



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 12:05 PM BST

PDB ID : 4X13  
Title : JC Polyomavirus genotype 3 VP1 in complex with LSTc pentasaccharide  
Authors : Stroh, L.J.; Stehle, T.  
Deposited on : 2014-11-24  
Resolution : 2.00 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

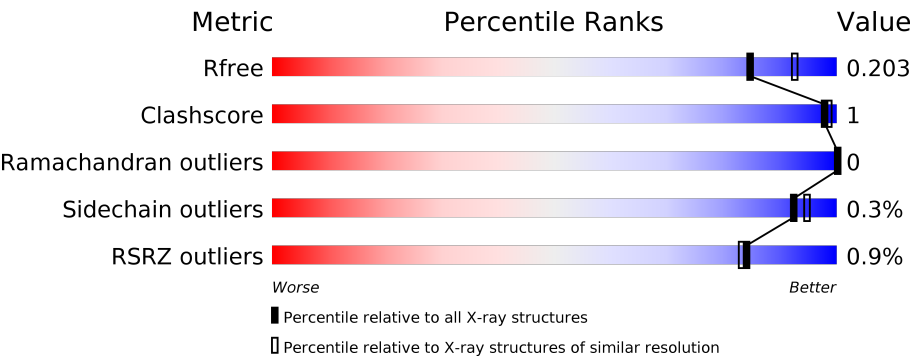
MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
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 2.00 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))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	272	<div><div>%</div><div><div></div><div>93%</div><div>• 5%</div></div></div>
1	B	272	<div><div>%</div><div><div></div><div>94%</div><div>6%</div></div></div>
1	C	272	<div><div>%</div><div><div></div><div>92%</div><div>• 6%</div></div></div>
1	D	272	<div><div>%</div><div><div></div><div>93%</div><div>• 6%</div></div></div>
1	E	272	<div><div>%</div><div><div></div><div>90%</div><div>• 5%</div></div></div>
2	F	5	<div><div></div><div><div>60%</div><div>40%</div></div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	5	 60% 40%
3	H	4	 50% 50%

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 11415 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 Major capsid protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	258	Total	C	N	O	S	0	4	0
			2008	1271	339	387	11			
1	B	256	Total	C	N	O	S	0	1	0
			1987	1256	341	379	11			
1	C	257	Total	C	N	O	S	0	4	0
			2007	1267	342	387	11			
1	D	256	Total	C	N	O	S	0	6	0
			2003	1267	340	385	11			
1	E	258	Total	C	N	O	S	0	3	0
			2017	1272	347	387	11			

There are 20 discrepancies between the modelled and reference sequences:

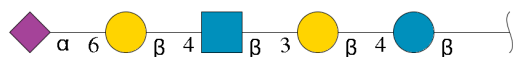
Chain	Residue	Modelled	Actual	Comment	Reference
A	18	GLY	-	expression tag	UNP P90498
A	19	SER	-	expression tag	UNP P90498
A	20	HIS	-	expression tag	UNP P90498
A	21	MET	-	expression tag	UNP P90498
B	18	GLY	-	expression tag	UNP P90498
B	19	SER	-	expression tag	UNP P90498
B	20	HIS	-	expression tag	UNP P90498
B	21	MET	-	expression tag	UNP P90498
C	18	GLY	-	expression tag	UNP P90498
C	19	SER	-	expression tag	UNP P90498
C	20	HIS	-	expression tag	UNP P90498
C	21	MET	-	expression tag	UNP P90498
D	18	GLY	-	expression tag	UNP P90498
D	19	SER	-	expression tag	UNP P90498
D	20	HIS	-	expression tag	UNP P90498
D	21	MET	-	expression tag	UNP P90498
E	18	GLY	-	expression tag	UNP P90498
E	19	SER	-	expression tag	UNP P90498
E	20	HIS	-	expression tag	UNP P90498

*Continued on next page...*

Continued from previous page...

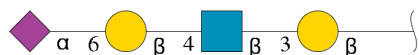
Chain	Residue	Modelled	Actual	Comment	Reference
E	21	MET	-	expression tag	UNP P90498

- Molecule 2 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	F	5	Total	C	N	O	0	0	0
			68	37	2	29			
2	G	5	Total	C	N	O	0	0	0
			68	37	2	29			

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	H	4	Total	C	N	O	0	0	0
			57	31	2	24			

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	1
			12	6	6		
4	D	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		
4	E	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	C	1	Total	C	O	0	0
			4	2	2		
5	D	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	K	0	0
			1	1		
6	A	1	Total	K	0	0
			1	1		
6	D	1	Total	K	0	0
			1	1		
6	C	1	Total	K	0	0
			1	1		
6	E	1	Total	K	0	0
			1	1		

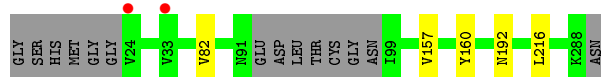
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	198	Total	O	0	0
			198	198		
7	B	204	Total	O	0	0
			204	204		
7	C	222	Total	O	0	0
			222	222		
7	D	222	Total	O	0	0
			222	222		
7	E	217	Total	O	0	0
			217	217		

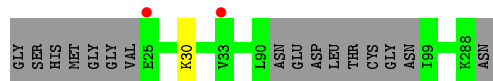
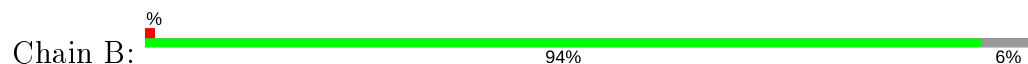
### 3 Residue-property plots [i](#)

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.

