



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2023 – 12:06 AM EDT

PDB ID : 6WQV  
Title : GH5-4 broad specificity endoglucanase from Ruminococcus champanellensis with bound cellotriose  
Authors : Bianchetti, C.M.; Bingman, C.A.; Smith, R.W.; Glasgow, E.M.; Fox, B.G.  
Deposited on : 2020-04-29  
Resolution : 1.45 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

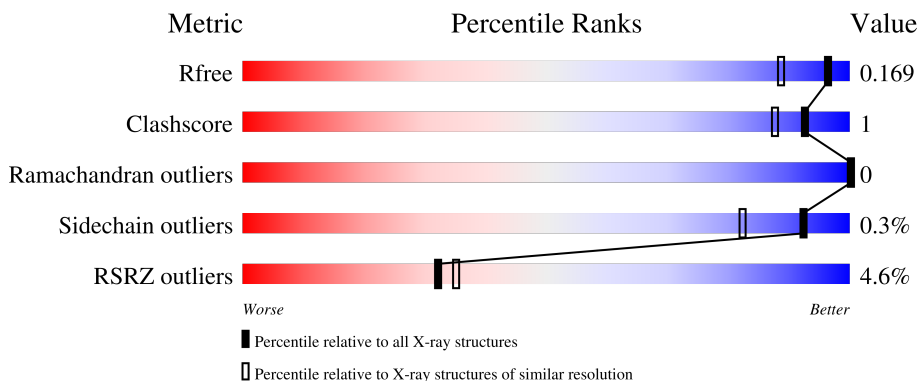
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.45 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.





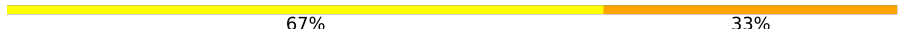
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1156 (1.46-1.46)
Clashscore	141614	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)
RSRZ outliers	127900	1139 (1.46-1.46)

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  $\geq 3$ , 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	354	 2% 95%
1	B	354	 2% 95%
1	C	354	 3% 94%
1	D	354	 12% 93%
2	E	3	 67% 33%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain	
2	F	3	 67%	33%
2	G	3	 67%	33%
2	H	3	 67%	33%

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 23647 atoms, of which 10823 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.

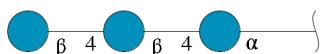
- Molecule 1 is a protein called Endoglucanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	348	5512	1772	2695	481	549	15	0	9	0
1	B	346	5512	1772	2693	481	549	17	0	11	0
1	C	344	5436	1749	2656	475	540	16	0	7	0
1	D	344	5399	1737	2639	471	535	17	0	5	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP D4LAX7
A	164	GLN	GLU	conflict	UNP D4LAX7
A	288	GLN	GLU	conflict	UNP D4LAX7
B	1	SER	-	expression tag	UNP D4LAX7
B	164	GLN	GLU	conflict	UNP D4LAX7
B	288	GLN	GLU	conflict	UNP D4LAX7
C	1	SER	-	expression tag	UNP D4LAX7
C	164	GLN	GLU	conflict	UNP D4LAX7
C	288	GLN	GLU	conflict	UNP D4LAX7
D	1	SER	-	expression tag	UNP D4LAX7
D	164	GLN	GLU	conflict	UNP D4LAX7
D	288	GLN	GLU	conflict	UNP D4LAX7

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	E	3	Total	C	H	O	0	0	0
			66	18	32	16			
2	F	3	Total	C	H	O	0	0	0
			66	18	32	16			
2	G	3	Total	C	H	O	0	0	0
			66	18	32	16			
2	H	3	Total	C	H	O	0	0	0
			66	18	32	16			

- Molecule 3 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
3	A	1	Total	C	H	O	0	0
			10	2	6	2		
3	B	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 4 is NITRATE ION (three-letter code: NO3) (formula: NO<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	N	O	0	0
			4	1	3		
4	C	1	Total	N	O	0	0
			4	1	3		
4	C	1	Total	N	O	0	0
			4	1	3		

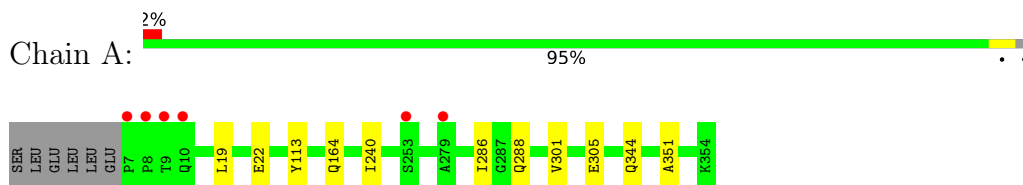
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	482	Total	O	0	21
			503	503		
5	B	461	Total	O	0	11
			472	472		
5	C	296	Total	O	0	4
			300	300		
5	D	214	Total	O	0	3
			217	217		

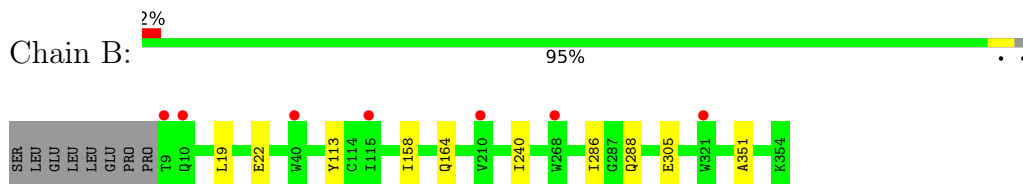
### 3 Residue-property plots [i](#)

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.

