



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 15, 2023 – 08:24 AM EDT

PDB ID : 1U0A  
Title : Crystal structure of the engineered beta-1,3-1,4-endoglucanase H(A16-M) in complex with beta-glucan tetrasaccharide  
Authors : Gaiser, O.J.; Piotukh, K.; Ponnuswamy, M.N.; Planas, A.; Borriss, R.; Heine-mann, U.  
Deposited on : 2004-07-13  
Resolution : 1.64 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  
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

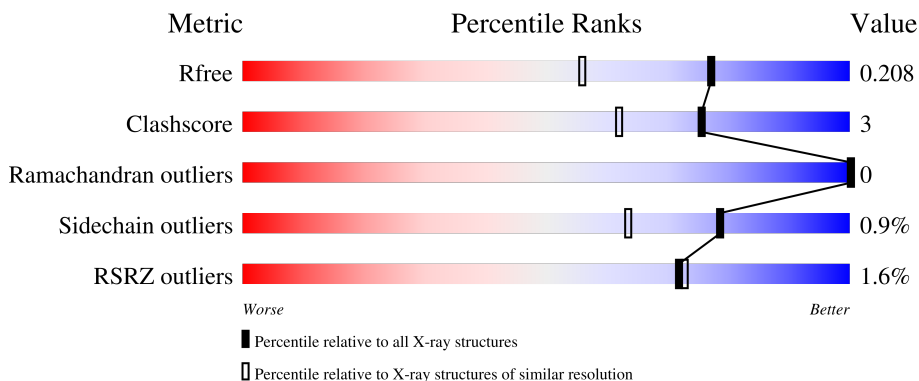
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.64 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.





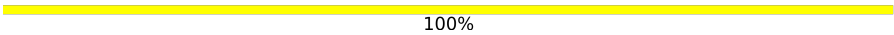
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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	214	 2% 93% 7%
1	B	214	 % 92% 7% .
1	C	214	 % 91% 9%
1	D	214	 2% 96% .
2	E	4	 100%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	4	 25% 75%
2	G	4	 25% 75%
2	H	4	 100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	BGC	E	4	X	-	-	-
2	BGC	F	4	X	-	-	-
2	BGC	G	4	X	-	-	-
2	BGC	H	4	X	-	-	-

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 8629 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 Beta-glucanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	214	1703	1093	276	329	5	0	3	0
1	B	214	1715	1099	278	333	5	0	6	0
1	C	214	1711	1098	278	330	5	0	5	0
1	D	214	1717	1101	278	333	5	0	7	0

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLN	-	SEE REMARK 999	UNP P23904
A	2	THR	-	SEE REMARK 999	UNP P23904
A	3	GLY	-	SEE REMARK 999	UNP P23904
A	4	GLY	-	SEE REMARK 999	UNP P23904
A	5	SER	-	SEE REMARK 999	UNP P23904
A	6	PHE	-	SEE REMARK 999	UNP P23904
A	7	PHE	-	SEE REMARK 999	UNP P23904
A	8	GLU	-	SEE REMARK 999	UNP P23904
A	9	PRO	-	SEE REMARK 999	UNP P23904
A	10	PHE	-	SEE REMARK 999	UNP P23904
A	11	ASN	-	SEE REMARK 999	UNP P23904
A	12	SER	-	SEE REMARK 999	UNP P23904
A	13	TYR	-	SEE REMARK 999	UNP P23904
A	14	ASN	-	SEE REMARK 999	UNP P23904
A	15	SER	-	SEE REMARK 999	UNP P23904
A	16	GLY	-	SEE REMARK 999	UNP P23904
A	105	GLN	GLU	engineered mutation	UNP P23904
A	109	GLN	GLU	engineered mutation	UNP P23904
B	301	GLN	-	SEE REMARK 999	UNP P23904
B	302	THR	-	SEE REMARK 999	UNP P23904
B	303	GLY	-	SEE REMARK 999	UNP P23904

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	304	GLY	-	SEE REMARK 999	UNP P23904
B	305	SER	-	SEE REMARK 999	UNP P23904
B	306	PHE	-	SEE REMARK 999	UNP P23904
B	307	PHE	-	SEE REMARK 999	UNP P23904
B	308	GLU	-	SEE REMARK 999	UNP P23904
B	309	PRO	-	SEE REMARK 999	UNP P23904
B	310	PHE	-	SEE REMARK 999	UNP P23904
B	311	ASN	-	SEE REMARK 999	UNP P23904
B	312	SER	-	SEE REMARK 999	UNP P23904
B	313	TYR	-	SEE REMARK 999	UNP P23904
B	314	ASN	-	SEE REMARK 999	UNP P23904
B	315	SER	-	SEE REMARK 999	UNP P23904
B	316	GLY	-	SEE REMARK 999	UNP P23904
B	405	GLN	GLU	engineered mutation	UNP P23904
B	409	GLN	GLU	engineered mutation	UNP P23904
C	601	GLN	-	SEE REMARK 999	UNP P23904
C	602	THR	-	SEE REMARK 999	UNP P23904
C	603	GLY	-	SEE REMARK 999	UNP P23904
C	604	GLY	-	SEE REMARK 999	UNP P23904
C	605	SER	-	SEE REMARK 999	UNP P23904
C	606	PHE	-	SEE REMARK 999	UNP P23904
C	607	PHE	-	SEE REMARK 999	UNP P23904
C	608	GLU	-	SEE REMARK 999	UNP P23904
C	609	PRO	-	SEE REMARK 999	UNP P23904
C	610	PHE	-	SEE REMARK 999	UNP P23904
C	611	ASN	-	SEE REMARK 999	UNP P23904
C	612	SER	-	SEE REMARK 999	UNP P23904
C	613	TYR	-	SEE REMARK 999	UNP P23904
C	614	ASN	-	SEE REMARK 999	UNP P23904
C	615	SER	-	SEE REMARK 999	UNP P23904
C	616	GLY	-	SEE REMARK 999	UNP P23904
C	705	GLN	GLU	engineered mutation	UNP P23904
C	709	GLN	GLU	engineered mutation	UNP P23904
D	901	GLN	-	SEE REMARK 999	UNP P23904
D	902	THR	-	SEE REMARK 999	UNP P23904
D	903	GLY	-	SEE REMARK 999	UNP P23904
D	904	GLY	-	SEE REMARK 999	UNP P23904
D	905	SER	-	SEE REMARK 999	UNP P23904
D	906	PHE	-	SEE REMARK 999	UNP P23904
D	907	PHE	-	SEE REMARK 999	UNP P23904
D	908	GLU	-	SEE REMARK 999	UNP P23904
D	909	PRO	-	SEE REMARK 999	UNP P23904