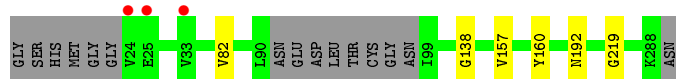
- Molecule 1: Major capsid protein



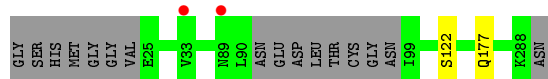
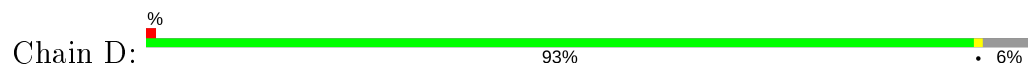
- Molecule 1: Major capsid protein



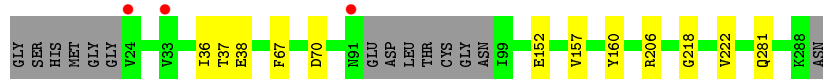
- Molecule 1: Major capsid protein



- Molecule 1: Major capsid protein

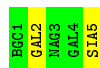


- Molecule 1: Major capsid protein



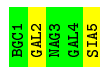
- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain F:  60% 40%



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-beta-D-glucopyranose

Chain G:  60% 40%



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose

Chain H:  50% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	150.12Å 96.28Å 128.14Å 90.00° 109.81° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 29.08 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-2.00) 99.7 (29.08-2.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.21 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.157 , 0.178 0.168 , 0.203	Depositor DCC
$R_{free}$ test set	5778 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	17.3	Xtriage
Anisotropy	0.556	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 49.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11415	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: GOL, BGC, NAG, K, EDO, SIA, GAL

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	0.53	0/2065	0.68	0/2812
1	B	0.53	0/2035	0.67	0/2768
1	C	0.53	0/2062	0.66	0/2807
1	D	0.54	0/2066	0.68	0/2812
1	E	0.54	0/2068	0.69	0/2813
All	All	0.53	0/10296	0.68	0/14012

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	2008	0	1947	2	0
1	B	1987	0	1934	0	0
1	C	2007	0	1943	3	0
1	D	2003	0	1951	2	0
1	E	2017	0	1957	7	0
2	F	68	0	58	0	0
2	G	68	0	58	0	0
3	H	57	0	49	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	12	0	16	0	0
4	B	12	0	16	0	0
4	C	18	0	24	1	0
4	D	24	0	32	0	0
4	E	6	0	8	0	0
5	A	12	0	18	0	0
5	B	24	0	36	0	0
5	C	4	0	6	0	0
5	D	4	0	6	0	0
5	E	16	0	24	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	E	1	0	0	0	0
7	A	198	0	0	0	0
7	B	204	0	0	0	0
7	C	222	0	0	1	0
7	D	222	0	0	1	0
7	E	217	0	0	0	0
All	All	11415	0	10083	14	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 (14) 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:E:218:GLY:O	1:E:222:VAL:HG11	2.02	0.60
1:D:177:GLN:HG2	7:D:562:HOH:O	2.10	0.52
1:E:38:GLU:OE2	1:E:281[A]:GLN:NE2	2.41	0.51
1:A:82:VAL:O	1:A:192[B]:ASN:ND2	2.39	0.50
1:C:82:VAL:O	1:C:192[A]:ASN:ND2	2.42	0.49
1:C:138:GLY:HA2	1:C:219:GLY:O	2.14	0.48
1:E:218:GLY:O	1:E:222:VAL:CG1	2.63	0.45
1:E:152:GLU:HB3	1:E:206[B]:ARG:CZ	2.48	0.44
1:D:122:SER:HB3	1:E:70:ASP:CG	2.39	0.42
1:E:36:ILE:HG22	1:E:37:THR:N	2.34	0.42
1:E:157:VAL:HG21	1:E:160:TYR:HA	2.02	0.41
1:A:157:VAL:HG21	1:A:160:TYR:HA	2.03	0.40
4:C:306:GOL:C1	7:C:512:HOH:O	2.69	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:157:VAL:HG21	1:C:160:TYR:HA	2.03	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	258/272 (95%)	248 (96%)	10 (4%)	0	100	100
1	B	253/272 (93%)	243 (96%)	10 (4%)	0	100	100
1	C	257/272 (94%)	247 (96%)	10 (4%)	0	100	100
1	D	258/272 (95%)	250 (97%)	8 (3%)	0	100	100
1	E	257/272 (94%)	248 (96%)	9 (4%)	0	100	100
All	All	1283/1360 (94%)	1236 (96%)	47 (4%)	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	223/237 (94%)	222 (100%)	1 (0%)	91	93
1	B	220/237 (93%)	219 (100%)	1 (0%)	88	92
1	C	223/237 (94%)	223 (100%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	224/237 (94%)	224 (100%)	0	100	100
1	E	224/237 (94%)	223 (100%)	1 (0%)	91	93
All	All	1114/1185 (94%)	1111 (100%)	3 (0%)	92	95

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

Mol	Chain	Res	Type
1	A	216	LEU
1	B	30	LYS
1	E	67	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	125	GLN
1	E	137	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