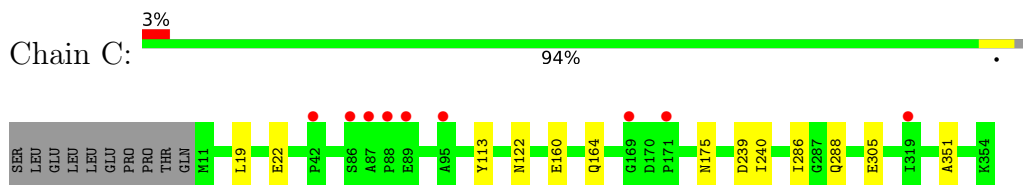
- Molecule 1: Endoglucanase



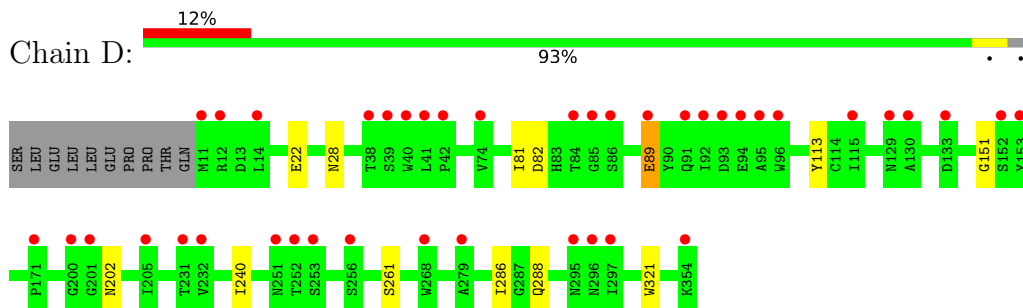
- Molecule 1: Endoglucanase



- Molecule 1: Endoglucanase



- Molecule 1: Endoglucanase



- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose



GLC1  
BGC2  
BGC3

- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain F:  67% 33%

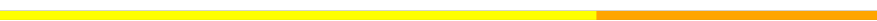
GLC1  
BGC2  
BGC3

- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain G:  67% 33%

GLC1  
BGC2  
BGC3

- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain H:  67% 33%

GLC1  
BGC2  
BGC3



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.61Å 83.17Å 85.21Å 93.98° 93.36° 89.27°	Depositor
Resolution (Å)	36.89 – 1.45 36.89 – 1.45	Depositor EDS
% Data completeness (in resolution range)	93.2 (36.89-1.45) 93.2 (36.89-1.45)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.65 (at 1.45Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.151 , 0.169 0.151 , 0.169	Depositor DCC
$R_{free}$ test set	2075 reflections (0.93%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.4	Xtrriage
Anisotropy	0.305	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 50.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.009 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	23647	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.29% 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: BGC, NO3, EDO, GLC

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	0.31	0/2889	0.57	0/3946
1	B	0.30	0/2889	0.56	0/3943
1	C	0.28	0/2850	0.52	0/3891
1	D	0.28	0/2830	0.50	0/3863
All	All	0.29	0/11458	0.54	0/15643

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	2817	2695	2687	8	0
1	B	2819	2693	2682	7	0
1	C	2780	2656	2649	9	0
1	D	2760	2639	2634	9	0
2	E	34	32	29	2	0
2	F	34	32	29	2	0
2	G	34	32	29	2	0
2	H	34	32	29	1	0
3	A	4	6	6	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	4	6	6	0	0
4	B	4	0	0	0	0
4	C	8	0	0	0	0
5	A	503	0	0	0	0
5	B	472	0	0	0	0
5	C	300	0	0	1	0
5	D	217	0	0	1	0
All	All	12824	10823	10780	33	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 (33) 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:C:22:GLU:OE1	1:C:113:TYR:OH	2.11	0.68
1:D:22:GLU:OE1	1:D:113:TYR:OH	2.12	0.67
1:A:22:GLU:OE1	1:A:113:TYR:OH	2.14	0.64
1:C:305:GLU:HG3	1:C:351:ALA:HB1	1.81	0.62
1:C:288:GLN:NE2	2:G:1:GLC:O1	2.35	0.60
1:C:122:ASN:ND2	5:C:501:HOH:O	2.33	0.59
1:D:288:GLN:NE2	2:H:1:GLC:O1	2.41	0.53
1:D:161:GLY:O	1:D:162[A]:MET:HG3	2.09	0.52
1:B:288:GLN:NE2	2:F:1:GLC:O1	2.42	0.52
1:A:305:GLU:HG3	1:A:351:ALA:HB1	1.92	0.51
1:C:164:GLN:OE1	2:G:1:GLC:H1	2.10	0.51
1:A:164:GLN:OE1	2:E:1:GLC:H1	2.10	0.51
1:A:19:LEU:C	1:A:19:LEU:HD23	2.31	0.50
1:B:305[A]:GLU:HG3	1:B:351:ALA:HB1	1.94	0.50
1:B:164:GLN:OE1	2:F:1:GLC:H1	2.13	0.49
1:D:89:GLU:HB2	5:D:501:HOH:O	2.12	0.48
1:A:305:GLU:CG	1:A:351:ALA:HB1	2.44	0.48
1:C:19:LEU:C	1:C:19:LEU:HD23	2.37	0.45
1:D:151:GLY:O	1:D:202:ASN:HB2	2.17	0.44
1:A:288:GLN:NE2	2:E:1:GLC:O1	2.50	0.43
1:B:19[A]:LEU:HD21	1:B:158:ILE:HD11	2.00	0.43
1:D:162[A]:MET:HE2	1:D:165:PRO:HB3	2.01	0.43
1:B:240:ILE:O	1:B:286:ILE:HA	2.19	0.43
1:A:301:VAL:O	1:A:305:GLU:HG3	2.19	0.43
1:D:28:ASN:HB2	1:D:321:TRP:HA	2.01	0.42
1:B:22:GLU:OE1	1:B:113:TYR:OH	2.28	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:81:ILE:HG23	1:D:82:ASP:N	2.35	0.41
1:D:240:ILE:O	1:D:286:ILE:HA	2.20	0.41
1:C:240:ILE:O	1:C:286:ILE:HA	2.21	0.41
1:A:240:ILE:O	1:A:286:ILE:HA	2.20	0.40
1:B:305[A]:GLU:CG	1:B:351:ALA:HB1	2.51	0.40
1:C:160:GLU:OE2	1:C:239:ASP:OD2	2.39	0.40
1:C:305:GLU:CG	1:C:351:ALA:HB1	2.50	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	355/354 (100%)	344 (97%)	11 (3%)	0	100	100
1	B	355/354 (100%)	344 (97%)	11 (3%)	0	100	100
1	C	349/354 (99%)	338 (97%)	11 (3%)	0	100	100
1	D	347/354 (98%)	333 (96%)	14 (4%)	0	100	100
All	All	1406/1416 (99%)	1359 (97%)	47 (3%)	0	100	100

