



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 3, 2023 – 02:02 PM EDT

PDB ID : 6P3R  
Title : crystal structure of human Fab H5.31 in complex with influenza A H5N1 Vietnam hemagglutinin head domain  
Authors : Dong, J.; Crowe, J.E.  
Deposited on : 2019-05-24  
Resolution : 3.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The 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

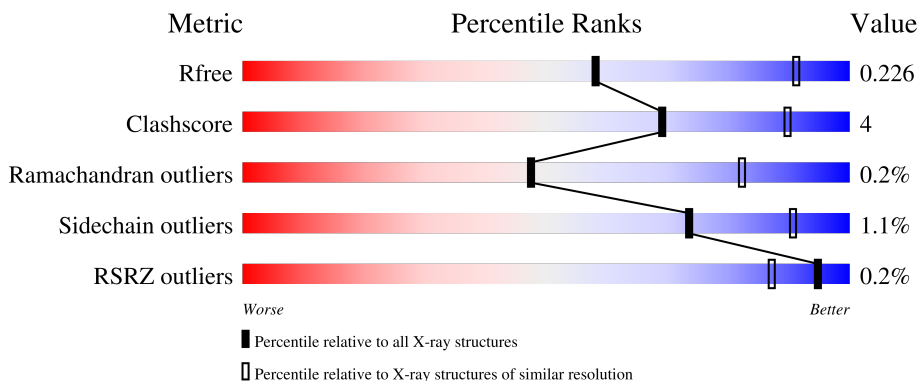
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.




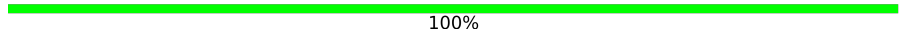

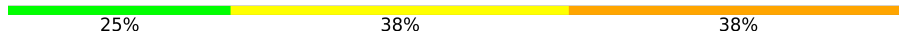
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	230	 87% 10%
1	C	230	 88% 12%
2	B	215	 93% 6%
2	D	215	 95% 5%
3	E	219	 77% 15% 8%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	L	219	 91% 5%
4	F	2	 100%
5	G	8	 12% 50% 38%
6	H	8	 25% 38% 38%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	NAG	F	1	-	-	-	X
4	NAG	F	2	-	-	-	X

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 9959 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 human Fab H5.31 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	224	Total	C	N	O	S	0	0	0
			1654	1051	282	315	6			
1	C	230	Total	C	N	O	S	0	0	0
			1712	1081	292	333	6			

- Molecule 2 is a protein called Human Fab H5.31 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	215	Total	C	N	O	S	0	0	0
			1621	1013	269	332	7			
2	D	215	Total	C	N	O	S	0	0	0
			1641	1027	273	334	7			

- Molecule 3 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	E	202	Total	C	N	O	S	0	0	0
			1486	938	264	279	5			
3	L	208	Total	C	N	O	S	0	0	0
			1506	954	261	285	6			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	269	ALA	-	expression tag	UNP Q1KHK2
E	270	ALA	-	expression tag	UNP Q1KHK2
E	271	HIS	-	expression tag	UNP Q1KHK2
E	272	HIS	-	expression tag	UNP Q1KHK2
E	273	HIS	-	expression tag	UNP Q1KHK2
E	274	HIS	-	expression tag	UNP Q1KHK2
E	275	HIS	-	expression tag	UNP Q1KHK2
E	276	HIS	-	expression tag	UNP Q1KHK2

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	269	ALA	-	expression tag	UNP Q1KHK2
L	270	ALA	-	expression tag	UNP Q1KHK2
L	271	HIS	-	expression tag	UNP Q1KHK2
L	272	HIS	-	expression tag	UNP Q1KHK2
L	273	HIS	-	expression tag	UNP Q1KHK2
L	274	HIS	-	expression tag	UNP Q1KHK2
L	275	HIS	-	expression tag	UNP Q1KHK2
L	276	HIS	-	expression tag	UNP Q1KHK2

