



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 9, 2020 – 05:27 PM BST

PDB ID : 3DWQ  
Title : Crystal structure of the A-subunit of the AB5 toxin from E. coli with Neu5Gc-2,3Gal-1,3GlcNAc  
Authors : Byres, E.; Paton, A.W.; Paton, J.C.; Loffing, J.C.; Smith, D.F.; Wilce, M.C.J.; Talbot, U.M.; Chong, D.C.; Yu, H.; Huang, S.; Chen, X.; Varki, N.M.; Varki, A.; Rossjohn, J.; Beddoe, T.  
Deposited on : 2008-07-22  
Resolution : 2.10 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

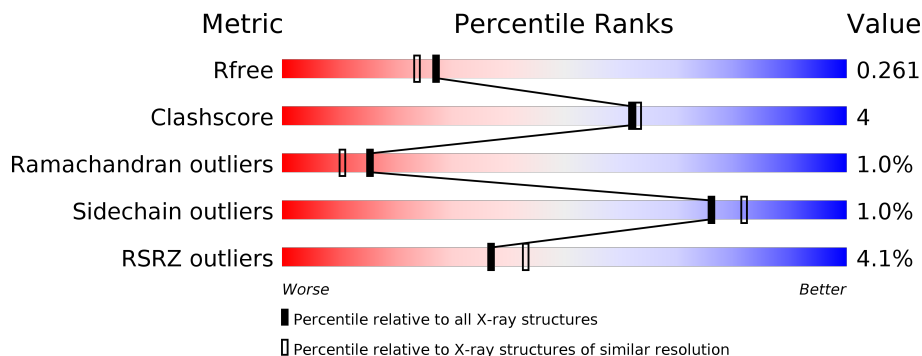
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	126	 4% 82% 10% 8%
1	B	126	 2% 76% 16% 7%
1	C	126	 3% 83% 12% 5%
1	D	126	 4% 83% 10% 7%
1	E	126	 6% 89% 8%
2	F	3	 100%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain	
2	G	3		
2	H	3		
2	I	3		

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	NGA	F	1	X	-	-	X
2	GAL	F	2	X	-	-	-
2	NGC	F	3	X	-	-	-
2	NGA	G	1	X	-	-	-
2	GAL	G	2	X	-	-	-
2	NGC	G	3	X	-	-	-
2	NGA	H	1	X	-	-	-
2	GAL	H	2	X	-	-	-
2	NGC	H	3	X	-	-	-
2	NGA	I	1	X	-	-	-
2	GAL	I	2	X	-	-	-
2	NGC	I	3	X	-	-	-
3	POL	C	130	-	-	-	X
4	AZI	C	131	-	-	-	X
4	AZI	D	131	-	-	-	X

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 5380 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 Subtilase cytotoxin, subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	116	Total 897	C 573	N 145	O 173	S 6	0	0	0
1	B	117	Total 905	C 577	N 147	O 175	S 6	0	0	0
1	C	120	Total 928	C 591	N 150	O 181	S 6	0	0	0
1	D	117	Total 905	C 577	N 147	O 175	S 6	0	0	0
1	E	116	Total 897	C 573	N 145	O 173	S 6	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

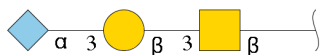
Chain	Residue	Modelled	Actual	Comment	Reference
A	119	LEU	-	expression tag	UNP Q3ZTX8
A	120	GLU	-	expression tag	UNP Q3ZTX8
A	121	HIS	-	expression tag	UNP Q3ZTX8
A	122	HIS	-	expression tag	UNP Q3ZTX8
A	123	HIS	-	expression tag	UNP Q3ZTX8
A	124	HIS	-	expression tag	UNP Q3ZTX8
A	125	HIS	-	expression tag	UNP Q3ZTX8
A	126	HIS	-	expression tag	UNP Q3ZTX8
B	119	LEU	-	expression tag	UNP Q3ZTX8
B	120	GLU	-	expression tag	UNP Q3ZTX8
B	121	HIS	-	expression tag	UNP Q3ZTX8
B	122	HIS	-	expression tag	UNP Q3ZTX8
B	123	HIS	-	expression tag	UNP Q3ZTX8
B	124	HIS	-	expression tag	UNP Q3ZTX8
B	125	HIS	-	expression tag	UNP Q3ZTX8
B	126	HIS	-	expression tag	UNP Q3ZTX8
C	119	LEU	-	expression tag	UNP Q3ZTX8
C	120	GLU	-	expression tag	UNP Q3ZTX8
C	121	HIS	-	expression tag	UNP Q3ZTX8

*Continued on next page...*

Continued from previous page...

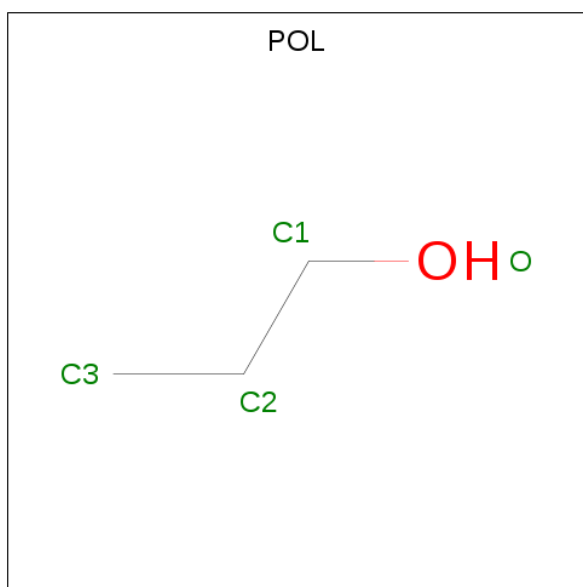
Chain	Residue	Modelled	Actual	Comment	Reference
C	122	HIS	-	expression tag	UNP Q3ZTX8
C	123	HIS	-	expression tag	UNP Q3ZTX8
C	124	HIS	-	expression tag	UNP Q3ZTX8
C	125	HIS	-	expression tag	UNP Q3ZTX8
C	126	HIS	-	expression tag	UNP Q3ZTX8
D	119	LEU	-	expression tag	UNP Q3ZTX8
D	120	GLU	-	expression tag	UNP Q3ZTX8
D	121	HIS	-	expression tag	UNP Q3ZTX8
D	122	HIS	-	expression tag	UNP Q3ZTX8
D	123	HIS	-	expression tag	UNP Q3ZTX8
D	124	HIS	-	expression tag	UNP Q3ZTX8
D	125	HIS	-	expression tag	UNP Q3ZTX8
D	126	HIS	-	expression tag	UNP Q3ZTX8
E	119	LEU	-	expression tag	UNP Q3ZTX8
E	120	GLU	-	expression tag	UNP Q3ZTX8
E	121	HIS	-	expression tag	UNP Q3ZTX8
E	122	HIS	-	expression tag	UNP Q3ZTX8
E	123	HIS	-	expression tag	UNP Q3ZTX8
E	124	HIS	-	expression tag	UNP Q3ZTX8
E	125	HIS	-	expression tag	UNP Q3ZTX8
E	126	HIS	-	expression tag	UNP Q3ZTX8