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	910	PHE	-	SEE REMARK 999	UNP P23904
D	911	ASN	-	SEE REMARK 999	UNP P23904
D	912	SER	-	SEE REMARK 999	UNP P23904
D	913	TYR	-	SEE REMARK 999	UNP P23904
D	914	ASN	-	SEE REMARK 999	UNP P23904
D	915	SER	-	SEE REMARK 999	UNP P23904
D	916	GLY	-	SEE REMARK 999	UNP P23904
D	1005	GLN	GLU	engineered mutation	UNP P23904
D	1009	GLN	GLU	engineered mutation	UNP P23904

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
2	E	4	Total C O 45 24 21	0	0	0
2	F	4	Total C O 45 24 21	0	0	0
2	G	4	Total C O 45 24 21	0	0	0
2	H	4	Total C O 45 24 21	0	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Ca 1 1	0	0
3	B	1	Total Ca 1 1	0	0
3	C	1	Total Ca 1 1	0	0
3	D	1	Total Ca 1 1	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total 2	Zn 2	0	0
4	B	1	Total 1	Zn 1	0	0

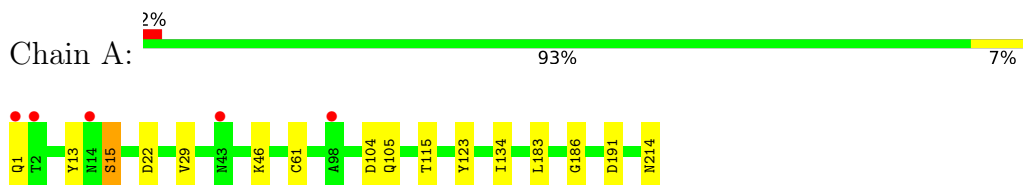
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	356	Total 356	O 356	0	0
5	B	426	Total 426	O 426	0	0
5	C	416	Total 416	O 416	0	0
5	D	398	Total 398	O 398	0	0

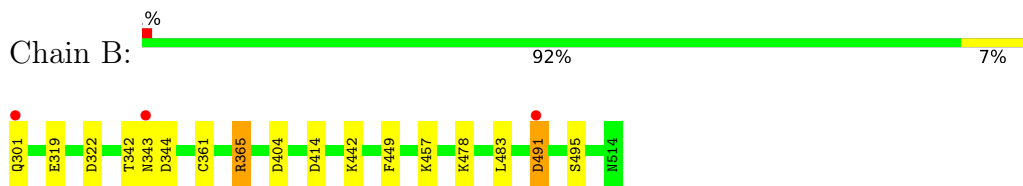
### 3 Residue-property plots [i](#)

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.

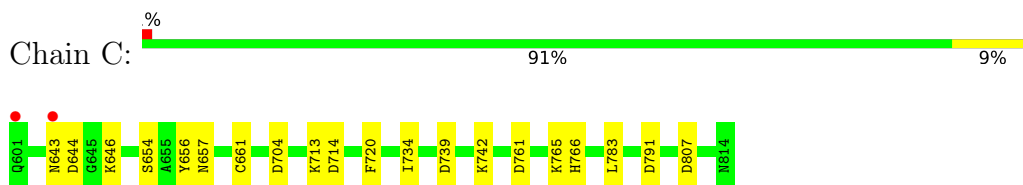
- Molecule 1: Beta-glucanase



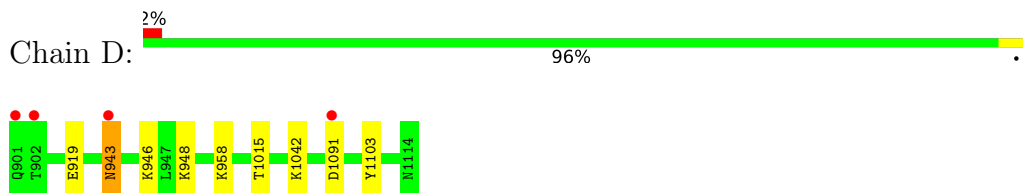
- Molecule 1: Beta-glucanase



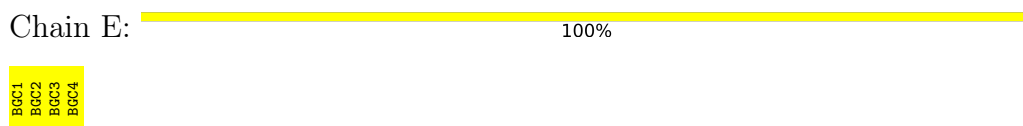
- Molecule 1: Beta-glucanase



- Molecule 1: Beta-glucanase



- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-beta-D-glucopyranose



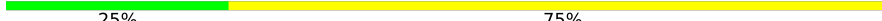


- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-beta-D-glucopyranose

Chain F:  25% 75%

BGC1  
BGC2  
BGC3  
BGC4

- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-beta-D-glucopyranose

Chain G:  25% 75%

BGC1  
BGC2  
BGC3  
BGC4

- Molecule 2: beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-4)-beta-D-glucopyranose-(1-3)-beta-D-glucopyranose