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	F	2	28	16	2	10	0	0	0

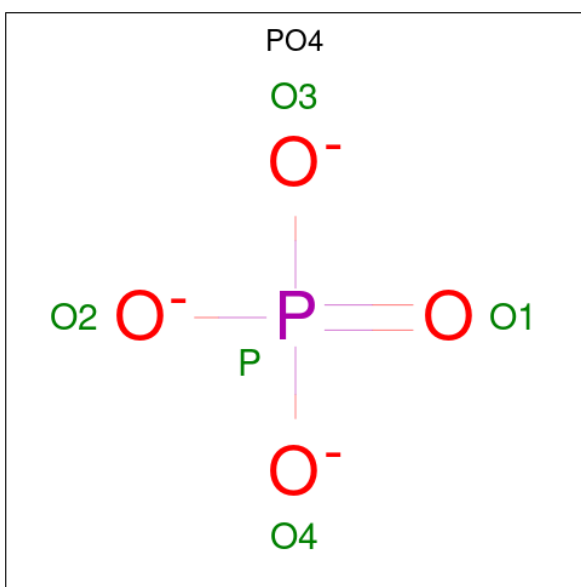
- Molecule 5 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-alpha-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	G	8	106	59	4	43	0	0	0

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

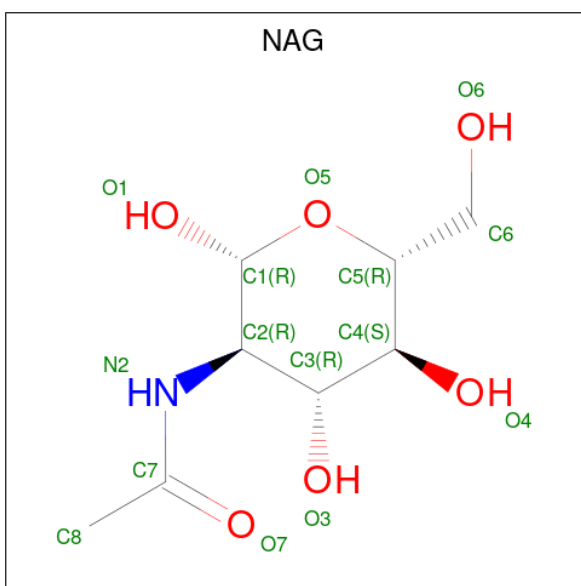
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
6	H	8	106	59	4	43	0	0	0

- Molecule 7 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total O P 5 4 1	0	0
7	D	1	Total O P 5 4 1	0	0
7	L	1	Total O P 5 4 1	0	0

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	C	1	Total	C	N	O	0	0
			14	8	1	5		
8	E	1	Total	C	N	O	0	0
			14	8	1	5		
8	L	1	Total	C	N	O	0	0
			14	8	1	5		


- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	11	Total	O	0	0
			11	11		
9	B	15	Total	O	0	0
			15	15		
9	C	5	Total	O	0	0
			5	5		
9	D	10	Total	O	0	0
			10	10		
9	L	1	Total	O	0	0
			1	1		

### 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: human Fab H5.31 heavy chain

Chain A:  87% 10% ..



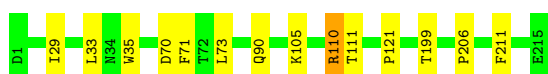
- Molecule 1: human Fab H5.31 heavy chain

Chain C:  88% 12%



- Molecule 2: Human Fab H5.31 light chain

Chain B:  93% 6%




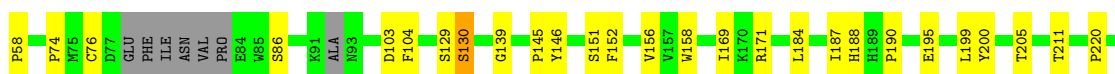
- Molecule 2: Human Fab H5.31 light chain