There are no Ramachandran outliers to report.

#### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	311/308 (101%)	310 (100%)	1 (0%)	92	82
1	B	310/308 (101%)	310 (100%)	0	100	100
1	C	305/308 (99%)	304 (100%)	1 (0%)	92	82
1	D	303/308 (98%)	300 (99%)	3 (1%)	76	52
All	All	1229/1232 (100%)	1224 (100%)	5 (0%)	92	80

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	344	GLN
1	C	175	ASN
1	D	89	GLU
1	D	261[A]	SER
1	D	261[B]	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	288	GLN
1	B	288	GLN
1	C	288	GLN
1	D	288	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](#)

12 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	E	1	2	12,12,12	1.70	2 (16%)	17,17,17	1.60	3 (17%)
2	BGC	E	2	2	11,11,12	1.81	3 (27%)	15,15,17	0.92	0
2	BGC	E	3	2	11,11,12	1.96	2 (18%)	15,15,17	0.82	1 (6%)
2	GLC	F	1	2	12,12,12	1.63	2 (16%)	17,17,17	1.46	2 (11%)
2	BGC	F	2	2	11,11,12	1.71	4 (36%)	15,15,17	1.01	0
2	BGC	F	3	2	11,11,12	2.02	2 (18%)	15,15,17	1.01	1 (6%)
2	GLC	G	1	2	12,12,12	1.67	2 (16%)	17,17,17	1.35	2 (11%)
2	BGC	G	2	2	11,11,12	1.84	5 (45%)	15,15,17	0.79	0
2	BGC	G	3	2	11,11,12	2.10	2 (18%)	15,15,17	1.05	1 (6%)
2	GLC	H	1	2	12,12,12	1.63	2 (16%)	17,17,17	1.30	2 (11%)
2	BGC	H	2	2	11,11,12	1.92	4 (36%)	15,15,17	1.02	1 (6%)
2	BGC	H	3	2	11,11,12	2.11	2 (18%)	15,15,17	1.17	1 (6%)

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	E	1	2	-	0/2/22/22	0/1/1/1
2	BGC	E	2	2	-	0/2/19/22	0/1/1/1
2	BGC	E	3	2	-	0/2/19/22	0/1/1/1
2	GLC	F	1	2	-	0/2/22/22	0/1/1/1
2	BGC	F	2	2	-	0/2/19/22	0/1/1/1
2	BGC	F	3	2	-	0/2/19/22	0/1/1/1
2	GLC	G	1	2	-	0/2/22/22	0/1/1/1
2	BGC	G	2	2	-	0/2/19/22	0/1/1/1
2	BGC	G	3	2	-	0/2/19/22	0/1/1/1
2	GLC	H	1	2	-	0/2/22/22	0/1/1/1
2	BGC	H	2	2	-	0/2/19/22	0/1/1/1
2	BGC	H	3	2	-	0/2/19/22	0/1/1/1

All (32) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	3	BGC	O5-C1	6.10	1.53	1.43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	3	BGC	O5-C1	5.99	1.53	1.43
2	F	3	BGC	O5-C1	5.78	1.52	1.43
2	E	3	BGC	O5-C1	5.51	1.52	1.43
2	H	2	BGC	O5-C1	4.27	1.50	1.43
2	E	1	GLC	O5-C1	4.11	1.53	1.42
2	G	1	GLC	O5-C1	4.11	1.53	1.42
2	F	1	GLC	O5-C1	4.04	1.53	1.42
2	H	1	GLC	O5-C1	4.01	1.52	1.42
2	E	2	BGC	O5-C1	3.68	1.49	1.43
2	G	2	BGC	O5-C1	3.36	1.49	1.43
2	F	2	BGC	O5-C1	3.32	1.49	1.43
2	H	2	BGC	O5-C5	2.99	1.49	1.43
2	E	2	BGC	O5-C5	2.98	1.49	1.43
2	G	2	BGC	O5-C5	2.91	1.49	1.43
2	F	2	BGC	O5-C5	2.78	1.49	1.43
2	E	3	BGC	C2-C3	-2.69	1.48	1.52
2	F	3	BGC	C2-C3	-2.67	1.48	1.52
2	G	3	BGC	C2-C3	-2.66	1.48	1.52
2	H	3	BGC	C2-C3	-2.56	1.48	1.52
2	G	1	GLC	O3-C3	2.43	1.48	1.43
2	G	2	BGC	C2-C3	-2.41	1.49	1.52
2	H	1	GLC	O3-C3	2.26	1.48	1.43
2	G	2	BGC	O4-C4	2.22	1.48	1.43
2	F	1	GLC	O3-C3	2.20	1.48	1.43
2	E	1	GLC	O3-C3	2.19	1.48	1.43
2	E	2	BGC	O4-C4	2.18	1.48	1.43
2	F	2	BGC	C2-C3	-2.13	1.49	1.52
2	H	2	BGC	O3-C3	2.12	1.48	1.43
2	H	2	BGC	O4-C4	2.10	1.47	1.43
2	F	2	BGC	O4-C4	2.05	1.47	1.43
2	G	2	BGC	O3-C3	2.02	1.47	1.43