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



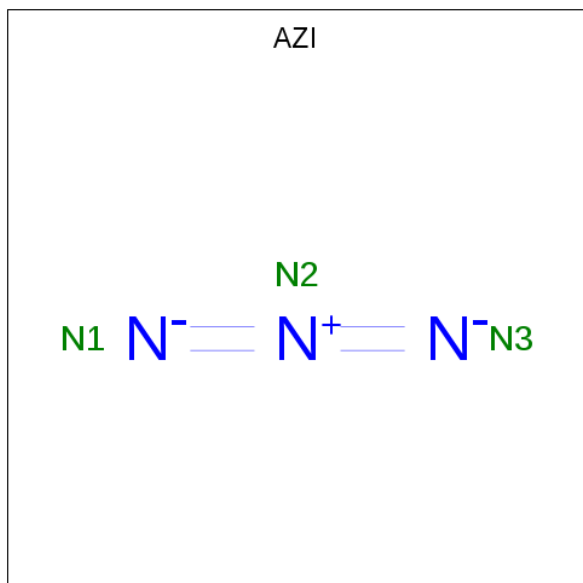
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	F	3	46	25	2	19	0	0	0
2	G	3	46	25	2	19	0	0	0
2	H	3	46	25	2	19	0	0	0
2	I	3	46	25	2	19	0	0	0

- Molecule 3 is N-PROPANOL (three-letter code: POL) (formula: C<sub>3</sub>H<sub>8</sub>O).



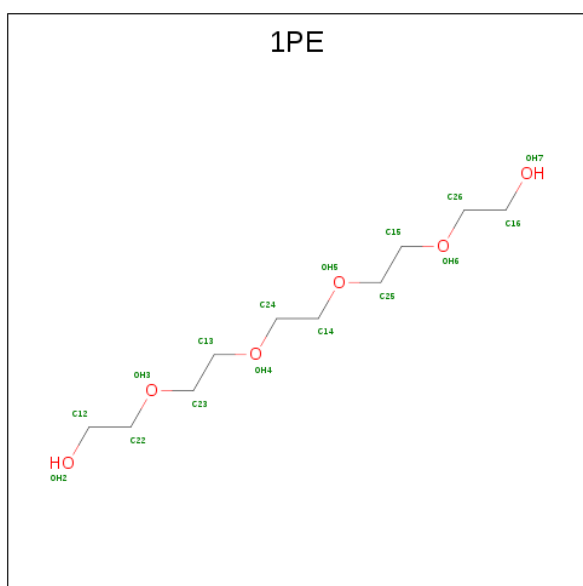
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	3	1		
3	C	1	Total	C	O	0	0
			4	3	1		
3	D	1	Total	C	O	0	0
			4	3	1		
3	E	1	Total	C	O	0	0
			4	3	1		

- Molecule 4 is AZIDE ION (three-letter code: AZI) (formula: N<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total N 3 3	0	0
4	C	1	Total N 3 3	0	0
4	D	1	Total N 3 3	0	0
4	E	1	Total N 3 3	0	0

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 16 10 6	0	0
5	B	1	Total C O 16 10 6	0	0
5	D	1	Total C O 16 10 6	0	0
5	D	1	Total C O 16 10 6	0	0
5	E	1	Total C O 16 10 6	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	112	Total O 112 112	0	0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	B	103	Total 103	O 103	0	0
6	C	108	Total 108	O 108	0	0
6	D	107	Total 107	O 107	0	0
6	E	126	Total 126	O 126	0	0





- Molecule 2: N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose

Chain F:  100%

NGA1  
GAL2  
NGC3

- Molecule 2: N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose

Chain G:  33% 67%

NGA1  
GAL2  
NGC3

- Molecule 2: N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose

Chain H:  33% 67%

NGA1  
GAL2  
NGC3

- Molecule 2: N-glycolyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-galactopyranose

Chain I:  33% 67%

NGA1  
GAL2  
NGC3

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	97.25Å 97.25Å 163.52Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.67 – 2.10 29.66 – 2.10	Depositor EDS
% Data completeness (in resolution range)	83.8 (29.67-2.10) 83.8 (29.66-2.10)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.74 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.211 , 0.262 0.211 , 0.261	Depositor DCC
$R_{free}$ test set	2177 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.0	Xtrriage
Anisotropy	0.024	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 50.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.044 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	5380	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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: AZI, NGA, NGC, 1PE, POL, 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	0.38	0/923	0.46	0/1255
1	B	0.39	0/931	0.51	0/1266
1	C	0.38	0/954	0.48	0/1297
1	D	0.37	0/931	0.48	0/1266
1	E	0.39	0/923	0.48	0/1255
All	All	0.38	0/4662	0.48	0/6339

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	897	0	852	8	0
1	B	905	0	858	15	0
1	C	928	0	882	8	0
1	D	905	0	858	12	0
1	E	897	0	852	2	0
2	F	46	0	29	0	0
2	G	46	0	29	1	0
2	H	46	0	29	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	I	46	0	29	1	0
3	A	4	0	5	0	0
3	C	4	0	5	0	0
3	D	4	0	6	0	0
3	E	4	0	5	0	0
4	A	3	0	0	0	0
4	C	3	0	0	0	0
4	D	3	0	0	0	0
4	E	3	0	0	0	0
5	A	16	0	22	0	0
5	B	16	0	22	1	0
5	D	32	0	44	1	0
5	E	16	0	22	0	0
6	A	112	0	0	1	0
6	B	103	0	0	2	0
6	C	108	0	0	1	0
6	D	107	0	0	0	0
6	E	126	0	0	1	0
All	All	5380	0	4549	41	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 4.