Chain D:  95% 5%



- Molecule 3: Hemagglutinin

Chain E:  77% 15% 8%





- Molecule 3: Hemagglutinin

Chain L:  91% 5%



- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:  100%




- Molecule 5: N-acetyl-alpha-neuraminic acid-(2-6)-alpha-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose-(1-2)-alpha-D-mannopyranose-(1-6)-[alpha-D-mannopyranose-(1-3)]alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-alpha-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:  12% 50% 38%



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

Chain H:  25% 38% 38%



## 4 Data and refinement statistics

Property	Value	Source
Space group	F 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	185.11Å 234.47Å 284.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.47 – 3.00 47.74 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.47-3.00) 100.0 (47.74-3.00)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.80 (at 3.01Å)	Xtrriage
Refinement program	PHENIX (1.15.2_3472: ???)	Depositor
R, $R_{free}$	0.192 , 0.226 0.192 , 0.226	Depositor DCC
$R_{free}$ test set	3062 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	82.9	Xtrriage
Anisotropy	0.088	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 68.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	9959	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	87.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GLA, NAG, GAL, MAN, BMA, SIA, PO4, NDG, Z9N

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	0.25	0/1692	0.47	0/2307
1	C	0.26	0/1751	0.47	0/2383
2	B	0.26	0/1656	0.47	0/2254
2	D	0.26	0/1676	0.48	0/2275
3	E	0.24	0/1526	0.43	0/2086
3	L	0.24	0/1549	0.42	0/2129
All	All	0.25	0/9850	0.46	0/13434

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

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	1654	0	1590	15	0
1	C	1712	0	1665	15	0
2	B	1621	0	1548	8	0
2	D	1641	0	1596	8	0
3	E	1486	0	1251	17	2
3	L	1506	0	1252	5	0
4	F	28	0	25	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	G	106	0	87	4	2
6	H	106	0	76	3	0
7	B	5	0	0	0	0
7	D	5	0	0	0	0
7	L	5	0	0	0	0
8	C	14	0	13	0	0
8	E	14	0	13	0	0
8	L	14	0	13	0	0
9	A	11	0	0	0	0
9	B	15	0	0	1	0
9	C	5	0	0	0	0
9	D	10	0	0	0	0
9	L	1	0	0	0	0
All	All	9959	0	9129	71	2

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 (71) 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:67:ARG:NH1	1:C:90:ASP:OD2	2.25	0.70
1:A:6:GLN:NE2	1:A:96:CYS:SG	2.62	0.70
1:C:6:GLN:HE21	1:C:96:CYS:H	1.40	0.69
3:E:74:PRO:HB2	3:E:146:TYR:HB2	1.75	0.68
2:B:70:ASP:OD2	2:D:18:ARG:NH2	2.21	0.67
1:A:135:PRO:HB3	1:A:161:TYR:HB3	1.79	0.64
1:A:6:GLN:HE22	1:A:120:GLY:HA3	1.63	0.63
3:E:190:PRO:O	3:E:225:ARG:NH2	2.33	0.61
1:C:135:PRO:HB3	1:C:161:TYR:HB3	1.82	0.61
5:G:2:NDG:HB	5:G:4:MAN:HO4	1.47	0.60
1:A:139:PRO:HD3	1:A:225:LYS:HE2	1.87	0.56
3:E:211:THR:HG22	3:E:248:ILE:HA	1.89	0.55
1:A:58:LYS:HD2	1:A:70:ILE:HD11	1.89	0.54
1:C:83:MET:HB3	1:C:86:LEU:HD21	1.91	0.52
5:G:6:GLA:O6	5:G:6:GLA:O4	2.16	0.51
3:E:58:PRO:HG2	3:E:86:SER:H	1.77	0.50
2:D:3:VAL:HG22	2:D:26:SER:HB3	1.93	0.49
1:A:6:GLN:NE2	1:A:96:CYS:H	2.11	0.49
1:A:83:MET:HB3	1:A:86:LEU:HD21	1.95	0.48
2:B:105:LYS:NZ	9:B:401:HOH:O	2.46	0.48