Chain H:  100%

BGC1  
BGC2  
BGC3  
BGC4

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.77Å 88.76Å 154.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.46 – 1.64 34.46 – 1.64	Depositor EDS
% Data completeness (in resolution range)	92.5 (34.46-1.64) 92.5 (34.46-1.64)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 1.64Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.161 , 0.197 0.174 , 0.208	Depositor DCC
$R_{free}$ test set	8310 reflections (7.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.5	Xtrriage
Anisotropy	0.117	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 40.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8629	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	21.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.29% 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: ZN, CA, BGC

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	0.60	0/1774	0.81	3/2411 (0.1%)
1	B	0.61	0/1801	0.87	5/2447 (0.2%)
1	C	0.62	0/1792	0.82	6/2436 (0.2%)
1	D	0.62	0/1807	0.82	1/2455 (0.0%)
All	All	0.61	0/7174	0.83	15/9749 (0.2%)

There are no bond length outliers.

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	491[A]	ASP	CB-CG-OD2	9.69	127.02	118.30
1	B	491[B]	ASP	CB-CG-OD2	9.69	127.02	118.30
1	A	104	ASP	CB-CG-OD2	7.24	124.82	118.30
1	B	322	ASP	CB-CG-OD2	6.50	124.15	118.30
1	C	791	ASP	CB-CG-OD2	5.85	123.56	118.30
1	C	739	ASP	CB-CG-OD2	5.75	123.48	118.30
1	C	807	ASP	CB-CG-OD1	5.68	123.41	118.30
1	C	704	ASP	CB-CG-OD2	5.63	123.36	118.30
1	C	714	ASP	CB-CG-OD2	5.52	123.27	118.30
1	A	22	ASP	CB-CG-OD2	5.38	123.14	118.30
1	B	404	ASP	CB-CG-OD2	5.32	123.08	118.30
1	D	1091	ASP	CB-CG-OD2	5.29	123.06	118.30
1	B	414	ASP	CB-CG-OD2	5.27	123.04	118.30
1	A	191	ASP	CB-CG-OD2	5.13	122.92	118.30
1	C	761	ASP	CB-CG-OD2	5.07	122.86	118.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1703	0	1572	11	1
1	B	1715	0	1579	8	1
1	C	1711	0	1579	14	0
1	D	1717	0	1585	9	0
2	E	45	0	39	0	0
2	F	45	0	39	0	0
2	G	45	0	39	0	0
2	H	45	0	39	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	2	0	0	0	0
4	B	1	0	0	0	0
5	A	356	0	0	3	0
5	B	426	0	0	4	0
5	C	416	0	0	8	0
5	D	398	0	0	8	0
All	All	8629	0	6471	42	1

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 3.

All (42) 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:13:TYR:OH	1:A:15[A]:SER:OG	1.97	0.80
1:C:720:PHE:CE1	1:C:734[A]:ILE:HD12	2.19	0.77
1:A:13:TYR:CZ	1:A:15[A]:SER:OG	2.37	0.74
1:C:646:LYS:HE3	5:C:3318:HOH:O	1.89	0.73
1:C:720:PHE:HE1	1:C:734[A]:ILE:HD12	1.60	0.64
1:B:478[B]:LYS:HE3	5:B:3029:HOH:O	1.97	0.63
1:C:643:ASN:ND2	5:C:3333:HOH:O	2.33	0.60
1:B:342:THR:OG1	1:B:344[A]:ASP:OD1	2.19	0.59
1:C:657[B]:ASN:OD1	5:C:2762:HOH:O	2.17	0.57