All (41) 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:31:CYS:HB2	6:C:199:HOH:O	1.60	1.02
1:D:10:MET:HE3	1:D:114:ARG:H	1.29	0.97
1:A:47:MET:HB3	1:A:50:SER:HB3	1.57	0.86
1:B:62:ASN:HB3	6:E:189:HOH:O	1.79	0.80
1:D:10:MET:HE3	1:D:114:ARG:N	2.04	0.70
1:D:14:VAL:HG12	1:D:36:GLN:HG3	1.78	0.66
1:D:10:MET:HE2	1:D:75:ARG:HD3	1.80	0.64
1:A:29:TYR:HB3	1:A:47:MET:HG3	1.88	0.56
2:I:2:GAL:C4	2:I:3:NGC:O6	2.54	0.55
1:A:63:GLN:HE22	1:C:65:LEU:HD13	1.72	0.55
2:H:2:GAL:C4	2:H:3:NGC:O6	2.57	0.53
1:B:56:SER:OG	1:E:23:GLN:HB2	2.09	0.52
1:E:30:PHE:CE2	1:E:46:SER:HB3	2.44	0.52
1:B:31:CYS:HB2	6:B:191:HOH:O	2.10	0.51
1:B:92:LEU:HD13	6:B:191:HOH:O	2.11	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:70:THR:HG22	5:D:132:IPE:H252	1.94	0.49
2:G:2:GAL:C4	2:G:3:NGC:O6	2.60	0.49
1:C:18:GLN:HB2	1:C:33:GLU:HB3	1.95	0.48
1:C:30:PHE:CE2	1:C:46:SER:HB3	2.49	0.48
1:D:10:MET:CE	1:D:75:ARG:HD3	2.42	0.47
1:A:56:SER:OG	1:C:23:GLN:HB2	2.14	0.47
1:A:62:ASN:HB3	6:A:140:HOH:O	2.13	0.47
1:B:67:PHE:CG	1:B:74:VAL:HG11	2.50	0.47
1:B:18:GLN:HB2	1:B:33:GLU:HB3	1.97	0.47
1:D:30:PHE:CE2	1:D:46:SER:HB3	2.50	0.46
1:B:30:PHE:CE2	1:B:46:SER:HB3	2.51	0.46
1:B:23:GLN:HB2	1:D:56:SER:OG	2.15	0.46
1:B:70:THR:HG22	5:B:127:IPE:H152	1.98	0.46
1:B:20:HIS:HD2	1:D:101:SER:OG	1.99	0.45
1:B:75:ARG:HB3	1:B:101:SER:HB2	1.99	0.45
1:B:47:MET:HB3	1:B:50:SER:HB3	1.98	0.45
1:B:29:TYR:HB3	1:B:47:MET:HG3	1.99	0.45
1:D:75:ARG:HB3	1:D:101:SER:HB2	1.98	0.45
1:C:31:CYS:SG	1:C:45:CYS:SG	3.15	0.44
1:B:37:SER:HA	1:B:39:GLY:N	2.33	0.44
1:A:30:PHE:CE2	1:A:46:SER:HB3	2.53	0.43
1:A:101:SER:OG	1:C:20:HIS:CD2	2.71	0.43
1:D:29:TYR:HB3	1:D:47:MET:HG3	2.01	0.42
1:C:75:ARG:HB3	1:C:101:SER:HB2	2.00	0.42
1:A:103:CYS:HA	1:A:109:CYS:HA	2.02	0.41
1:B:20:HIS:CD2	1:D:101:SER:OG	2.74	0.41

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	114/126 (90%)	110 (96%)	2 (2%)	2 (2%)	8	4
1	B	115/126 (91%)	109 (95%)	5 (4%)	1 (1%)	17	12
1	C	118/126 (94%)	115 (98%)	1 (1%)	2 (2%)	9	4
1	D	115/126 (91%)	112 (97%)	3 (3%)	0	100	100
1	E	114/126 (90%)	112 (98%)	1 (1%)	1 (1%)	17	12
All	All	576/630 (91%)	558 (97%)	12 (2%)	6 (1%)	15	11

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	51	SER
1	A	51	SER
1	C	50	SER
1	A	50	SER
1	E	51	SER
1	B	51	SER

### 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	98/108 (91%)	97 (99%)	1 (1%)	76	82
1	B	99/108 (92%)	98 (99%)	1 (1%)	76	82
1	C	102/108 (94%)	100 (98%)	2 (2%)	55	60
1	D	99/108 (92%)	98 (99%)	1 (1%)	76	82
1	E	98/108 (91%)	98 (100%)	0	100	100
All	All	496/540 (92%)	491 (99%)	5 (1%)	76	82

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

Mol	Chain	Res	Type
1	A	40	SER
1	B	37	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	27	LYS
1	C	107	THR
1	D	40	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (13) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	63	GLN
1	A	72	GLN
1	B	20	HIS
1	B	26	ASN
1	B	63	GLN
1	B	72	GLN
1	C	20	HIS
1	C	36	GLN
1	C	63	GLN
1	C	72	GLN
1	C	117	ASN
1	D	49	ASN
1	E	63	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	NGA	F	1	3,2	14,14,15	1.09	1 (7%)	17,19,21	0.76	0
2	GAL	F	2	2	11,11,12	0.59	0	15,15,17	1.68	2 (13%)
2	NGC	F	3	2	18,21,22	0.69	0	22,29,32	1.93	3 (13%)
2	NGA	G	1	3,2	14,14,15	1.06	1 (7%)	17,19,21	1.26	3 (17%)
2	GAL	G	2	2	11,11,12	0.70	0	15,15,17	1.46	2 (13%)
2	NGC	G	3	2	18,21,22	0.89	1 (5%)	22,29,32	2.22	7 (31%)
2	NGA	H	1	3,2	14,14,15	1.12	1 (7%)	17,19,21	1.10	2 (11%)
2	GAL	H	2	2	11,11,12	0.76	0	15,15,17	1.59	2 (13%)
2	NGC	H	3	2	18,21,22	0.80	0	22,29,32	2.19	7 (31%)
2	NGA	I	1	3,2	14,14,15	1.10	1 (7%)	17,19,21	1.69	6 (35%)
2	GAL	I	2	2	11,11,12	0.60	0	15,15,17	1.77	4 (26%)
2	NGC	I	3	2	18,21,22	1.01	1 (5%)	22,29,32	2.87	8 (36%)

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	NGA	F	1	3,2	5/5/5/7	5/6/23/26	0/1/1/1
2	GAL	F	2	2	5/5/4/5	2/2/19/22	0/1/1/1
2	NGC	F	3	2	5/5/8/9	3/15/36/40	0/1/1/1
2	NGA	G	1	3,2	5/5/5/7	5/6/23/26	0/1/1/1
2	GAL	G	2	2	5/5/4/5	1/2/19/22	0/1/1/1
2	NGC	G	3	2	5/5/8/9	4/15/36/40	1/1/1/1
2	NGA	H	1	3,2	5/5/5/7	4/6/23/26	0/1/1/1
2	GAL	H	2	2	5/5/4/5	2/2/19/22	0/1/1/1
2	NGC	H	3	2	5/5/8/9	5/15/36/40	1/1/1/1
2	NGA	I	1	3,2	5/5/5/7	3/6/23/26	0/1/1/1
2	GAL	I	2	2	5/5/4/5	2/2/19/22	0/1/1/1
2	NGC	I	3	2	5/5/8/9	5/15/36/40	1/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	1	NGA	C7-N2	3.51	1.46	1.34
2	H	1	NGA	C7-N2	3.47	1.46	1.34

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1	NGA	C7-N2	3.44	1.46	1.34
2	G	1	NGA	C7-N2	3.39	1.46	1.34
2	I	3	NGC	C4-C5	2.76	1.55	1.53
2	G	3	NGC	C4-C5	2.54	1.55	1.53