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:91:THR:HG23	1:C:126:THR:HA	1.95	0.48
5:G:2:NDG:H8C1	5:G:2:NDG:H2	1.62	0.48
3:E:205:THR:HG21	3:E:255:ASN:OD1	2.13	0.48
3:E:130:SER:OG	3:E:171:ARG:NH2	2.46	0.47
1:C:83:MET:HE1	1:C:125:VAL:HG21	1.96	0.47
3:E:188:HIS:HB2	3:E:257:ILE:HD11	1.97	0.46
1:C:168:VAL:HG22	1:C:214:VAL:HG22	1.97	0.46
3:E:195:GLU:HG2	3:E:199:LEU:HD13	1.96	0.46
6:H:5:NDG:H2	6:H:5:NDG:H8C1	1.50	0.46
2:D:33:LEU:HD22	2:D:71:PHE:CG	2.51	0.45
3:E:184:LEU:O	3:E:259:PRO:HB3	2.15	0.45
2:B:199:THR:HG22	2:B:206:PRO:HB3	1.98	0.45
1:A:79:LEU:HD12	1:A:79:LEU:HA	1.86	0.45
1:C:71:SER:HB3	1:C:80:TYR:HB2	1.98	0.45
3:E:205:THR:HG23	3:E:220:PRO:HG2	1.99	0.45
5:G:5:NDG:H2	5:G:5:NDG:H8C1	1.52	0.45
1:A:6:GLN:NE2	1:A:120:GLY:HA3	2.31	0.45
3:E:139:GLY:HA3	3:E:158:TRP:HB3	1.98	0.45
1:C:100:PHE:CE2	1:C:102:GLU:HB3	2.52	0.45
1:C:146:SER:HA	2:D:118:PHE:HB3	1.99	0.45
3:E:187:ILE:HD11	3:E:220:PRO:HG3	1.99	0.44
2:B:29:ILE:HG21	2:B:90:GLN:HG3	1.99	0.44
6:H:5:NDG:H4	6:H:6:GAL:H2	1.86	0.44
3:E:103:ASP:OD1	3:E:104:PHE:N	2.50	0.44
2:B:110:ARG:HD3	2:B:111:THR:O	2.18	0.44
3:L:184:LEU:O	3:L:259:PRO:HB3	2.17	0.43
1:A:28:THR:HG22	1:A:30:SER:H	1.82	0.43
2:B:33:LEU:HD22	2:B:71:PHE:CG	2.54	0.43
3:L:130:SER:O	3:L:130:SER:OG	2.32	0.43
3:E:130:SER:O	3:E:171:ARG:NH2	2.52	0.43
2:B:35:TRP:CE2	2:B:73:LEU:HB2	2.54	0.43
2:B:121:PRO:HB3	2:B:211:PHE:CE2	2.54	0.42
3:L:205:THR:HG21	3:L:255:ASN:OD1	2.19	0.42
1:A:64:VAL:HB	1:A:68:PHE:CG	2.55	0.42
1:A:18:LEU:HD12	1:A:18:LEU:HA	1.92	0.42
1:C:148:SER:HB2	2:D:118:PHE:CE1	2.55	0.42
1:C:230:LYS:HE3	2:D:124:ASP:OD2	2.20	0.42
3:E:151:SER:OG	3:E:152:PHE:N	2.52	0.42
3:L:169:ILE:O	3:L:251:GLU:HA	2.20	0.42
2:D:165:VAL:HG12	2:D:177:LEU:HD12	2.01	0.42
1:C:11:LEU:HB2	1:C:163:PRO:HG3	2.02	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:98:ARG:HB3	1:A:118:VAL:CG1	2.50	0.41
1:C:72:ARG:HG2	1:C:74:ASN:OD1	2.20	0.41
2:D:33:LEU:HD11	2:D:88:CYS:HB2	2.02	0.41
3:E:156:VAL:CG1	3:E:257:ILE:HG22	2.51	0.41
1:A:159:LYS:HA	1:A:193:SER:HB2	2.02	0.41
1:C:200:VAL:HG11	1:C:210:TYR:CE1	2.56	0.41
1:A:28:THR:HG22	1:A:30:SER:N	2.36	0.40
3:E:169:ILE:O	3:E:251:GLU:HA	2.22	0.40
3:L:196:GLN:NE2	3:L:255:ASN:OD1	2.48	0.40
6:H:3:BMA:H62	6:H:4:BMA:H2	1.94	0.40