9 non-standard protein/DNA/RNA residues 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	G	3	2	14,14,15	0.68	0	17,19,21	0.83	0
2	GAL	G	4	2	11,11,12	0.71	0	15,15,17	0.64	0
3	GAL	H	1	3	12,12,12	0.58	0	17,17,17	0.70	0
2	GAL	F	2	2	11,11,12	0.83	0	15,15,17	1.15	1 (6%)
3	GAL	H	3	3	11,11,12	0.63	0	15,15,17	0.78	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	GAL	G	2	2	11,11,12	0.72	0	15,15,17	1.19	1 (6%)
2	GAL	F	4	2	11,11,12	0.80	0	15,15,17	0.63	0
2	NAG	F	3	2	14,14,15	0.64	0	17,19,21	1.00	0
3	NAG	H	2	3	14,14,15	0.61	0	17,19,21	1.03	2 (11%)

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	G	3	2	-	0/6/23/26	0/1/1/1
2	GAL	G	4	2	-	1/2/19/22	0/1/1/1
3	GAL	H	1	3	-	1/2/22/22	0/1/1/1
2	GAL	F	2	2	-	0/2/19/22	0/1/1/1
3	GAL	H	3	3	-	1/2/19/22	0/1/1/1
2	GAL	G	2	2	-	1/2/19/22	0/1/1/1
2	GAL	F	4	2	-	1/2/19/22	0/1/1/1
2	NAG	F	3	2	-	0/6/23/26	0/1/1/1
3	NAG	H	2	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	2	GAL	C1-C2-C3	2.71	112.99	109.67
2	F	2	GAL	O3-C3-C2	-2.39	105.41	109.99
3	H	2	NAG	O5-C1-C2	-2.30	107.66	111.29
3	H	2	NAG	C1-C2-N2	-2.24	106.66	110.49

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	H	3	GAL	O5-C5-C6-O6
2	G	4	GAL	O5-C5-C6-O6
3	H	1	GAL	C4-C5-C6-O6
2	F	4	GAL	O5-C5-C6-O6
2	G	2	GAL	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates ⓘ

14 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	BGC	F	1	2	12,12,12	0.46	0	17,17,17	0.75	0
2	GAL	F	2	2	11,11,12	0.83	0	15,15,17	1.15	1 (6%)
2	NAG	F	3	2	14,14,15	0.64	0	17,19,21	1.00	0
2	GAL	F	4	2	11,11,12	0.80	0	15,15,17	0.63	0
2	SIA	F	5	2	17,20,21	0.58	0	21,28,31	0.82	1 (4%)
2	BGC	G	1	2	12,12,12	0.46	0	17,17,17	0.74	0
2	GAL	G	2	2	11,11,12	0.72	0	15,15,17	1.19	1 (6%)
2	NAG	G	3	2	14,14,15	0.68	0	17,19,21	0.83	0
2	GAL	G	4	2	11,11,12	0.71	0	15,15,17	0.64	0
2	SIA	G	5	2	17,20,21	0.45	0	21,28,31	0.90	1 (4%)
3	GAL	H	1	3	12,12,12	0.58	0	17,17,17	0.70	0
3	NAG	H	2	3	14,14,15	0.61	0	17,19,21	1.03	2 (11%)
3	GAL	H	3	3	11,11,12	0.63	0	15,15,17	0.78	0
3	SIA	H	4	3	17,20,21	0.42	0	21,28,31	0.78	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	BGC	F	1	2	-	2/2/22/22	0/1/1/1
2	GAL	F	2	2	-	0/2/19/22	0/1/1/1
2	NAG	F	3	2	-	0/6/23/26	0/1/1/1
2	GAL	F	4	2	-	1/2/19/22	0/1/1/1
2	SIA	F	5	2	-	0/14/34/38	0/1/1/1
2	BGC	G	1	2	-	2/2/22/22	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GAL	G	2	2	-	1/2/19/22	0/1/1/1
2	NAG	G	3	2	-	0/6/23/26	0/1/1/1
2	GAL	G	4	2	-	1/2/19/22	0/1/1/1
2	SIA	G	5	2	-	0/14/34/38	0/1/1/1
3	GAL	H	1	3	-	1/2/22/22	0/1/1/1
3	NAG	H	2	3	-	0/6/23/26	0/1/1/1
3	GAL	H	3	3	-	1/2/19/22	0/1/1/1
3	SIA	H	4	3	-	0/14/34/38	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	2	GAL	C1-C2-C3	2.71	112.99	109.67
2	G	5	SIA	C6-O6-C2	2.40	116.47	111.34
2	F	2	GAL	O3-C3-C2	-2.39	105.41	109.99
2	F	5	SIA	C6-O6-C2	2.38	116.44	111.34
3	H	2	NAG	O5-C1-C2	-2.30	107.66	111.29
3	H	2	NAG	C1-C2-N2	-2.24	106.66	110.49
3	H	4	SIA	C6-O6-C2	2.03	115.69	111.34

There are no chirality outliers.