All (46) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	3	NGC	C5-N5-C10	8.24	135.03	122.90
2	F	3	NGC	O6-C2-C3	6.00	120.40	109.87
2	G	3	NGC	C5-N5-C10	5.60	131.15	122.90
2	I	2	GAL	O3-C3-C4	5.41	122.85	110.35
2	F	2	GAL	C1-C2-C3	5.31	116.20	109.67
2	H	3	NGC	C5-N5-C10	5.20	130.55	122.90
2	G	3	NGC	O6-C2-C3	5.14	118.90	109.87
2	I	3	NGC	C3-C2-C1	4.93	122.69	111.93
2	H	2	GAL	O3-C3-C4	4.84	121.55	110.35
2	I	3	NGC	O6-C2-C3	4.70	118.12	109.87
2	H	3	NGC	O6-C2-C3	4.62	117.97	109.87
2	F	3	NGC	C5-N5-C10	4.54	129.58	122.90
2	H	3	NGC	C3-C2-C1	4.38	121.49	111.93
2	G	2	GAL	O3-C3-C4	4.34	120.39	110.35
2	G	3	NGC	C3-C2-C1	4.10	120.89	111.93
2	I	1	NGA	C2-N2-C7	3.76	128.26	122.90
2	I	3	NGC	C3-C4-C5	-3.66	107.03	111.46
2	I	3	NGC	C4-C3-C2	-3.43	103.66	109.81
2	I	3	NGC	C4-C5-N5	3.25	116.81	110.38
2	F	3	NGC	C3-C2-C1	3.22	118.95	111.93
2	H	3	NGC	C3-C4-C5	-3.15	107.66	111.46
2	I	3	NGC	C8-C7-C6	3.05	118.81	113.03
2	G	3	NGC	C4-C3-C2	-2.84	104.71	109.81
2	G	3	NGC	C6-C5-N5	2.80	115.57	110.91
2	H	3	NGC	C4-C3-C2	-2.74	104.90	109.81
2	H	3	NGC	C6-C5-N5	2.69	115.38	110.91
2	I	2	GAL	C1-O5-C5	2.62	115.74	112.19
2	I	1	NGA	C8-C7-N2	-2.55	111.78	116.10
2	G	1	NGA	C8-C7-N2	-2.51	111.86	116.10
2	G	2	GAL	C1-O5-C5	2.38	115.42	112.19
2	I	1	NGA	O7-C7-N2	2.37	126.31	121.95
2	F	2	GAL	O3-C3-C2	-2.36	105.47	109.99
2	H	1	NGA	O5-C5-C6	2.36	110.91	107.20
2	I	3	NGC	C11-C10-N5	-2.34	112.30	116.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	1	NGA	C4-C3-C2	2.23	114.28	111.02
2	H	2	GAL	O5-C5-C6	2.21	110.67	107.20
2	G	3	NGC	C8-C7-C6	2.19	117.18	113.03
2	I	2	GAL	C2-C3-C4	-2.17	107.15	110.89
2	I	1	NGA	O5-C5-C6	2.16	110.60	107.20
2	G	3	NGC	C3-C4-C5	-2.14	108.87	111.46
2	I	1	NGA	C3-C4-C5	2.09	113.97	110.24
2	G	1	NGA	O5-C5-C6	2.03	110.39	107.20
2	G	1	NGA	C1-C2-N2	2.03	113.95	110.49
2	H	3	NGC	C8-C7-C6	2.02	116.87	113.03
2	H	1	NGA	C8-C7-N2	-2.01	112.69	116.10
2	I	2	GAL	O5-C5-C6	2.01	110.35	107.20

All (60) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	G	2	GAL	C2
2	G	2	GAL	C5
2	G	2	GAL	C3
2	G	2	GAL	C1
2	G	2	GAL	C4
2	F	3	NGC	C5
2	F	3	NGC	C8
2	F	3	NGC	C6
2	F	3	NGC	C4
2	F	3	NGC	C7
2	G	1	NGA	C2
2	G	1	NGA	C5
2	G	1	NGA	C3
2	G	1	NGA	C1
2	G	1	NGA	C4
2	H	3	NGC	C5
2	H	3	NGC	C8
2	H	3	NGC	C6
2	H	3	NGC	C4
2	H	3	NGC	C7
2	I	3	NGC	C5
2	I	3	NGC	C8
2	I	3	NGC	C6
2	I	3	NGC	C4
2	I	3	NGC	C7
2	I	2	GAL	C2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atom
2	I	2	GAL	C5
2	I	2	GAL	C3
2	I	2	GAL	C1
2	I	2	GAL	C4
2	G	3	NGC	C5
2	G	3	NGC	C8
2	G	3	NGC	C6
2	G	3	NGC	C4
2	G	3	NGC	C7
2	I	1	NGA	C2
2	I	1	NGA	C5
2	I	1	NGA	C3
2	I	1	NGA	C1
2	I	1	NGA	C4
2	F	1	NGA	C2
2	F	1	NGA	C5
2	F	1	NGA	C3
2	F	1	NGA	C1
2	F	1	NGA	C4
2	F	2	GAL	C2
2	F	2	GAL	C5
2	F	2	GAL	C3
2	F	2	GAL	C1
2	F	2	GAL	C4
2	H	1	NGA	C2
2	H	1	NGA	C5
2	H	1	NGA	C3
2	H	1	NGA	C1
2	H	1	NGA	C4
2	H	2	GAL	C2
2	H	2	GAL	C5
2	H	2	GAL	C3
2	H	2	GAL	C1
2	H	2	GAL	C4

All (41) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	3	NGC	O10-C10-C11-O11
2	G	1	NGA	C8-C7-N2-C2
2	G	1	NGA	O7-C7-N2-C2
2	H	3	NGC	O8-C8-C9-O9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	H	3	NGC	O10-C10-C11-O11
2	I	3	NGC	C4-C5-N5-C10
2	F	1	NGA	C3-C2-N2-C7
2	F	1	NGA	C8-C7-N2-C2
2	F	1	NGA	O7-C7-N2-C2
2	H	1	NGA	C8-C7-N2-C2
2	H	1	NGA	O7-C7-N2-C2
2	I	3	NGC	O8-C8-C9-O9
2	G	3	NGC	O8-C8-C9-O9
2	G	1	NGA	O5-C5-C6-O6
2	I	2	GAL	O5-C5-C6-O6
2	H	3	NGC	C7-C8-C9-O9
2	I	3	NGC	C7-C8-C9-O9
2	G	3	NGC	C7-C8-C9-O9
2	I	2	GAL	C4-C5-C6-O6
2	F	1	NGA	C4-C5-C6-O6
2	F	1	NGA	O5-C5-C6-O6
2	I	1	NGA	C4-C5-C6-O6
2	I	1	NGA	O5-C5-C6-O6
2	G	1	NGA	C4-C5-C6-O6
2	G	2	GAL	O5-C5-C6-O6
2	F	3	NGC	C4-C5-N5-C10
2	H	3	NGC	C4-C5-N5-C10
2	I	1	NGA	C3-C2-N2-C7
2	G	1	NGA	C1-C2-N2-C7
2	I	3	NGC	O10-C10-C11-O11
2	G	3	NGC	C6-C5-N5-C10
2	G	3	NGC	C4-C5-N5-C10
2	I	3	NGC	C5-C6-C7-C8
2	H	2	GAL	C4-C5-C6-O6
2	F	2	GAL	C4-C5-C6-O6
2	H	1	NGA	C1-C2-N2-C7
2	F	2	GAL	O5-C5-C6-O6
2	F	3	NGC	C6-C5-N5-C10
2	H	3	NGC	C6-C5-N5-C10
2	H	2	GAL	O5-C5-C6-O6
2	H	1	NGA	C3-C2-N2-C7