All (2) 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)
3:E:195:GLU:OE2	5:G:7:SIA:C9[3_555]	1.38	0.82
3:E:195:GLU:OE2	5:G:7:SIA:O9[3_555]	1.68	0.52

## 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	220/230 (96%)	213 (97%)	6 (3%)	1 (0%)	29	68
1	C	228/230 (99%)	224 (98%)	4 (2%)	0	100	100
2	B	213/215 (99%)	203 (95%)	10 (5%)	0	100	100
2	D	213/215 (99%)	205 (96%)	8 (4%)	0	100	100
3	E	196/219 (90%)	181 (92%)	13 (7%)	2 (1%)	15	53
3	L	204/219 (93%)	188 (92%)	16 (8%)	0	100	100
All	All	1274/1328 (96%)	1214 (95%)	57 (4%)	3 (0%)	47	82

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	129	SER
3	E	145	PRO
1	A	30	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	173/193 (90%)	169 (98%)	4 (2%)	50	80
1	C	185/193 (96%)	183 (99%)	2 (1%)	73	90
2	B	183/189 (97%)	182 (100%)	1 (0%)	88	96
2	D	188/189 (100%)	188 (100%)	0	100	100
3	E	134/197 (68%)	131 (98%)	3 (2%)	52	81
3	L	135/197 (68%)	134 (99%)	1 (1%)	84	94
All	All	998/1158 (86%)	987 (99%)	11 (1%)	73	90

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

Mol	Chain	Res	Type
1	A	58	LYS
1	A	103	ARG
1	A	154	LEU
1	A	186	LEU
2	B	110	ARG
1	C	58	LYS
1	C	154	LEU
3	E	76	CYS
3	E	130	SER
3	E	200	TYR
3	L	144	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

18 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
4	NAG	F	1	1,4	14,14,15	0.29	0	17,19,21	0.48	0
4	NAG	F	2	4	14,14,15	0.27	0	17,19,21	0.41	0
5	NAG	G	1	5	14,14,15	0.33	0	17,19,21	0.76	0
5	NDG	G	2	5	14,14,15	0.35	0	17,19,21	0.92	1 (5%)
5	MAN	G	3	5	11,11,12	0.30	0	15,15,17	1.22	2 (13%)
5	MAN	G	4	5	11,11,12	0.39	0	15,15,17	1.65	2 (13%)
5	NDG	G	5	5	14,14,15	0.28	0	17,19,21	0.65	0
5	GLA	G	6	5	11,11,12	0.24	0	15,15,17	1.01	0
5	SIA	G	7	5	20,20,21	0.55	0	24,28,31	1.05	2 (8%)
5	MAN	G	8	5	11,11,12	0.43	0	15,15,17	1.71	2 (13%)
6	NAG	H	1	6,3	13,13,15	0.33	0	16,18,21	0.52	0
6	NAG	H	2	6	14,14,15	0.30	0	17,19,21	0.63	0
6	BMA	H	3	6	11,11,12	0.30	0	15,15,17	0.71	1 (6%)
6	BMA	H	4	6	11,11,12	0.67	0	15,15,17	1.40	2 (13%)
6	NDG	H	5	6	14,14,15	0.29	0	17,19,21	0.83	0
6	GAL	H	6	6	11,11,12	0.43	0	15,15,17	1.58	3 (20%)
6	SIA	H	7	6	20,20,21	0.55	0	24,28,31	1.08	2 (8%)
6	Z9N	H	8	6	11,12,12	0.85	1 (9%)	10,18,18	4.33	1 (10%)

