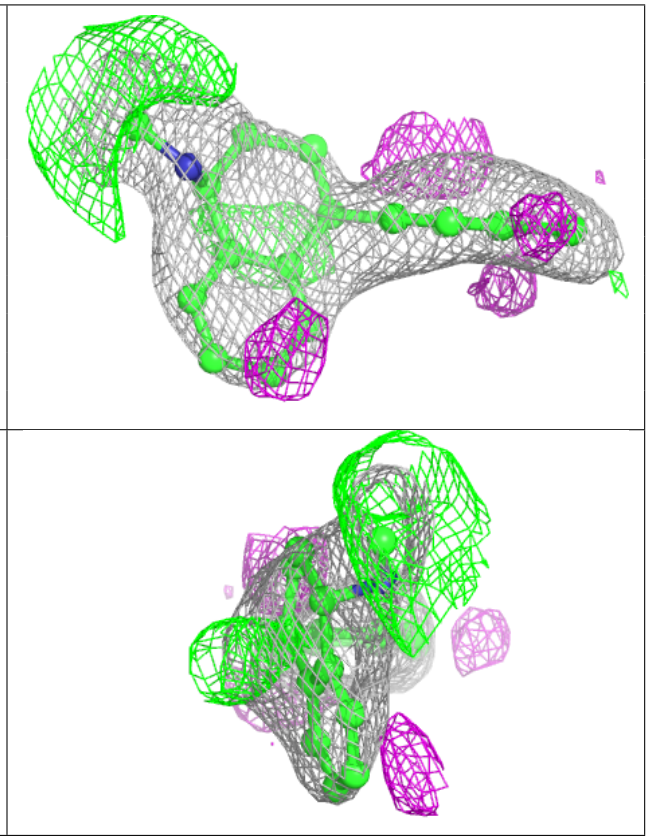
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	A	504	14/15	0.66	0.48	119,134,143,143	0
4	NAG	A	502	14/15	0.71	0.46	129,140,144,146	0
5	GOL	B	708	6/6	0.74	0.23	91,101,102,102	0
5	GOL	A	505	6/6	0.79	0.23	95,105,107,109	0
6	SRE	A	508	20/20	0.83	0.22	94,103,117,120	0
7	DMS	B	706	4/4	0.87	0.36	138,140,140,141	0
4	NAG	A	503	14/15	0.88	0.18	96,108,111,113	0
5	GOL	A	507	6/6	0.89	0.16	94,101,108,110	0
5	GOL	A	506	6/6	0.94	0.30	49,63,67,67	0
5	GOL	B	707	6/6	0.95	0.23	63,71,76,80	0
4	NAG	A	501	14/15	0.96	0.13	56,75,80,80	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around SRE A 508:

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



6.5 Other polymers (i)

There are no such residues in this entry.