All (3) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	3	NGC	C2-C3-C4-C5-C6-O6

*Continued on next page...*

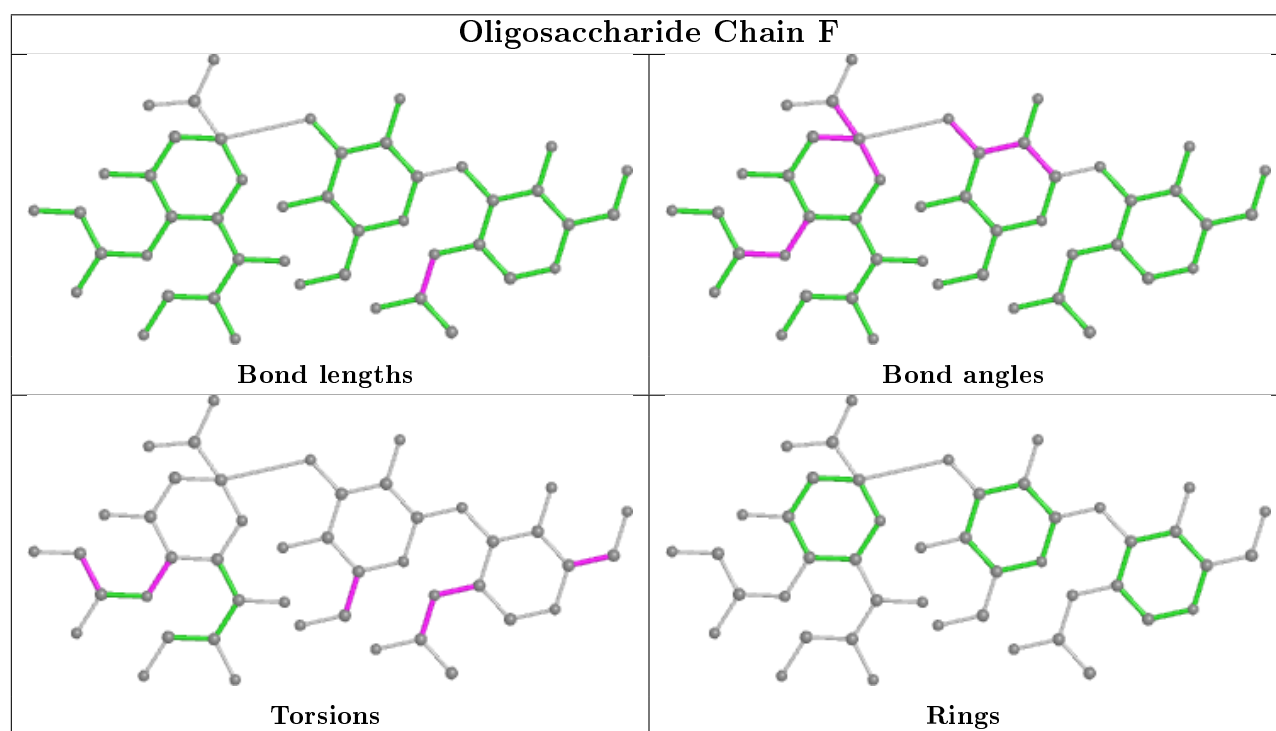
Continued from previous page...

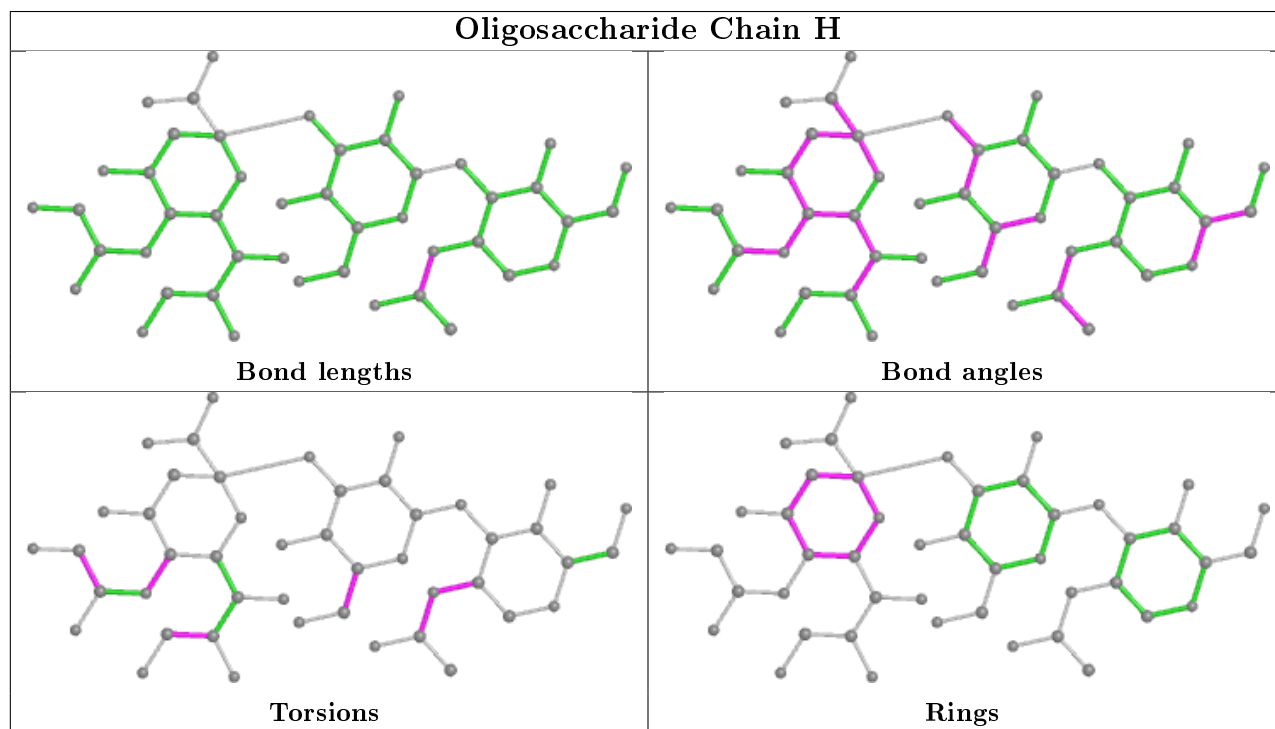
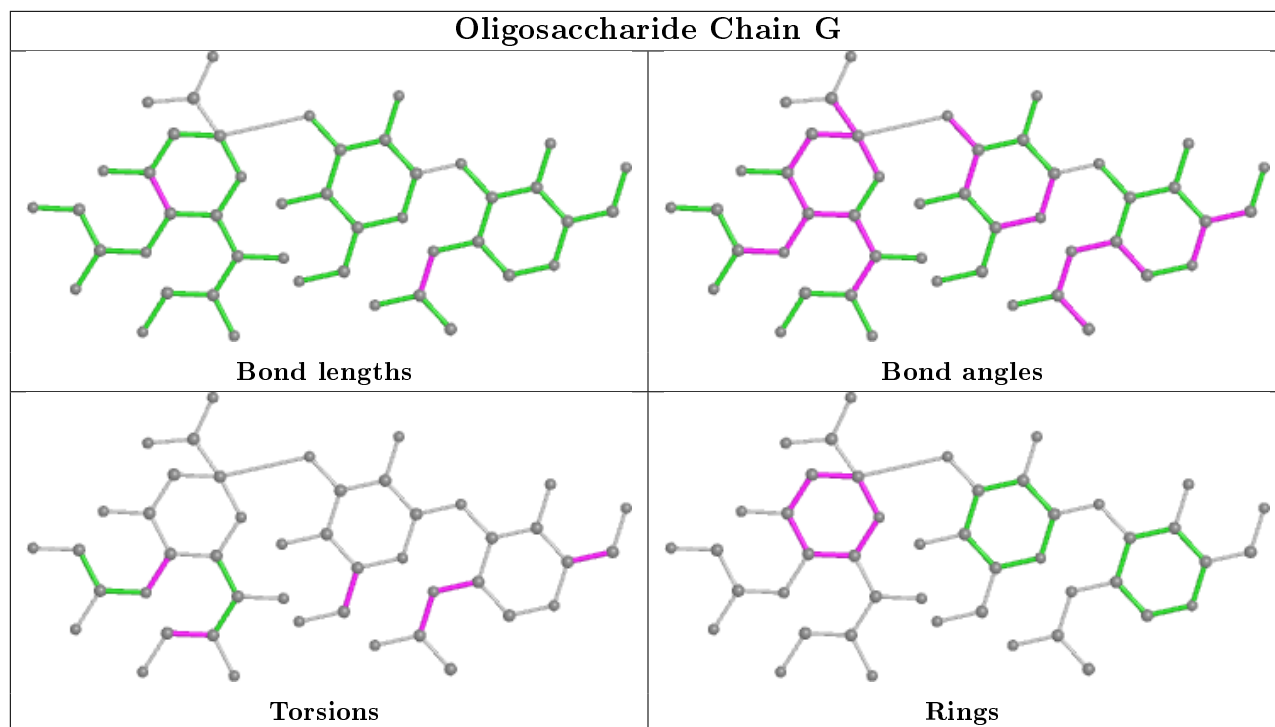
Mol	Chain	Res	Type	Atoms
2	H	3	NGC	C2-C3-C4-C5-C6-O6
2	I	3	NGC	C2-C3-C4-C5-C6-O6