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:946:LYS:HE3	5:D:3455:HOH:O	2.04	0.56
1:A:1:GLN:HE21	1:A:1:GLN:HA	1.71	0.55
1:A:13:TYR:OH	5:A:3454:HOH:O	2.18	0.55
1:A:115:THR:HA	5:A:3325:HOH:O	2.07	0.55
1:C:644:ASP:HB2	5:C:3288:HOH:O	2.07	0.55
1:A:46:LYS:HE3	5:A:3539:HOH:O	2.07	0.55
1:D:958:LYS:NZ	5:D:3487:HOH:O	2.39	0.55
1:D:1015:THR:HA	5:D:2365:HOH:O	2.07	0.54
1:C:713:LYS:NZ	5:C:2450:HOH:O	2.05	0.52
1:C:734[A]:ILE:HD11	1:C:766:HIS:HB2	1.93	0.51
1:B:319:GLU:HG3	1:B:365:ARG:HG3	1.93	0.49
1:A:1:GLN:HA	1:A:1:GLN:NE2	2.28	0.48
1:C:661:CYS:HB2	1:C:783:LEU:O	2.14	0.48
1:C:713:LYS:HE3	5:C:2846:HOH:O	2.13	0.48
1:C:742:LYS:HE3	5:C:3261:HOH:O	2.14	0.47
1:C:713:LYS:CE	5:C:2450:HOH:O	2.56	0.47
1:B:301:GLN:NE2	5:B:3001:HOH:O	2.47	0.47
1:D:919[B]:GLU:CD	5:D:2661:HOH:O	2.54	0.45
1:A:61:CYS:HB2	1:A:183:LEU:O	2.18	0.43
1:D:1042[B]:LYS:HD3	5:D:2892:HOH:O	2.18	0.43
1:C:734[A]:ILE:HD13	1:C:765:LYS:HB3	2.00	0.43
1:D:943[A]:ASN:ND2	5:D:3107:HOH:O	2.51	0.43
1:B:442[B]:LYS:NZ	5:B:3387:HOH:O	2.51	0.43
1:A:29:VAL:O	1:A:186:GLY:HA2	2.19	0.42
1:B:361:CYS:HB2	1:B:483:LEU:O	2.19	0.42
1:B:449:PHE:HA	1:B:457:LYS:O	2.20	0.42
1:D:948:LYS:HE2	1:D:1103:TYR:CG	2.55	0.42
1:D:946:LYS:NZ	5:D:2763:HOH:O	2.38	0.41
1:A:134:ILE:HD13	1:A:134:ILE:HA	1.92	0.41
1:C:656:TYR:CE2	1:C:657[B]:ASN:ND2	2.88	0.41
1:B:495[A]:SER:HB2	5:B:2395:HOH:O	2.20	0.41
1:A:105:GLN:HG2	1:A:123:TYR:HB2	2.03	0.41
1:D:1042[B]:LYS:HE3	5:D:2312:HOH:O	2.20	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:214:ASN:O	1:B:301:GLN:N[2_565]	2.18	0.02

## 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	215/214 (100%)	209 (97%)	6 (3%)	0	100	100
1	B	218/214 (102%)	214 (98%)	4 (2%)	0	100	100
1	C	217/214 (101%)	213 (98%)	4 (2%)	0	100	100
1	D	219/214 (102%)	214 (98%)	5 (2%)	0	100	100
All	All	869/856 (102%)	850 (98%)	19 (2%)	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	179/176 (102%)	177 (99%)	2 (1%)	73	55
1	B	182/176 (103%)	177 (97%)	5 (3%)	44	18
1	C	181/176 (103%)	179 (99%)	2 (1%)	73	55
1	D	183/176 (104%)	181 (99%)	2 (1%)	73	55
All	All	725/704 (103%)	714 (98%)	11 (2%)	78	42

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15[A]	SER
1	A	15[B]	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	343[A]	ASN
1	B	343[B]	ASN
1	B	365	ARG
1	B	491[A]	ASP
1	B	491[B]	ASP
1	C	654[A]	SER
1	C	654[B]	SER
1	D	943[A]	ASN
1	D	943[B]	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1	GLN
1	A	152	GLN
1	A	214	ASN
1	B	301	GLN
1	C	668	ASN

### 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](#)

16 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	E	1	2	12,12,12	0.51	0	17,17,17	1.19	1 (5%)
2	BGC	E	2	2	11,11,12	1.11	1 (9%)	15,15,17	2.17	4 (26%)
2	BGC	E	3	2	11,11,12	0.92	0	15,15,17	2.29	4 (26%)
2	BGC	E	4	2	11,11,12	0.83	0	15,15,17	2.73	4 (26%)
2	BGC	F	1	2	12,12,12	0.45	0	17,17,17	0.71	0
2	BGC	F	2	2	11,11,12	1.20	1 (9%)	15,15,17	1.88	2 (13%)
2	BGC	F	3	2	11,11,12	0.96	1 (9%)	15,15,17	2.35	4 (26%)
2	BGC	F	4	2	11,11,12	0.59	0	15,15,17	2.83	5 (33%)
2	BGC	G	1	2	12,12,12	0.54	0	17,17,17	0.72	0
2	BGC	G	2	2	11,11,12	0.98	1 (9%)	15,15,17	2.12	4 (26%)
2	BGC	G	3	2	11,11,12	0.75	0	15,15,17	2.63	4 (26%)
2	BGC	G	4	2	11,11,12	0.75	0	15,15,17	2.67	3 (20%)
2	BGC	H	1	2	12,12,12	0.65	0	17,17,17	1.00	1 (5%)
2	BGC	H	2	2	11,11,12	0.80	1 (9%)	15,15,17	1.96	4 (26%)
2	BGC	H	3	2	11,11,12	1.00	1 (9%)	15,15,17	2.24	3 (20%)
2	BGC	H	4	2	11,11,12	0.64	0	15,15,17	2.56	4 (26%)

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	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	BGC	E	4	2	1/1/4/5	1/2/19/22	0/1/1/1
2	BGC	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	-	1/2/19/22	0/1/1/1
2	BGC	F	4	2	1/1/4/5	2/2/19/22	0/1/1/1
2	BGC	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	BGC	G	4	2	1/1/4/5	2/2/19/22	0/1/1/1
2	BGC	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
2	BGC	H	4	2	1/1/4/5	2/2/19/22	0/1/1/1



All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	2	BGC	O5-C1	-3.41	1.38	1.43
2	E	2	BGC	O5-C1	-3.14	1.38	1.43
2	F	3	BGC	O5-C1	-2.78	1.39	1.43
2	G	2	BGC	O5-C1	-2.53	1.39	1.43
2	H	2	BGC	O5-C1	-2.22	1.40	1.43
2	H	3	BGC	O5-C1	-2.10	1.40	1.43