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	1	GLC	C1-C2-C3	4.51	119.67	110.31
2	F	1	GLC	C1-C2-C3	3.92	118.46	110.31
2	H	3	BGC	C1-C2-C3	3.50	113.97	109.67
2	G	1	GLC	C1-C2-C3	3.39	117.35	110.31
2	G	3	BGC	C1-C2-C3	3.19	113.59	109.67
2	F	3	BGC	C1-C2-C3	3.19	113.59	109.67
2	H	1	GLC	O5-C1-C2	2.97	115.58	110.28

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	1	GLC	C1-C2-C3	2.94	116.40	110.31
2	E	1	GLC	O1-C1-C2	2.83	117.01	109.03
2	F	1	GLC	O3-C3-C2	-2.40	104.80	110.35
2	E	3	BGC	C1-C2-C3	2.24	112.42	109.67
2	H	2	BGC	C1-O5-C5	2.09	115.03	112.19
2	G	1	GLC	O5-C1-C2	2.09	114.01	110.28
2	E	1	GLC	C3-C4-C5	-2.01	106.65	110.24

There are no chirality outliers.

There are no torsion outliers.

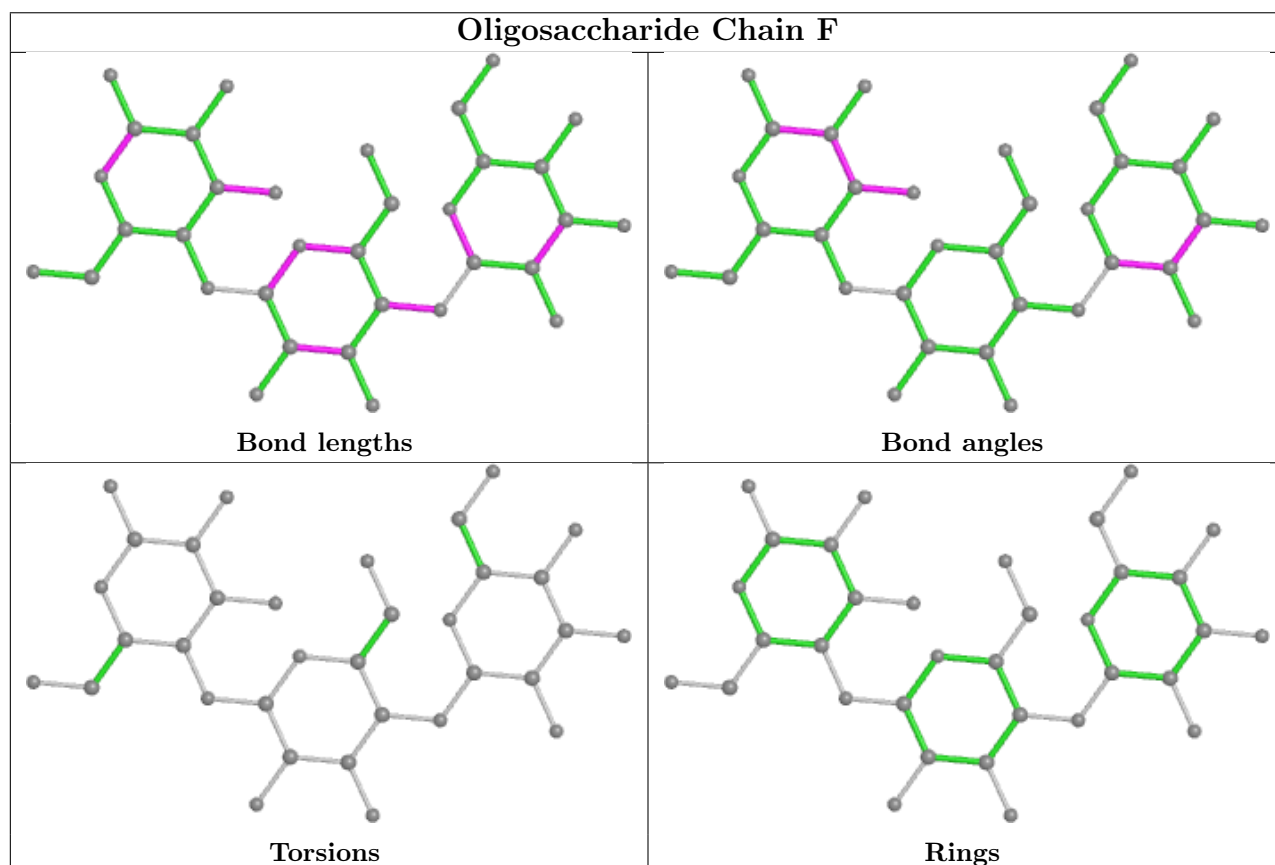
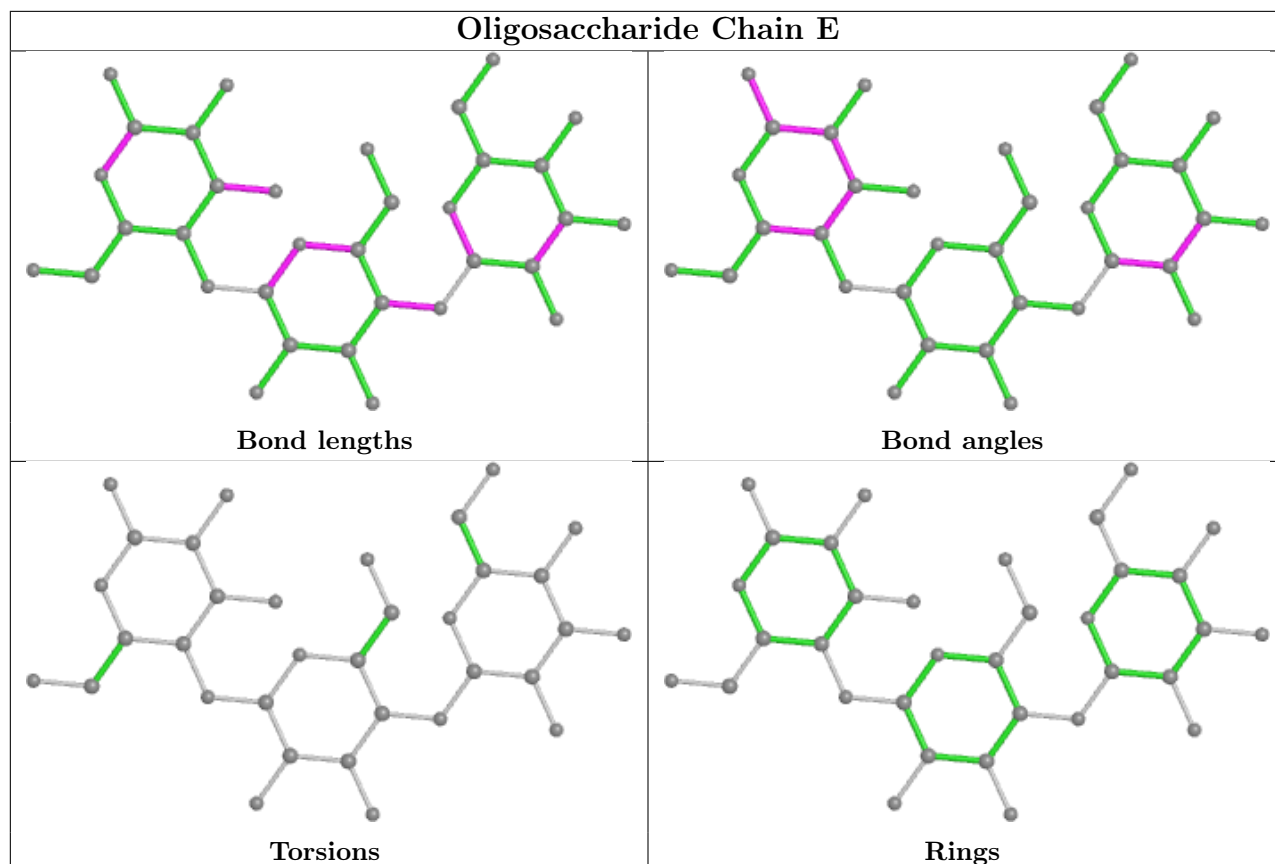
There are no ring outliers.