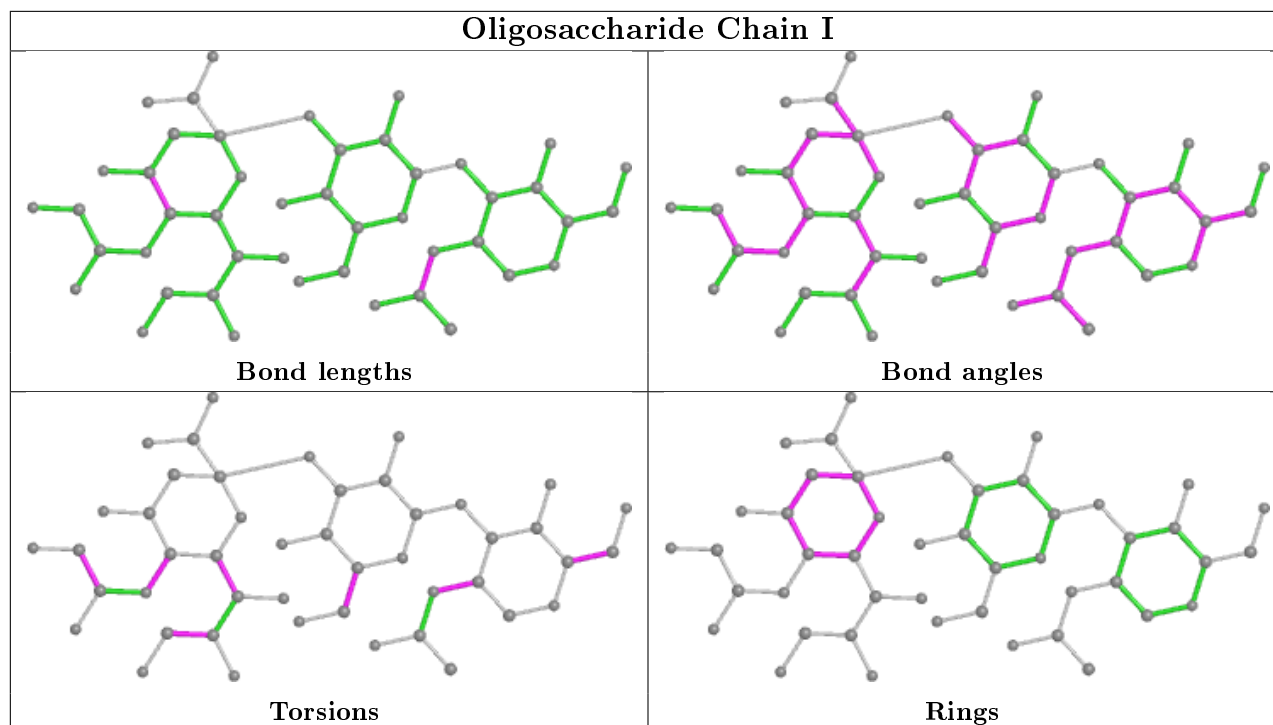
6 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	2	GAL	1	0
2	I	3	NGC	1	0
2	I	2	GAL	1	0
2	H	3	NGC	1	0
2	H	2	GAL	1	0
2	G	3	NGC	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](#)

13 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	$\# Z  > 2$
3	POL	E	130	2,4	3,3,3	0.33	0	2,2,2	0.43	0
5	1PE	A	132	-	15,15,15	0.47	0	14,14,14	0.24	0
3	POL	C	130	2,4	3,3,3	0.40	0	2,2,2	0.22	0
5	1PE	B	127	-	15,15,15	0.47	0	14,14,14	0.25	0
4	AZI	E	131	3	0,2,2	0.00	-	0,1,1	0.00	-
5	1PE	D	132	-	15,15,15	0.47	0	14,14,14	0.25	0
4	AZI	C	131	3	0,2,2	0.00	-	0,1,1	0.00	-
5	1PE	E	132	-	15,15,15	0.47	0	14,14,14	0.27	0
4	AZI	A	131	3	0,2,2	0.00	-	0,1,1	0.00	-
5	1PE	D	133	-	15,15,15	0.47	0	14,14,14	0.25	0
3	POL	D	130	2,4	3,3,3	0.42	0	2,2,2	0.29	0
3	POL	A	130	2,4	3,3,3	0.37	0	2,2,2	0.27	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	AZI	D	131	3	0,2,2	0.00	-	0,1,1	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	POL	E	130	2,4	-	1/1/1/1	-
5	1PE	A	132	-	-	8/13/13/13	-
5	1PE	B	127	-	-	9/13/13/13	-
3	POL	C	130	2,4	-	0/1/1/1	-
5	1PE	E	132	-	-	8/13/13/13	-
5	1PE	D	132	-	-	7/13/13/13	-
3	POL	D	130	2,4	-	0/1/1/1	-
3	POL	A	130	2,4	-	1/1/1/1	-
5	1PE	D	133	-	-	9/13/13/13	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (43) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	D	132	1PE	OH5-C14-C24-OH4
5	D	132	1PE	OH6-C15-C25-OH5
5	E	132	1PE	OH6-C15-C25-OH5
5	D	133	1PE	OH4-C13-C23-OH3
5	A	132	1PE	OH4-C13-C23-OH3
5	B	127	1PE	OH5-C14-C24-OH4
5	A	132	1PE	OH6-C15-C25-OH5
5	E	132	1PE	OH4-C13-C23-OH3
5	D	133	1PE	OH5-C14-C24-OH4
5	A	132	1PE	OH7-C16-C26-OH6
5	D	133	1PE	OH2-C12-C22-OH3
5	D	133	1PE	OH6-C15-C25-OH5
5	B	127	1PE	C12-C22-OH3-C23
5	B	127	1PE	C15-C25-OH5-C14
5	D	132	1PE	C12-C22-OH3-C23
5	E	132	1PE	C12-C22-OH3-C23