All (47) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	4	BGC	O5-C1-C2	8.43	123.78	110.77
2	F	4	BGC	O5-C1-C2	8.06	123.21	110.77
2	G	3	BGC	O5-C1-C2	8.04	123.19	110.77
2	G	4	BGC	O5-C1-C2	7.70	122.66	110.77
2	F	3	BGC	C1-O5-C5	-7.44	102.11	112.19
2	H	3	BGC	O5-C1-C2	7.15	121.80	110.77
2	E	3	BGC	O5-C1-C2	7.00	121.58	110.77
2	H	4	BGC	O5-C1-C2	6.50	120.80	110.77
2	G	2	BGC	C1-O5-C5	-6.25	103.73	112.19
2	F	2	BGC	C1-O5-C5	-6.00	104.06	112.19
2	E	2	BGC	C1-O5-C5	-5.96	104.11	112.19
2	H	2	BGC	C1-O5-C5	-5.51	104.72	112.19
2	G	4	BGC	C1-O5-C5	5.47	119.60	112.19
2	H	4	BGC	C2-C3-C4	-4.70	102.75	110.89
2	E	4	BGC	C2-C3-C4	-4.00	103.97	110.89
2	F	4	BGC	C1-O5-C5	3.99	117.60	112.19
2	H	4	BGC	C1-O5-C5	3.85	117.41	112.19
2	E	4	BGC	C1-O5-C5	3.80	117.35	112.19
2	F	4	BGC	C2-C3-C4	-3.73	104.44	110.89
2	F	4	BGC	C1-C2-C3	3.66	114.16	109.67
2	E	2	BGC	O5-C1-C2	3.52	116.21	110.77
2	E	1	BGC	C4-C3-C2	-3.50	104.72	110.82
2	G	2	BGC	C1-C2-C3	-3.27	105.65	109.67
2	E	2	BGC	C1-C2-C3	-3.20	105.73	109.67
2	G	3	BGC	O2-C2-C1	3.10	115.49	109.15
2	E	3	BGC	O4-C4-C3	-3.00	103.42	110.35
2	G	4	BGC	C2-C3-C4	-2.89	105.89	110.89
2	H	2	BGC	O5-C1-C2	2.86	115.19	110.77
2	H	2	BGC	C1-C2-C3	-2.76	106.27	109.67
2	E	3	BGC	C1-C2-C3	2.67	112.94	109.67
2	F	3	BGC	O5-C5-C6	2.63	111.32	107.20
2	G	3	BGC	C1-O5-C5	2.57	115.67	112.19

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	2	BGC	O5-C5-C6	2.50	111.12	107.20
2	G	3	BGC	O5-C5-C6	-2.41	103.43	107.20
2	F	3	BGC	O4-C4-C3	-2.40	104.79	110.35
2	F	2	BGC	O4-C4-C3	-2.30	105.04	110.35
2	H	3	BGC	O4-C4-C3	-2.27	105.10	110.35
2	F	4	BGC	O5-C5-C4	2.22	116.22	110.83
2	E	4	BGC	O5-C5-C6	2.20	110.65	107.20
2	E	3	BGC	O2-C2-C1	2.19	113.63	109.15
2	G	2	BGC	O5-C5-C6	2.17	110.61	107.20
2	F	3	BGC	O2-C2-C1	2.17	113.59	109.15
2	G	2	BGC	O5-C1-C2	2.16	114.10	110.77
2	H	1	BGC	C4-C3-C2	-2.11	107.13	110.82
2	H	4	BGC	C3-C4-C5	-2.11	106.48	110.24
2	H	3	BGC	C1-C2-C3	2.06	112.20	109.67
2	H	2	BGC	O5-C5-C6	2.05	110.42	107.20

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	E	4	BGC	C1
2	F	4	BGC	C1
2	G	4	BGC	C1
2	H	4	BGC	C1