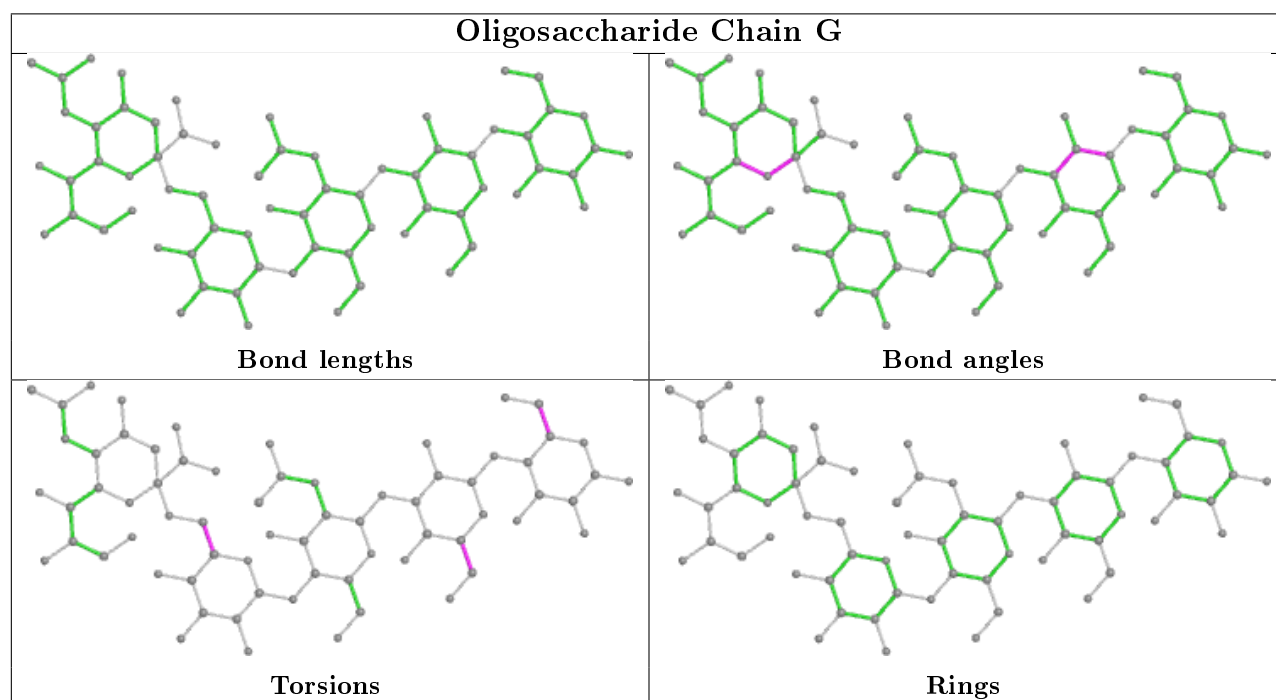
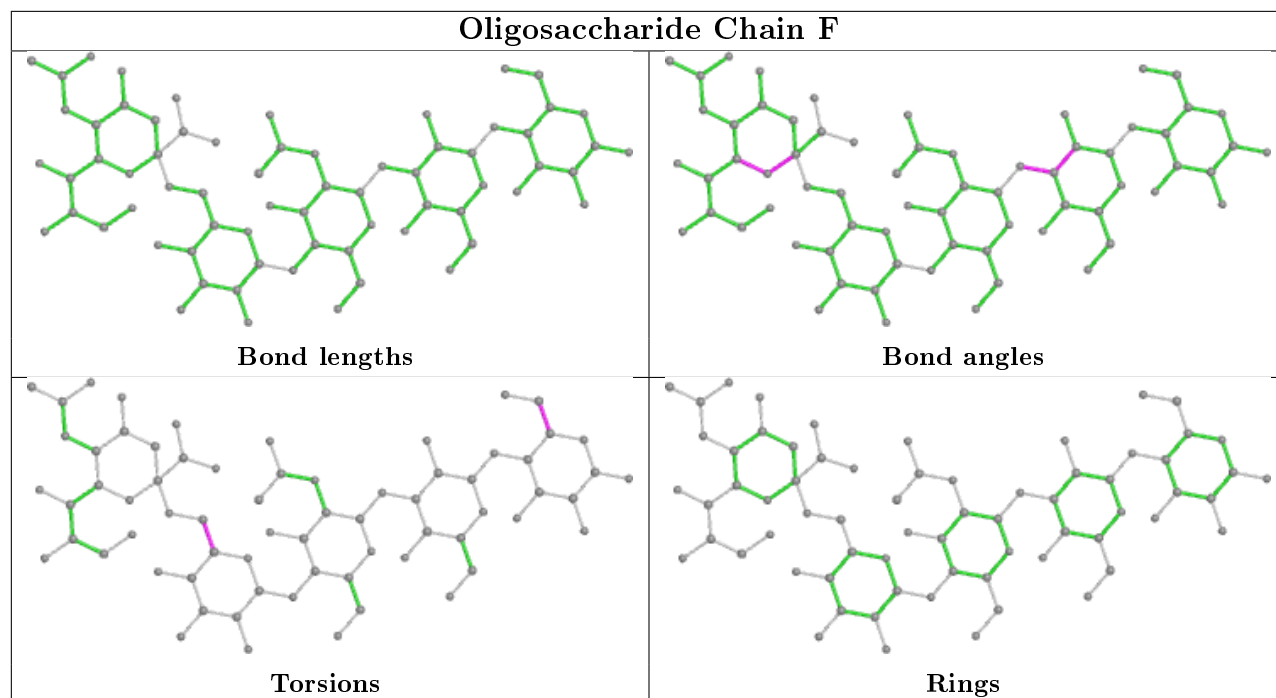
All (9) torsion outliers are listed below:

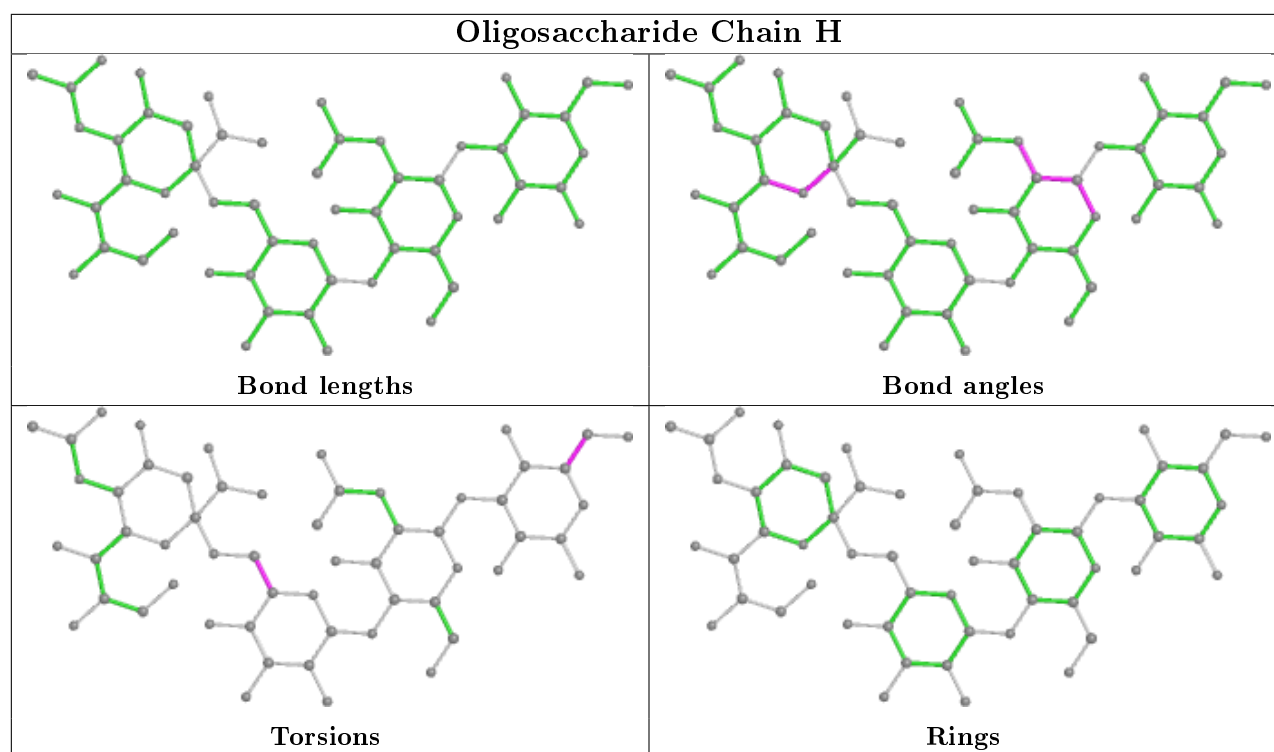
Mol	Chain	Res	Type	Atoms
2	F	1	BGC	O5-C5-C6-O6
2	G	1	BGC	O5-C5-C6-O6
2	F	1	BGC	C4-C5-C6-O6
2	G	1	BGC	C4-C5-C6-O6
3	H	3	GAL	O5-C5-C6-O6
2	G	4	GAL	O5-C5-C6-O6
3	H	1	GAL	C4-C5-C6-O6
2	F	4	GAL	O5-C5-C6-O6
2	G	2	GAL	O5-C5-C6-O6

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](#)