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	D	133	1PE	C14-C24-OH4-C13
5	E	132	1PE	OH5-C14-C24-OH4
5	B	127	1PE	C25-C15-OH6-C26
5	D	132	1PE	C23-C13-OH4-C24
5	E	132	1PE	C23-C13-OH4-C24
5	E	132	1PE	C25-C15-OH6-C26
5	E	132	1PE	C16-C26-OH6-C15
5	A	132	1PE	C23-C13-OH4-C24
5	B	127	1PE	C14-C24-OH4-C13
5	D	133	1PE	C15-C25-OH5-C14
5	A	132	1PE	C25-C15-OH6-C26
5	B	127	1PE	C23-C13-OH4-C24
5	B	127	1PE	C16-C26-OH6-C15
5	E	132	1PE	C14-C24-OH4-C13
5	D	133	1PE	C23-C13-OH4-C24
5	A	132	1PE	C24-C14-OH5-C25
5	D	133	1PE	C16-C26-OH6-C15
5	D	132	1PE	C24-C14-OH5-C25
5	D	132	1PE	OH4-C13-C23-OH3
5	A	132	1PE	C12-C22-OH3-C23
5	D	132	1PE	OH7-C16-C26-OH6
5	B	127	1PE	OH4-C13-C23-OH3
5	B	127	1PE	OH6-C15-C25-OH5
5	D	133	1PE	C13-C23-OH3-C22
3	A	130	POL	O-C1-C2-C3
3	E	130	POL	O-C1-C2-C3
5	A	132	1PE	OH5-C14-C24-OH4

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	127	1PE	1	0
5	D	132	1PE	1	0

## 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

### 6.1 Protein, DNA and RNA chains

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	116/126 (92%)	-0.21	5 (4%) 35 41	15, 20, 29, 30	0
1	B	117/126 (92%)	-0.26	2 (1%) 70 74	15, 20, 29, 32	0
1	C	120/126 (95%)	-0.17	4 (3%) 46 53	15, 20, 29, 37	0
1	D	117/126 (92%)	-0.03	5 (4%) 35 41	15, 20, 29, 32	0
1	E	116/126 (92%)	0.12	8 (6%) 16 21	15, 20, 28, 30	0
All	All	586/630 (93%)	-0.11	24 (4%) 37 43	15, 20, 29, 37	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	38	ALA	4.4
1	E	39	GLY	4.1
1	D	37	SER	4.0
1	A	37	SER	3.9
1	D	117	ASN	3.4
1	E	38	ALA	3.3
1	D	38	ALA	3.3
1	C	50	SER	3.1
1	E	32	ILE	2.8
1	A	51	SER	2.8
1	C	119	LEU	2.8
1	A	50	SER	2.7
1	B	38	ALA	2.7
1	A	116	LYS	2.6
1	C	120	GLU	2.6
1	E	37	SER	2.5
1	C	51	SER	2.5
1	E	51	SER	2.5
1	E	60	LEU	2.4
1	D	39	GLY	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	E	44	ALA	2.3
1	E	100	LEU	2.1
1	D	40	SER	2.0
1	B	37	SER	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](#)

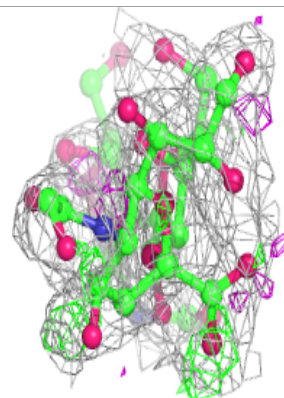
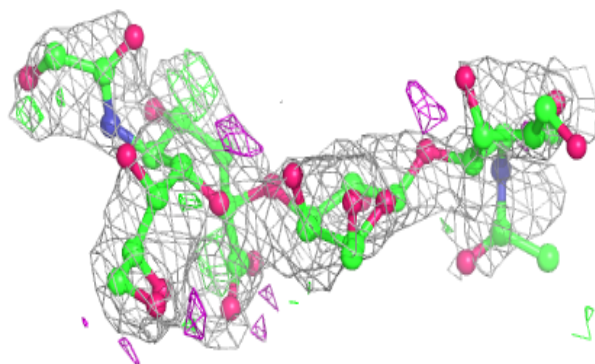
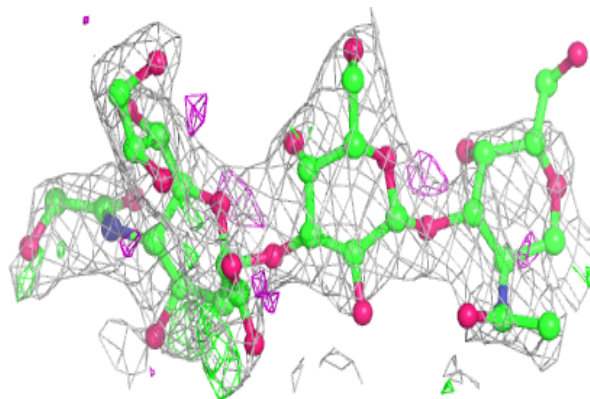
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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	NGC	F	3	21/22	0.72	0.23	66,67,68,68	0
2	NGC	H	3	21/22	0.72	0.22	45,47,47,48	0
2	NGA	F	1	14/15	0.72	0.48	71,72,72,72	0
2	NGA	H	1	14/15	0.77	0.38	54,56,57,57	0
2	NGA	G	1	14/15	0.79	0.27	47,50,50,50	0
2	GAL	F	2	11/12	0.83	0.25	68,69,70,70	0
2	NGC	G	3	21/22	0.85	0.17	32,34,35,36	0
2	GAL	H	2	11/12	0.85	0.28	49,52,52,53	0
2	GAL	G	2	11/12	0.87	0.24	39,44,45,45	0
2	NGA	I	1	14/15	0.87	0.12	27,28,29,30	0
2	NGC	I	3	21/22	0.89	0.13	22,26,30,31	0
2	GAL	I	2	11/12	0.89	0.14	25,25,26,28	0