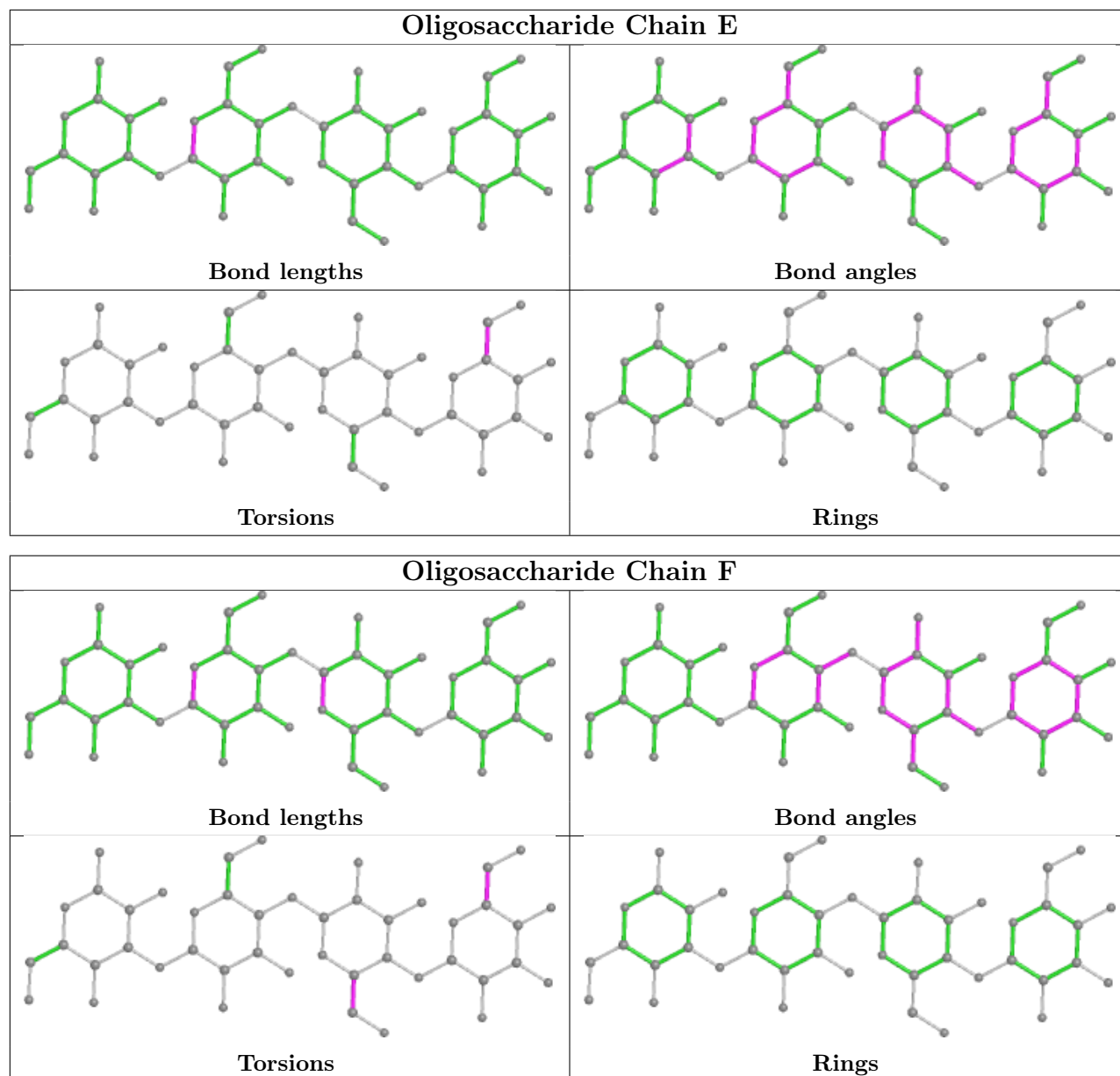
All (8) torsion outliers are listed below:

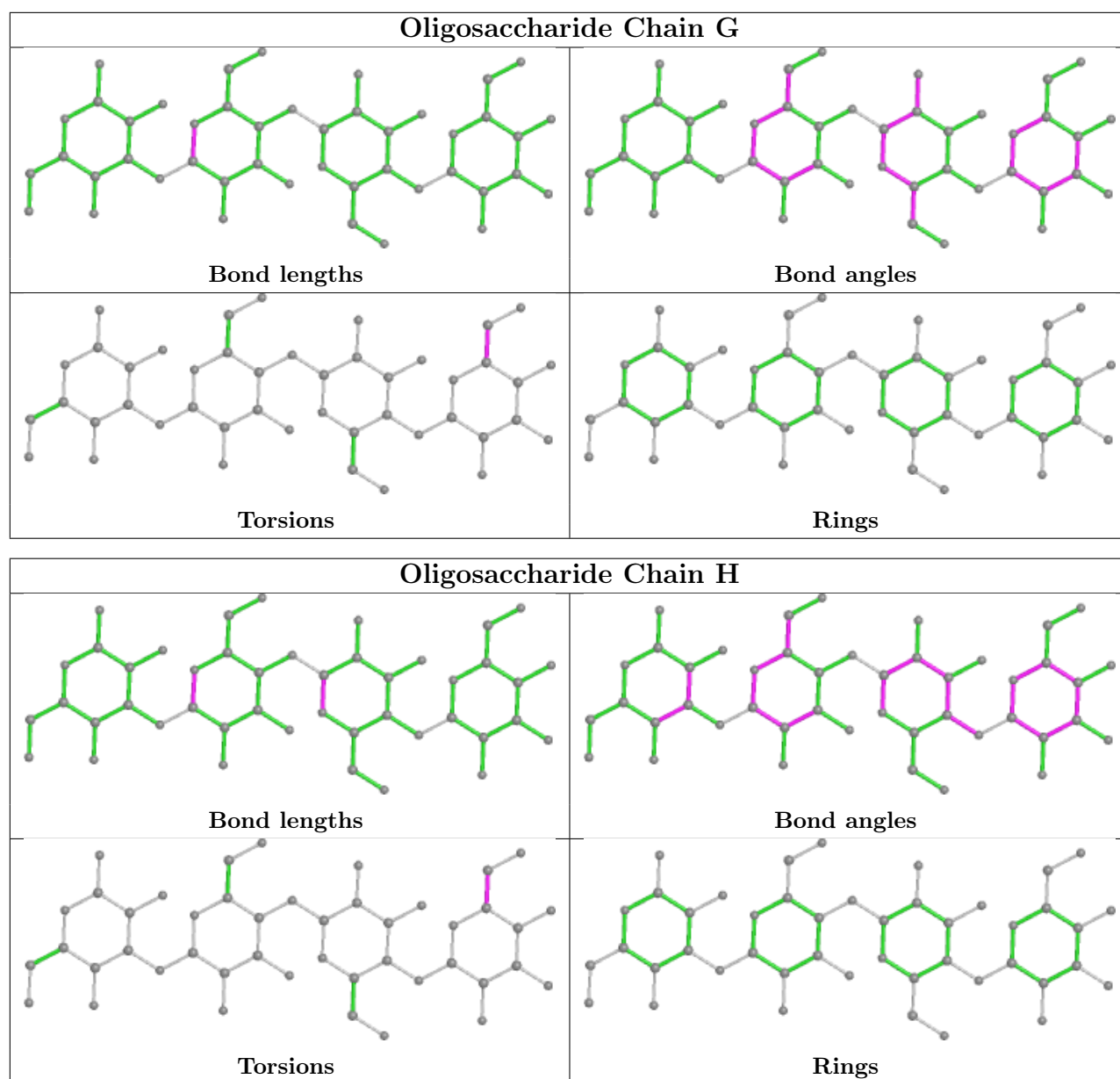
Mol	Chain	Res	Type	Atoms
2	H	4	BGC	O5-C5-C6-O6
2	H	4	BGC	C4-C5-C6-O6
2	F	4	BGC	O5-C5-C6-O6
2	G	4	BGC	O5-C5-C6-O6
2	F	4	BGC	C4-C5-C6-O6
2	F	3	BGC	C4-C5-C6-O6
2	E	4	BGC	O5-C5-C6-O6
2	G	4	BGC	C4-C5-C6-O6