Of 32 ligands modelled in this entry, 5 are monoatomic - leaving 27 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
5	EDO	B	313	-	3,3,3	0.47	0	2,2,2	0.37	0
5	EDO	E	307	-	3,3,3	0.43	0	2,2,2	0.47	0
5	EDO	A	303	-	3,3,3	0.45	0	2,2,2	0.38	0
4	GOL	C	306	-	5,5,5	0.34	0	5,5,5	0.77	0
4	GOL	E	305	-	5,5,5	0.08	0	5,5,5	0.57	0
5	EDO	A	302	-	3,3,3	0.40	0	2,2,2	0.52	0
4	GOL	A	305	-	5,5,5	0.14	0	5,5,5	0.73	0
4	GOL	B	312	-	5,5,5	0.41	0	5,5,5	0.57	0
5	EDO	B	306	-	3,3,3	0.44	0	2,2,2	0.50	0
5	EDO	E	308	-	3,3,3	0.51	0	2,2,2	0.28	0
5	EDO	E	306	-	3,3,3	0.36	0	2,2,2	0.54	0
5	EDO	B	310	-	3,3,3	0.44	0	2,2,2	0.35	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GOL	C	308	-	5,5,5	0.21	0	5,5,5	0.58	0
4	GOL	B	311	-	5,5,5	0.26	0	5,5,5	0.33	0
4	GOL	D	302	-	5,5,5	0.23	0	5,5,5	0.37	0
5	EDO	A	304	-	3,3,3	0.30	0	2,2,2	0.43	0
4	GOL	D	301[B]	-	5,5,5	0.46	0	5,5,5	0.67	0
4	GOL	D	301[A]	-	5,5,5	0.30	0	5,5,5	0.12	0
4	GOL	C	309	-	5,5,5	0.21	0	5,5,5	0.46	0
5	EDO	B	307	-	3,3,3	0.48	0	2,2,2	0.23	0
5	EDO	E	309	-	3,3,3	0.59	0	2,2,2	0.36	0
4	GOL	D	304	-	5,5,5	0.31	0	5,5,5	0.43	0
5	EDO	C	307	-	3,3,3	0.38	0	2,2,2	0.62	0
5	EDO	B	308	-	3,3,3	0.27	0	2,2,2	0.49	0
4	GOL	A	301	-	5,5,5	0.30	0	5,5,5	0.30	0
5	EDO	B	309	-	3,3,3	0.64	0	2,2,2	0.40	0
5	EDO	D	303	-	3,3,3	0.34	0	2,2,2	0.60	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
5	EDO	B	313	-	-	0/1/1/1	-
5	EDO	E	307	-	-	1/1/1/1	-
5	EDO	A	303	-	-	1/1/1/1	-
4	GOL	C	306	-	-	3/4/4/4	-
4	GOL	E	305	-	-	2/4/4/4	-
5	EDO	A	302	-	-	0/1/1/1	-
4	GOL	A	305	-	-	2/4/4/4	-
4	GOL	B	312	-	-	0/4/4/4	-
5	EDO	B	306	-	-	1/1/1/1	-
5	EDO	E	308	-	-	1/1/1/1	-
5	EDO	E	306	-	-	1/1/1/1	-
5	EDO	B	310	-	-	0/1/1/1	-
4	GOL	C	308	-	-	3/4/4/4	-
4	GOL	B	311	-	-	2/4/4/4	-
4	GOL	D	302	-	-	0/4/4/4	-
5	EDO	A	304	-	-	0/1/1/1	-
4	GOL	D	301[B]	-	-	2/4/4/4	-
4	GOL	D	301[A]	-	-	2/4/4/4	-
4	GOL	C	309	-	-	2/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	B	307	-	-	0/1/1/1	-
5	EDO	E	309	-	-	1/1/1/1	-
4	GOL	D	304	-	-	0/4/4/4	-
5	EDO	C	307	-	-	1/1/1/1	-
5	EDO	B	308	-	-	1/1/1/1	-
4	GOL	A	301	-	-	3/4/4/4	-
5	EDO	B	309	-	-	0/1/1/1	-
5	EDO	D	303	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (30) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	311	GOL	O1-C1-C2-C3
4	D	301[B]	GOL	O1-C1-C2-C3
4	D	301[A]	GOL	O2-C2-C3-O3
4	C	309	GOL	C1-C2-C3-O3
4	A	305	GOL	O1-C1-C2-C3
4	A	301	GOL	C1-C2-C3-O3
4	D	301[A]	GOL	C1-C2-C3-O3
4	C	309	GOL	O2-C2-C3-O3
5	E	307	EDO	O1-C1-C2-O2
5	A	303	EDO	O1-C1-C2-O2
5	E	306	EDO	O1-C1-C2-O2
5	C	307	EDO	O1-C1-C2-O2
5	B	306	EDO	O1-C1-C2-O2
4	B	311	GOL	O1-C1-C2-O2
4	D	301[B]	GOL	O1-C1-C2-O2
4	C	306	GOL	O1-C1-C2-C3
4	C	308	GOL	C1-C2-C3-O3
4	A	305	GOL	O1-C1-C2-O2
5	E	308	EDO	O1-C1-C2-O2
5	B	308	EDO	O1-C1-C2-O2
5	E	309	EDO	O1-C1-C2-O2
4	A	301	GOL	O2-C2-C3-O3
5	D	303	EDO	O1-C1-C2-O2
4	E	305	GOL	O1-C1-C2-C3
4	C	308	GOL	O2-C2-C3-O3
4	C	306	GOL	O1-C1-C2-O2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	E	305	GOL	O1-C1-C2-O2
4	C	306	GOL	C1-C2-C3-O3
4	C	308	GOL	O1-C1-C2-C3
4	A	301	GOL	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	306	GOL	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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	258/272 (94%)	-0.45	2 (0%) 86 85	12, 19, 33, 59	0
1	B	256/272 (94%)	-0.51	2 (0%) 86 85	12, 19, 35, 56	0
1	C	257/272 (94%)	-0.53	3 (1%) 79 78	11, 18, 34, 69	0
1	D	256/272 (94%)	-0.54	2 (0%) 86 85	11, 18, 33, 47	0
1	E	258/272 (94%)	-0.47	3 (1%) 79 78	12, 18, 35, 66	0
All	All	1285/1360 (94%)	-0.50	12 (0%) 84 83	11, 19, 34, 69	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	33	VAL	6.2
1	A	33	VAL	6.2
1	C	24	VAL	5.9
1	E	33	VAL	4.8
1	D	33	VAL	4.0
1	C	33	VAL	3.9
1	A	24	VAL	3.1
1	B	25	GLU	2.3
1	E	24	VAL	2.3
1	C	25	GLU	2.3
1	D	89	ASN	2.0
1	E	91	ASN	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	GAL	H	1	12/12	0.90	0.29	34,37,41,44	0
3	GAL	H	3	11/12	0.93	0.13	24,26,30,31	0
2	GAL	G	4	11/12	0.95	0.12	19,22,25,26	0
3	NAG	H	2	14/15	0.95	0.16	27,32,34,35	0
2	GAL	G	2	11/12	0.96	0.16	22,26,29,34	0
2	GAL	F	4	11/12	0.96	0.11	19,21,25,26	0
2	NAG	F	3	14/15	0.96	0.10	22,25,30,33	0
2	GAL	F	2	11/12	0.96	0.14	24,27,30,33	0
2	NAG	G	3	14/15	0.97	0.10	20,22,28,29	0