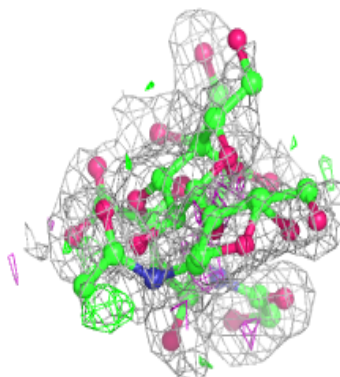
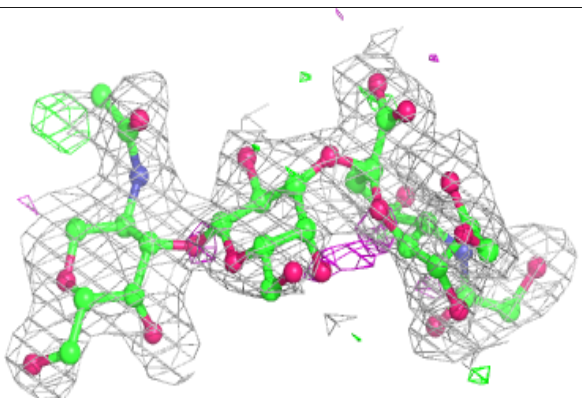
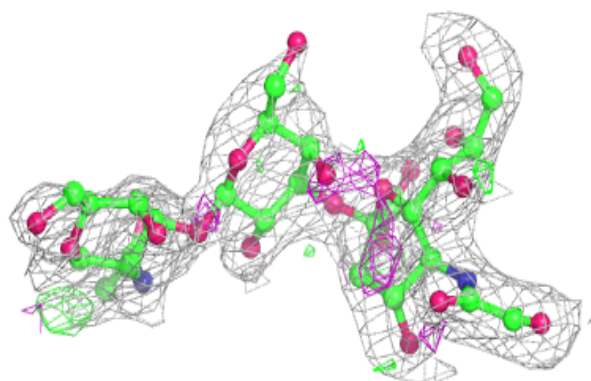
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**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)

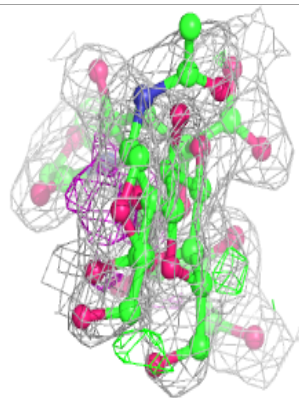
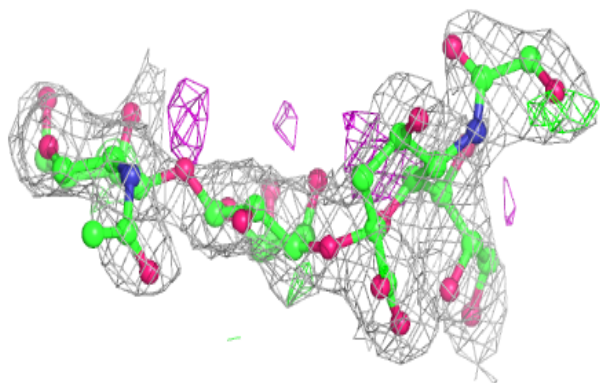
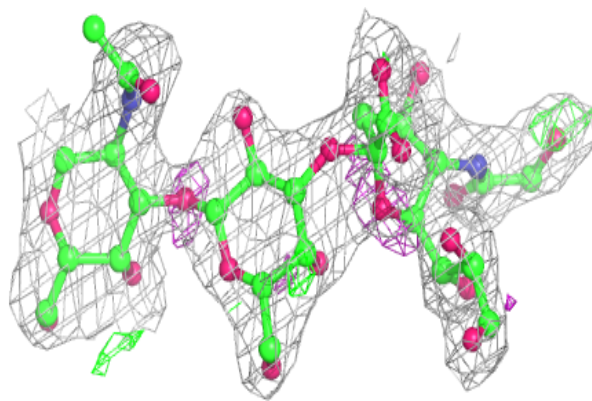
**Electron density around Chain G:**

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

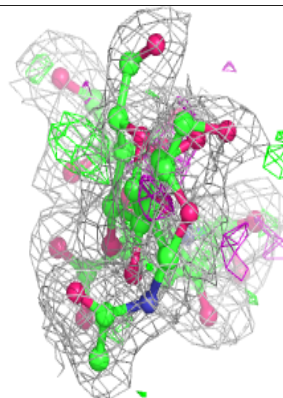
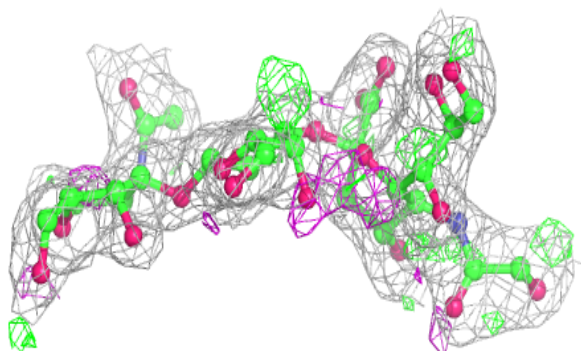
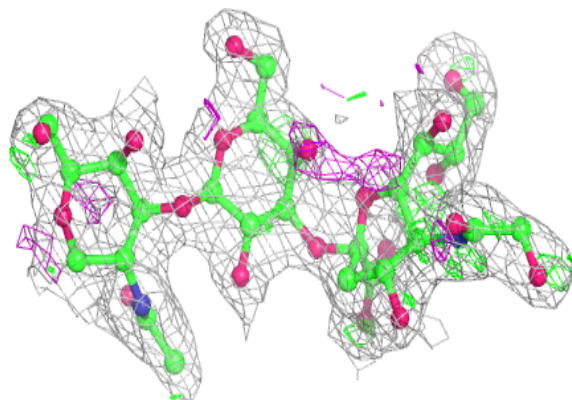


**Electron density around Chain H:**

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

**Electron density around Chain I:**

$2mF_o-DF_c$  (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
5	1PE	E	132	16/16	0.22	0.39	76,78,80,80	0
5	1PE	B	127	16/16	0.34	0.39	82,83,85,85	0
4	AZI	A	131	3/3	0.62	0.32	72,72,72,72	0
5	1PE	D	132	16/16	0.66	0.23	57,59,59,59	0
4	AZI	E	131	3/3	0.67	0.25	36,36,37,37	0
4	AZI	D	131	3/3	0.67	0.54	59,59,60,60	0
5	1PE	D	133	16/16	0.73	0.27	70,70,70,70	0
3	POL	C	130	4/4	0.73	0.49	51,52,52,53	0
5	1PE	A	132	16/16	0.77	0.20	51,52,53,53	0
4	AZI	C	131	3/3	0.78	0.44	53,53,53,53	0
3	POL	A	130	4/4	0.83	0.45	72,72,72,72	0
3	POL	D	130	4/4	0.84	0.51	57,58,58,59	0
3	POL	E	130	4/4	0.93	0.12	30,31,33,35	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.