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 7 ligands modelled in this entry, 7 are monoatomic - leaving 0 for Mogul analysis.

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 [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	214/214 (100%)	0.10	5 (2%) 60 60	12, 19, 35, 55	0
1	B	214/214 (100%)	-0.17	3 (1%) 75 76	12, 17, 24, 46	0
1	C	214/214 (100%)	-0.13	2 (0%) 84 85	12, 18, 25, 45	0
1	D	214/214 (100%)	-0.20	4 (1%) 66 67	11, 16, 25, 47	0
All	All	856/856 (100%)	-0.10	14 (1%) 72 73	11, 18, 28, 55	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	GLN	6.6
1	D	901	GLN	5.4
1	C	601	GLN	4.7
1	A	43	ASN	4.0
1	C	643	ASN	3.6
1	B	301	GLN	3.6
1	B	343[A]	ASN	3.5
1	B	491[A]	ASP	3.3
1	A	14	ASN	3.0
1	A	2	THR	2.9
1	D	902	THR	2.9
1	D	943[A]	ASN	2.4
1	A	98	ALA	2.4
1	D	1091	ASP	2.3

### 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

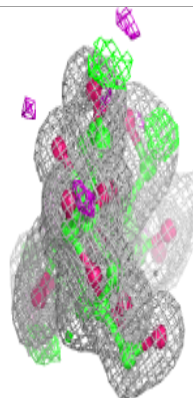
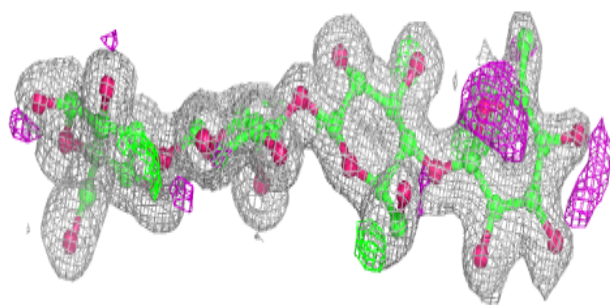
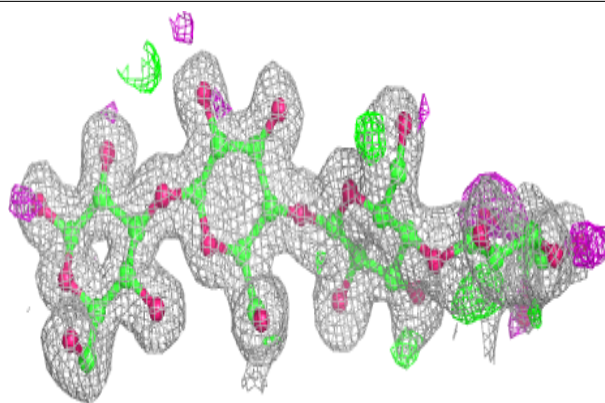
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	BGC	E	4	11/12	0.64	0.32	35,38,40,41	0
2	BGC	G	4	11/12	0.72	0.28	31,34,36,36	0
2	BGC	F	4	11/12	0.79	0.18	25,28,30,32	0
2	BGC	H	4	11/12	0.81	0.29	33,36,38,39	0
2	BGC	E	3	11/12	0.86	0.14	23,26,30,32	0
2	BGC	E	2	11/12	0.92	0.09	17,19,22,22	0
2	BGC	G	3	11/12	0.94	0.09	20,23,26,27	0
2	BGC	E	1	12/12	0.94	0.08	15,18,19,19	0
2	BGC	H	2	11/12	0.94	0.08	15,17,20,21	0
2	BGC	G	1	12/12	0.94	0.10	16,17,18,21	0
2	BGC	H	1	12/12	0.95	0.09	15,16,18,19	0
2	BGC	F	3	11/12	0.95	0.08	15,19,22,23	0
2	BGC	H	3	11/12	0.95	0.09	17,21,24,28	0
2	BGC	G	2	11/12	0.95	0.07	17,19,20,21	0
2	BGC	F	1	12/12	0.96	0.07	13,14,15,15	0
2	BGC	F	2	11/12	0.96	0.07	14,15,16,18	0