### 6.3 Carbohydrates

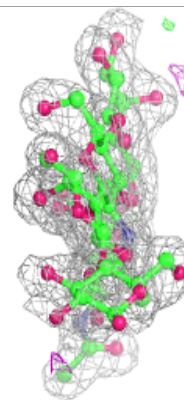
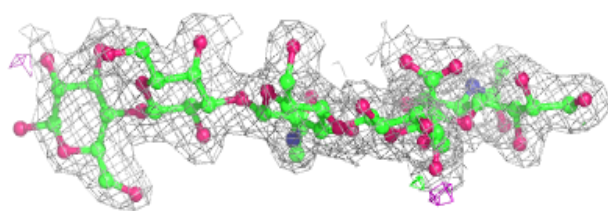
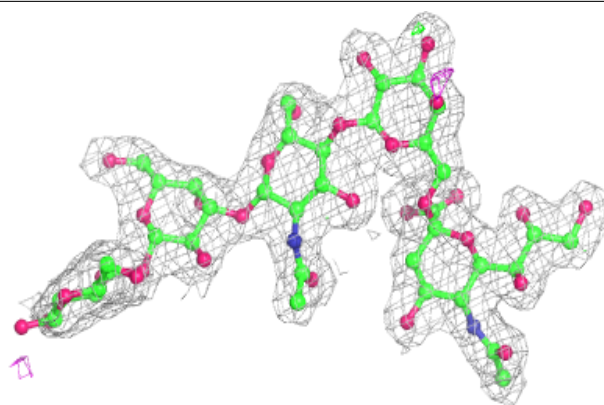
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	BGC	G	1	12/12	0.84	0.30	36,51,60,65	0
2	BGC	F	1	12/12	0.87	0.30	45,56,65,82	0
3	GAL	H	1	12/12	0.90	0.29	34,37,41,44	0
3	SIA	H	4	20/21	0.93	0.12	18,24,29,31	0
3	GAL	H	3	11/12	0.93	0.13	24,26,30,31	0
2	GAL	G	4	11/12	0.95	0.12	19,22,25,26	0
3	NAG	H	2	14/15	0.95	0.16	27,32,34,35	0
2	SIA	F	5	20/21	0.96	0.07	17,19,23,24	0
2	GAL	G	2	11/12	0.96	0.16	22,26,29,34	0
2	GAL	F	4	11/12	0.96	0.11	19,21,25,26	0
2	NAG	F	3	14/15	0.96	0.10	22,25,30,33	0
2	GAL	F	2	11/12	0.96	0.14	24,27,30,33	0
2	NAG	G	3	14/15	0.97	0.10	20,22,28,29	0
2	SIA	G	5	20/21	0.97	0.06	14,17,19,20	0

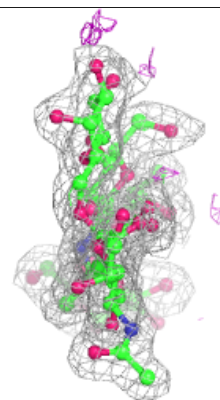
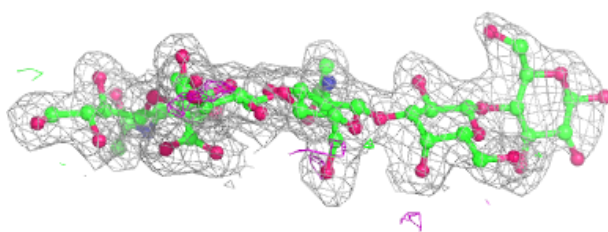
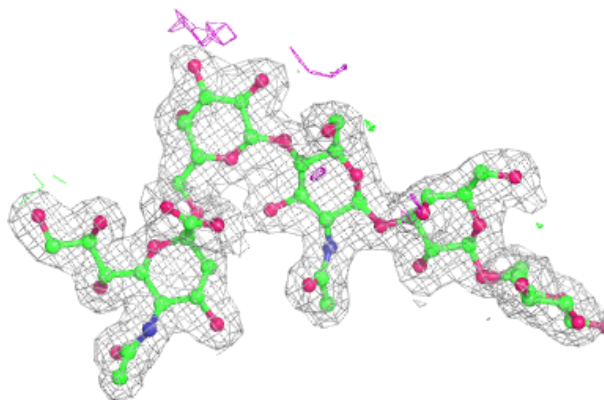
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