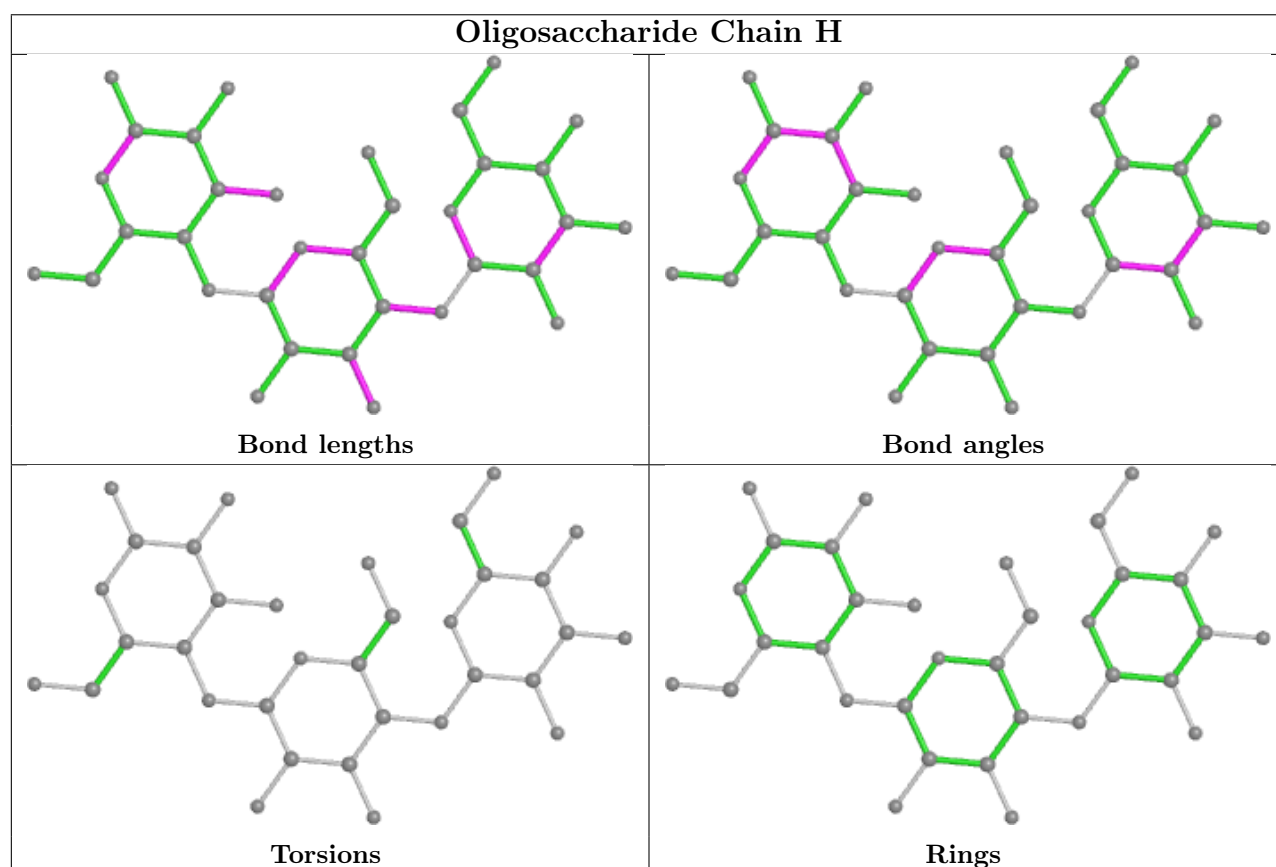
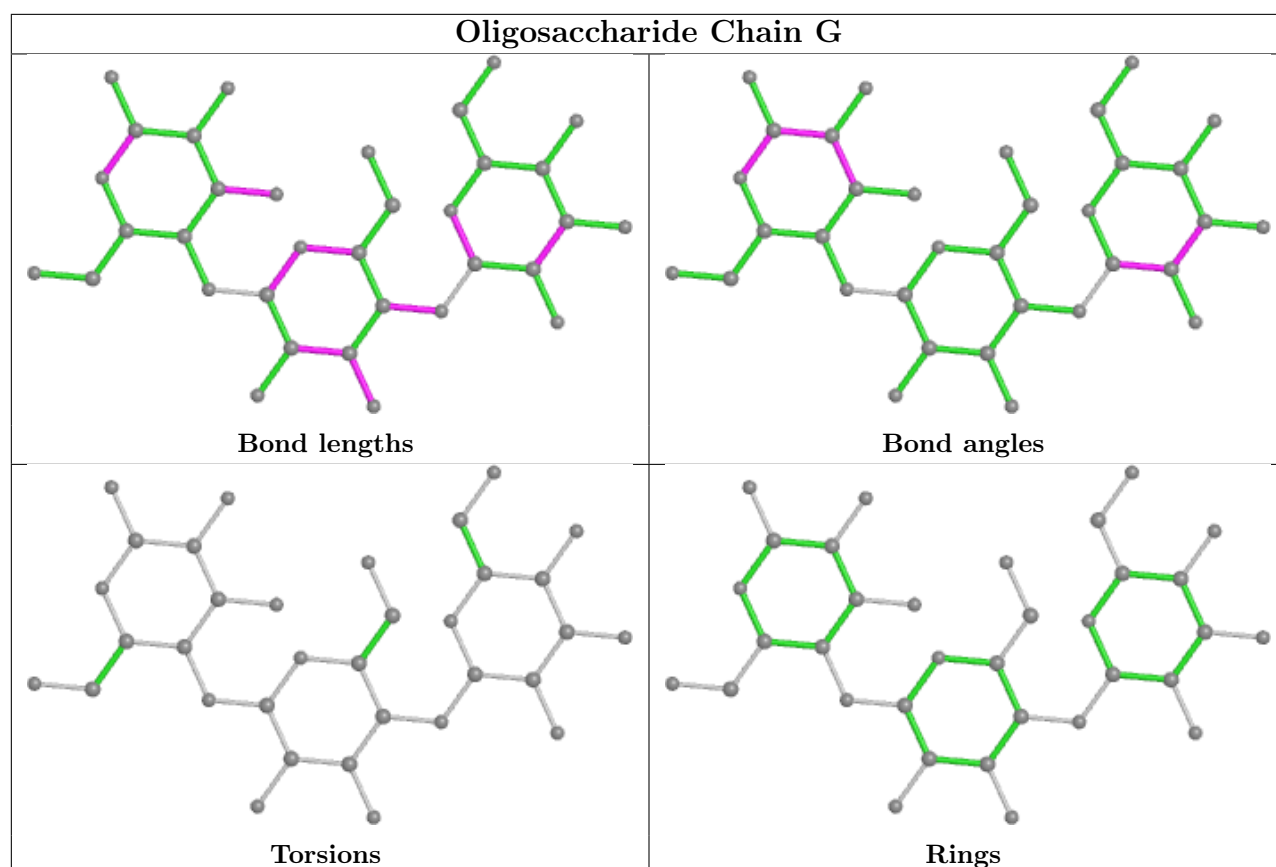
4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	1	GLC	2	0
2	G	1	GLC	2	0
2	E	1	GLC	2	0
2	H	1	GLC	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.