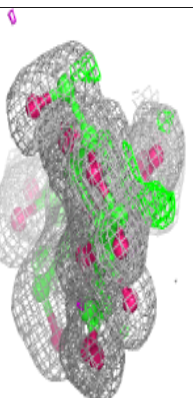
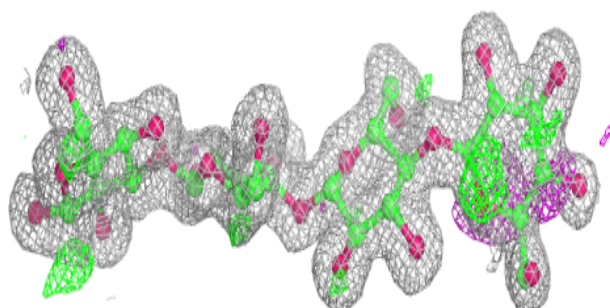
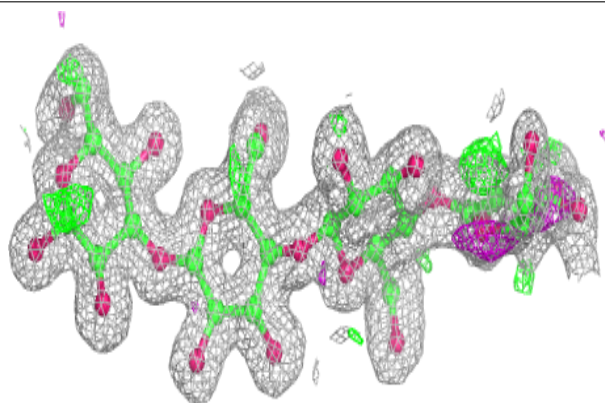
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**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:**

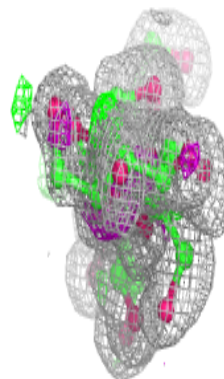
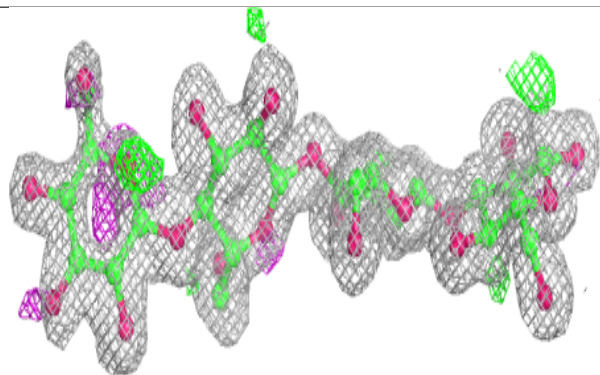
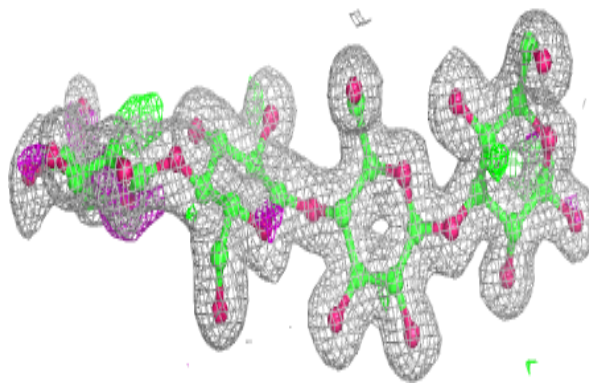
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



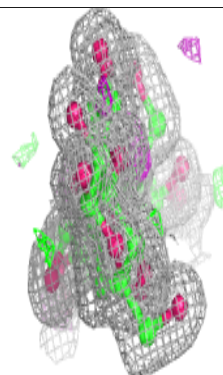
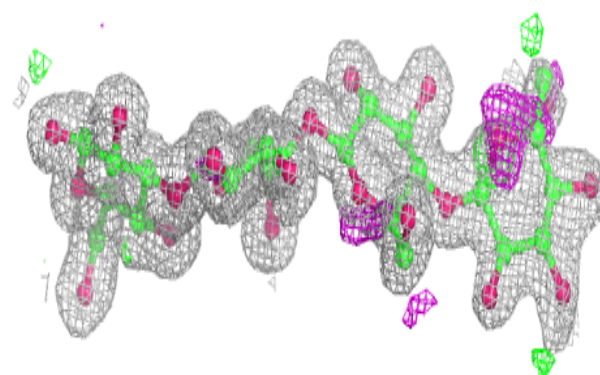
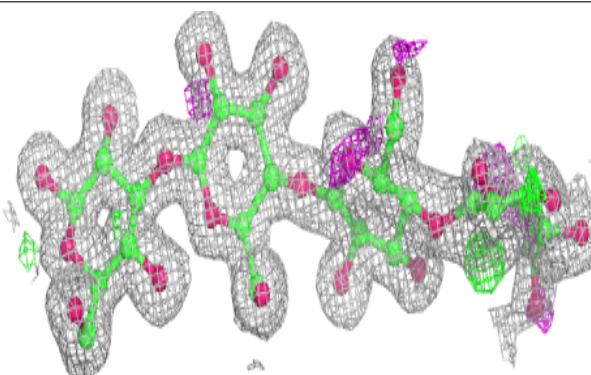


**Electron density around Chain G:**

$2mF_o-DF_c$  (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( $\text{\AA}^2$ )	Q<0.9
3	CA	A	5004	1/1	0.90	0.09	27,27,27,27	0
3	CA	B	5002	1/1	0.98	0.06	20,20,20,20	0
3	CA	C	5003	1/1	0.98	0.04	20,20,20,20	0
3	CA	D	5001	1/1	0.99	0.06	13,13,13,13	0
4	ZN	A	5011	1/1	0.99	0.07	23,23,23,23	0
4	ZN	A	5005	1/1	1.00	0.06	16,16,16,16	0
4	ZN	B	5006	1/1	1.00	0.04	13,13,13,13	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.