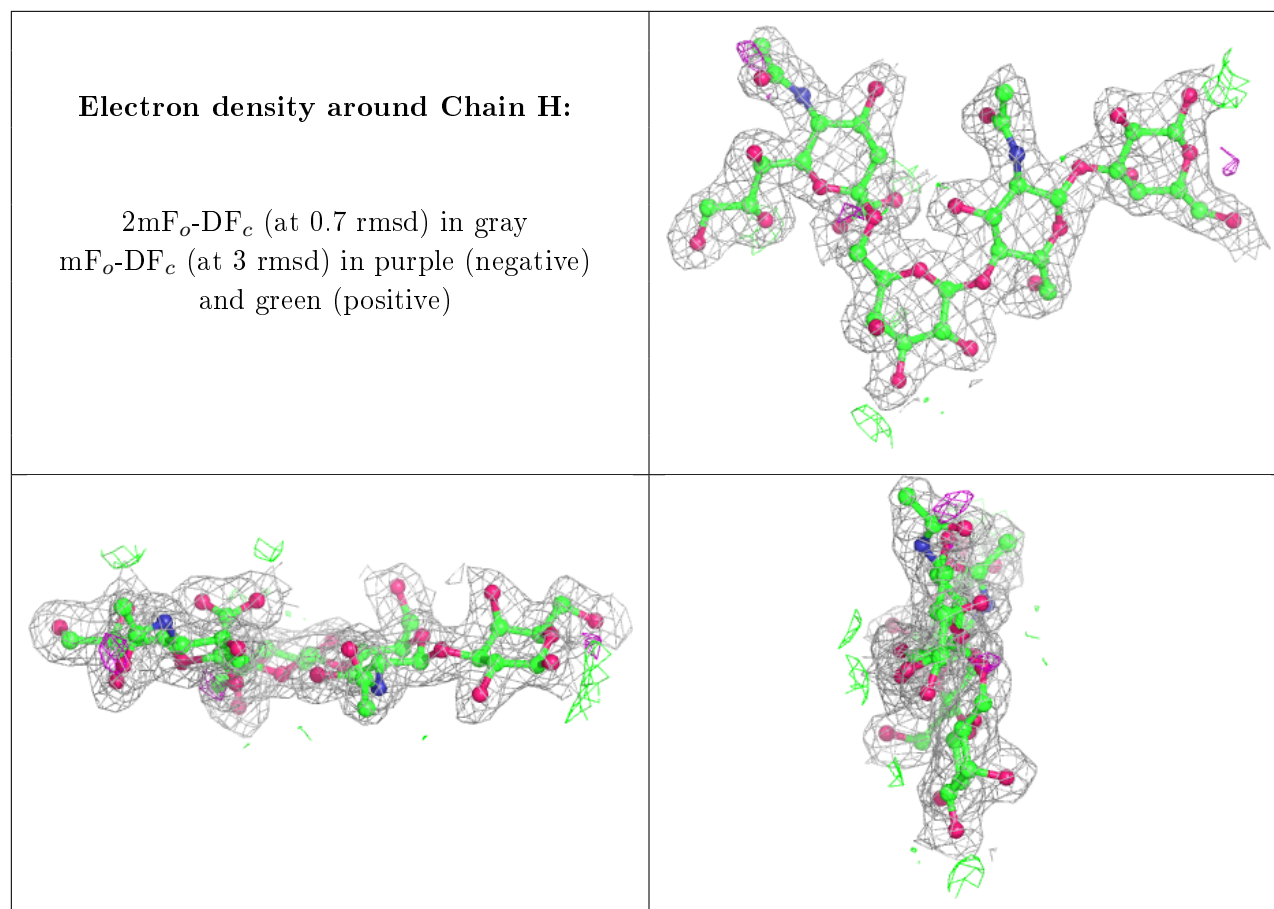
**Electron density around Chain F:**

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

**Electron density around Chain G:**

$2mF_o-DF_c$  (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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
5	EDO	B	307	4/4	0.76	0.15	41,49,51,52	0
4	GOL	C	308	6/6	0.77	0.20	39,44,48,52	0
5	EDO	B	313	4/4	0.79	0.17	48,50,51,52	0
4	GOL	D	304	6/6	0.81	0.12	41,43,43,47	0
4	GOL	B	311	6/6	0.83	0.21	25,32,36,37	6
5	EDO	A	302	4/4	0.84	0.12	41,43,47,55	0
5	EDO	E	309	4/4	0.87	0.14	36,39,42,43	0
4	GOL	C	306	6/6	0.87	0.22	27,34,36,37	0
5	EDO	B	310	4/4	0.88	0.20	41,45,45,46	0
5	EDO	B	308	4/4	0.88	0.20	25,30,30,44	0
5	EDO	D	303	4/4	0.88	0.14	43,46,47,48	0
5	EDO	E	307	4/4	0.89	0.10	40,42,42,51	0
4	GOL	C	309	6/6	0.90	0.18	27,31,33,41	6

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	EDO	E	308	4/4	0.90	0.12	44,45,46,48	0
4	GOL	B	312	6/6	0.90	0.16	27,34,37,38	0
4	GOL	D	302	6/6	0.90	0.16	38,42,45,53	0
5	EDO	C	307	4/4	0.90	0.14	39,44,45,50	0
4	GOL	D	301[B]	6/6	0.90	0.18	27,35,40,42	6
4	GOL	D	301[A]	6/6	0.90	0.18	11,18,20,28	6
4	GOL	A	305	6/6	0.91	0.13	29,31,36,44	6
5	EDO	A	303	4/4	0.91	0.11	41,41,44,45	0
5	EDO	A	304	4/4	0.91	0.17	23,28,29,41	0
5	EDO	B	306	4/4	0.91	0.09	36,39,41,43	0
5	EDO	B	309	4/4	0.91	0.14	21,25,33,39	0
4	GOL	A	301	6/6	0.91	0.23	31,34,37,46	0
4	GOL	E	305	6/6	0.92	0.14	38,40,45,49	0
5	EDO	E	306	4/4	0.93	0.11	44,46,46,46	0
6	K	E	310	1/1	0.97	0.07	25,25,25,25	1
6	K	C	310	1/1	0.98	0.06	21,21,21,21	1
6	K	D	305	1/1	0.99	0.10	25,25,25,25	0
6	K	A	306	1/1	0.99	0.05	18,18,18,18	1
6	K	B	314	1/1	0.99	0.07	28,28,28,28	0

## 6.5 Other polymers

There are no such residues in this entry.