## 5.6 Ligand geometry [i](#)

5 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	NO3	C	403	-	1,3,3	0.60	0	0,3,3	-	-
4	NO3	B	402	-	1,3,3	0.50	0	0,3,3	-	-
3	EDO	A	401	-	3,3,3	0.45	0	2,2,2	0.48	0
4	NO3	C	402	-	1,3,3	0.51	0	0,3,3	-	-
3	EDO	B	401	-	3,3,3	0.47	0	2,2,2	0.36	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	B	401	-	-	0/1/1/1	-
3	EDO	A	401	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

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 [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	348/354 (98%)	-0.10	6 (1%) 70 70	10, 15, 27, 46	0
1	B	346/354 (97%)	0.01	7 (2%) 65 67	11, 17, 28, 51	0
1	C	344/354 (97%)	0.13	9 (2%) 56 58	13, 24, 38, 51	0
1	D	344/354 (97%)	0.70	41 (11%) 4 5	16, 33, 47, 63	0
All	All	1382/1416 (97%)	0.18	63 (4%) 32 35	10, 21, 42, 63	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	9	THR	5.7
1	D	201	GLY	5.5
1	D	232	VAL	4.8
1	C	88	PRO	4.5
1	D	92	ILE	4.5
1	B	10	GLN	4.4
1	D	205	ILE	4.1
1	D	85	GLY	4.0
1	D	40	TRP	4.0
1	A	8	PRO	3.9
1	D	279	ALA	3.9
1	D	14	LEU	3.8
1	A	7	PRO	3.6
1	C	171	PRO	3.6
1	D	89	GLU	3.5
1	D	39	SER	3.4
1	A	9	THR	3.3
1	D	41	LEU	3.2
1	D	74	VAL	3.2
1	D	12	ARG	3.0
1	D	297	ILE	3.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	115	ILE	3.0
1	C	87	ALA	2.9
1	D	231	THR	2.9
1	C	89	GLU	2.8
1	D	11	MET	2.8
1	D	130	ALA	2.8
1	D	252	THR	2.8
1	D	115	ILE	2.7
1	D	93	ASP	2.7
1	B	40	TRP	2.6
1	D	133	ASP	2.6
1	C	95	ALA	2.5
1	D	253	SER	2.5
1	D	354	LYS	2.5
1	D	94	GLU	2.5
1	A	10	GLN	2.5
1	D	42	PRO	2.5
1	D	38	THR	2.5
1	B	268	TRP	2.5
1	D	295	ASN	2.5
1	D	152	SER	2.4
1	B	210	VAL	2.4
1	D	95	ALA	2.4
1	D	268	TRP	2.4
1	D	84	THR	2.3
1	C	86	SER	2.2
1	D	96	TRP	2.2
1	A	253	SER	2.2
1	D	153	TYR	2.2
1	D	251	ASN	2.2
1	B	321	TRP	2.2
1	C	42	PRO	2.2
1	C	169	GLY	2.1
1	D	296	ASN	2.1
1	D	86	SER	2.1
1	D	129	ASN	2.1
1	D	256	SER	2.1
1	D	91	GLN	2.1
1	D	200	GLY	2.1
1	C	319	ILE	2.1
1	A	279	ALA	2.1
1	D	171	PRO	2.0

## 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](#)

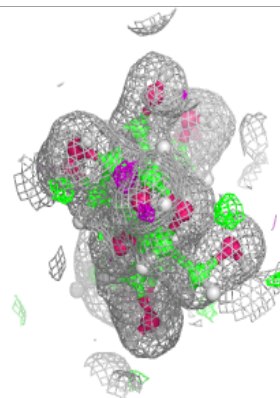
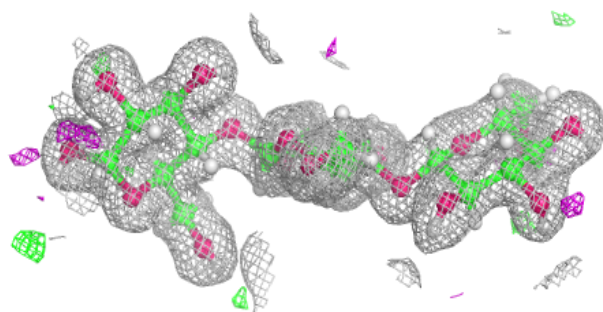
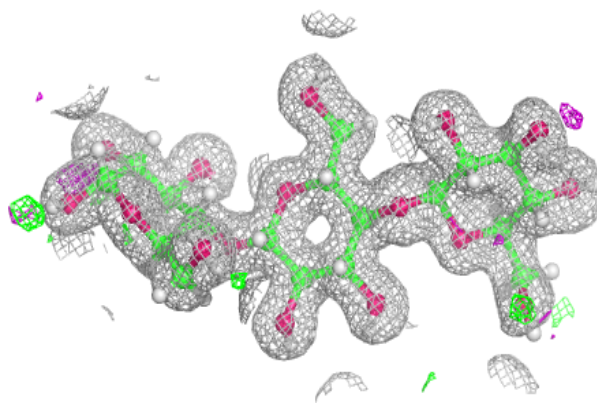
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
2	GLC	H	1	12/12	0.83	0.13	22,30,37,37	0
2	BGC	G	3	11/12	0.84	0.12	30,38,44,46	0
2	BGC	H	2	11/12	0.85	0.11	27,31,36,38	0
2	BGC	H	3	11/12	0.85	0.17	34,44,51,56	0
2	BGC	F	3	11/12	0.88	0.14	24,35,42,45	0
2	BGC	E	3	11/12	0.90	0.10	20,31,36,39	0
2	GLC	G	1	12/12	0.91	0.11	17,23,28,30	0
2	BGC	G	2	11/12	0.92	0.09	19,22,25,30	0
2	BGC	F	2	11/12	0.95	0.08	14,17,21,25	0
2	GLC	F	1	12/12	0.95	0.10	12,18,23,27	0
2	GLC	E	1	12/12	0.96	0.09	11,14,19,23	0
2	BGC	E	2	11/12	0.97	0.06	12,14,17,23	0

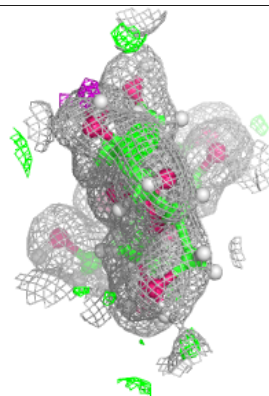
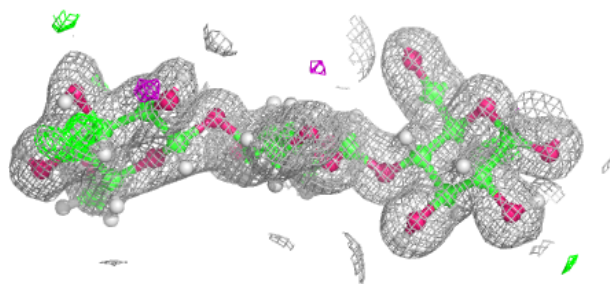
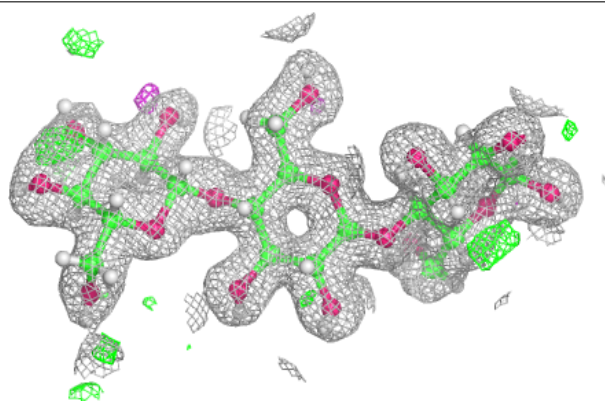
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain E:**

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

**Electron density around Chain F:**

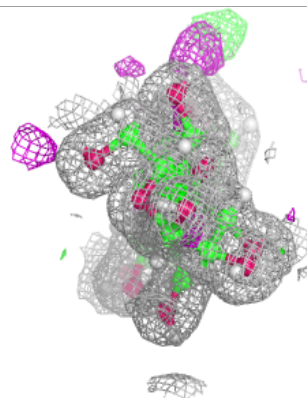
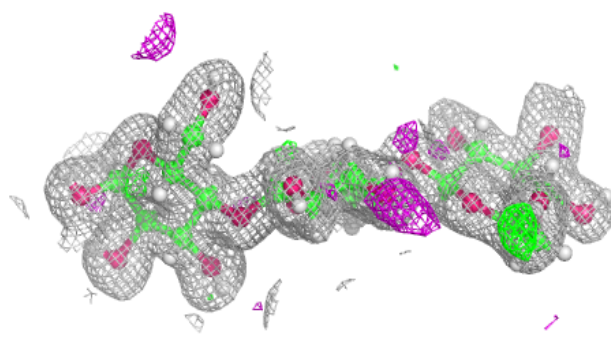
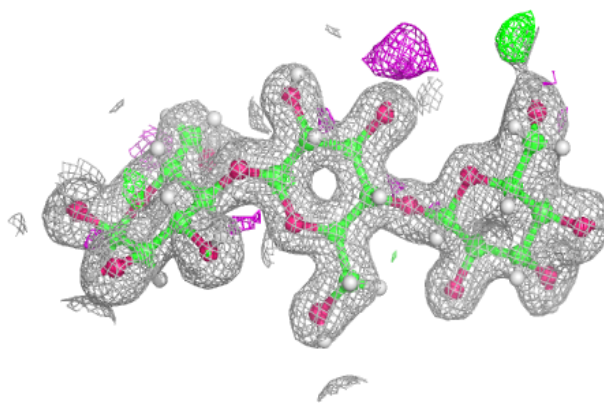
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



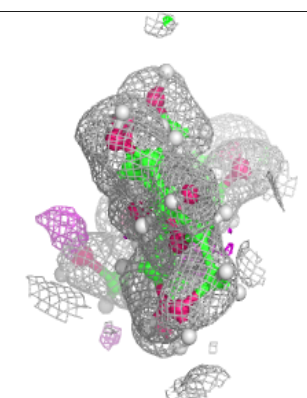
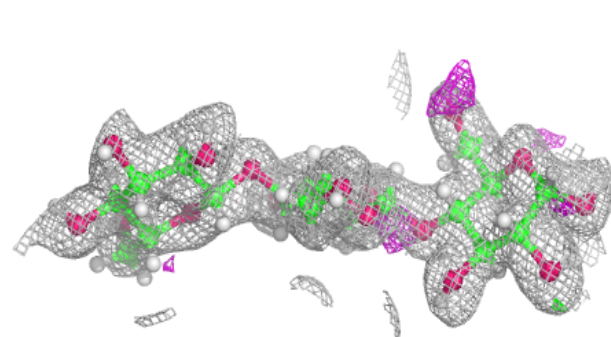
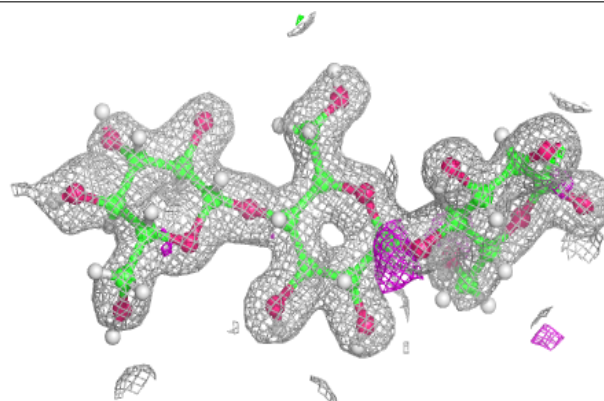


**Electron density around Chain G:**

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

**Electron density around Chain H:**

$2mF_o-DF_c$  (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, 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
3	EDO	B	401	4/4	0.79	0.14	52,62,66,67	0
3	EDO	A	401	4/4	0.82	0.29	33,42,51,51	0
4	NO3	C	403	4/4	0.87	0.18	32,33,33,34	0
4	NO3	C	402	4/4	0.98	0.07	17,17,18,20	0
4	NO3	B	402	4/4	0.99	0.06	17,19,21,21	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.