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
4	NAG	F	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	F	2	4	-	2/6/23/26	0/1/1/1
5	NAG	G	1	5	-	0/6/23/26	0/1/1/1
5	NDG	G	2	5	-	2/6/23/26	0/1/1/1
5	MAN	G	3	5	-	0/2/19/22	0/1/1/1
5	MAN	G	4	5	-	1/2/19/22	0/1/1/1
5	NDG	G	5	5	-	4/6/23/26	0/1/1/1
5	GLA	G	6	5	-	2/2/19/22	0/1/1/1
5	SIA	G	7	5	-	2/18/34/38	0/1/1/1
5	MAN	G	8	5	-	0/2/19/22	0/1/1/1
6	NAG	H	1	6,3	-	2/4/21/26	0/1/1/1
6	NAG	H	2	6	-	0/6/23/26	0/1/1/1
6	BMA	H	3	6	-	0/2/19/22	0/1/1/1
6	BMA	H	4	6	-	1/2/19/22	0/1/1/1
6	NDG	H	5	6	-	3/6/23/26	0/1/1/1
6	GAL	H	6	6	-	2/2/19/22	0/1/1/1
6	SIA	H	7	6	-	2/18/34/38	0/1/1/1
6	Z9N	H	8	6	-	2/5/24/24	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	8	Z9N	O2-C2	2.06	1.44	1.40

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	H	8	Z9N	O2-C2-O5	-13.43	83.57	109.50
5	G	8	MAN	C1-O5-C5	4.83	118.73	112.19
5	G	4	MAN	C1-O5-C5	4.78	118.66	112.19
6	H	4	BMA	O2-C2-C3	3.64	117.43	110.14
6	H	6	GAL	C1-C2-C3	3.61	114.10	109.67
5	G	3	MAN	C1-O5-C5	3.49	116.92	112.19
5	G	7	SIA	C6-O6-C2	3.05	117.86	111.34
5	G	8	MAN	C1-C2-C3	2.77	113.08	109.67
6	H	7	SIA	C6-O6-C2	2.70	117.12	111.34
6	H	6	GAL	O2-C2-C3	-2.57	104.99	110.14

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	3	MAN	C1-C2-C3	2.55	112.80	109.67
5	G	4	MAN	C3-C4-C5	2.54	114.78	110.24
6	H	3	BMA	O5-C5-C6	2.44	111.03	107.20
6	H	6	GAL	O5-C5-C6	2.33	110.86	107.20
6	H	4	BMA	C2-C3-C4	2.28	114.83	110.89
5	G	2	NDG	O4-C4-C3	2.18	115.38	110.35
5	G	7	SIA	O6-C2-C1	2.10	111.81	107.70
6	H	7	SIA	O6-C2-C1	2.09	111.81	107.70

There are no chirality outliers.

All (25) torsion outliers are listed below:

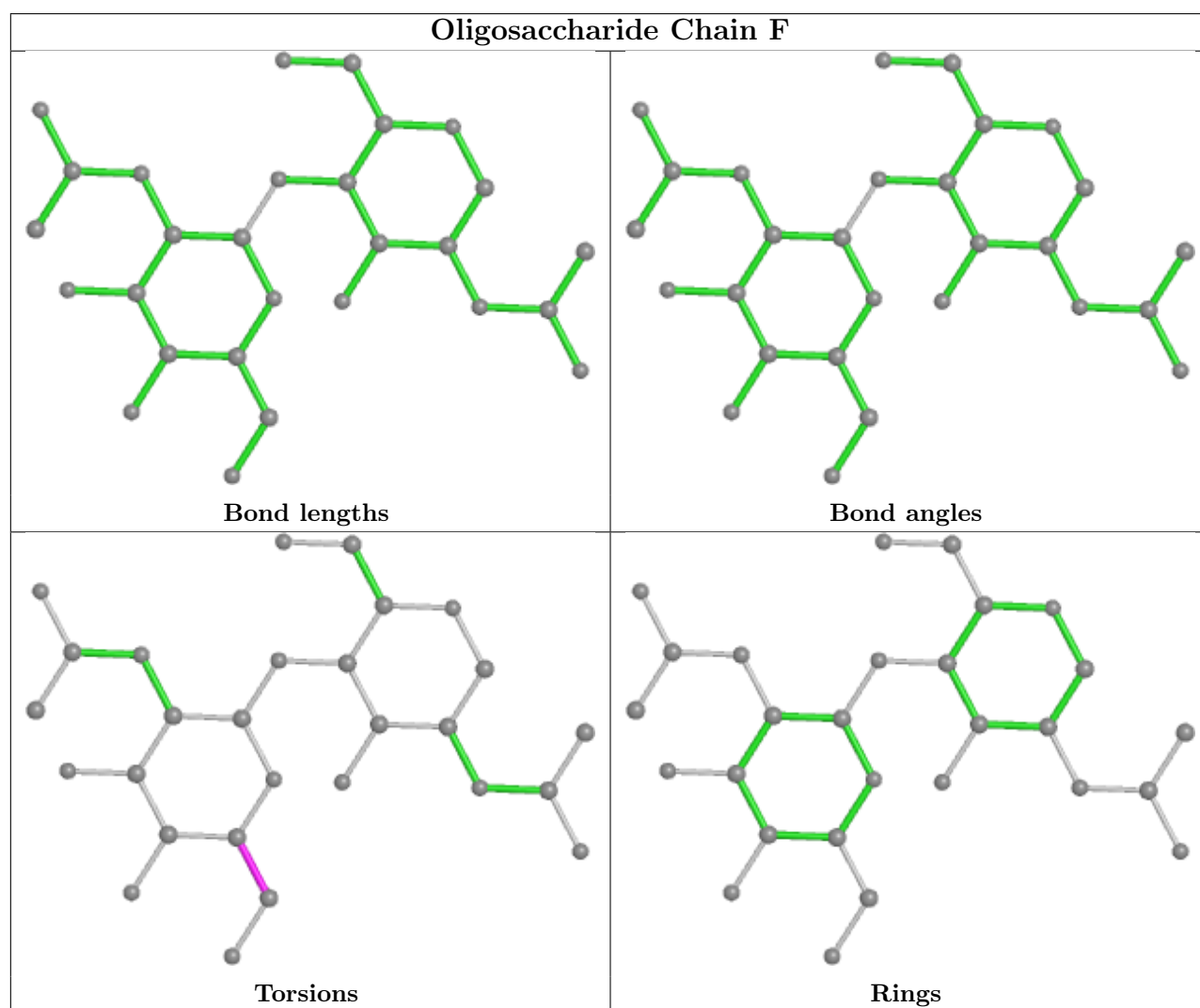
Mol	Chain	Res	Type	Atoms
6	H	7	SIA	C11-C10-N5-C5
6	H	7	SIA	O10-C10-N5-C5
6	H	8	Z9N	O1-C1-C2-O2
5	G	7	SIA	C11-C10-N5-C5
5	G	6	GLA	C4-C5-C6-O6
4	F	2	NAG	O5-C5-C6-O6
6	H	6	GAL	O5-C5-C6-O6
5	G	7	SIA	O10-C10-N5-C5
5	G	2	NDG	C8-C7-N2-C2
5	G	2	NDG	O7-C7-N2-C2
5	G	5	NDG	C8-C7-N2-C2
5	G	5	NDG	O7-C7-N2-C2
6	H	1	NAG	C8-C7-N2-C2
6	H	1	NAG	O7-C7-N2-C2
6	H	5	NDG	C8-C7-N2-C2
6	H	5	NDG	O7-C7-N2-C2
5	G	5	NDG	O5-C5-C6-O6
5	G	6	GLA	O5-C5-C6-O6
4	F	2	NAG	C4-C5-C6-O6
6	H	5	NDG	O5-C5-C6-O6
5	G	4	MAN	O5-C5-C6-O6
6	H	4	BMA	O5-C5-C6-O6
6	H	6	GAL	C4-C5-C6-O6
5	G	5	NDG	C4-C5-C6-O6
6	H	8	Z9N	O1-C1-C2-O5

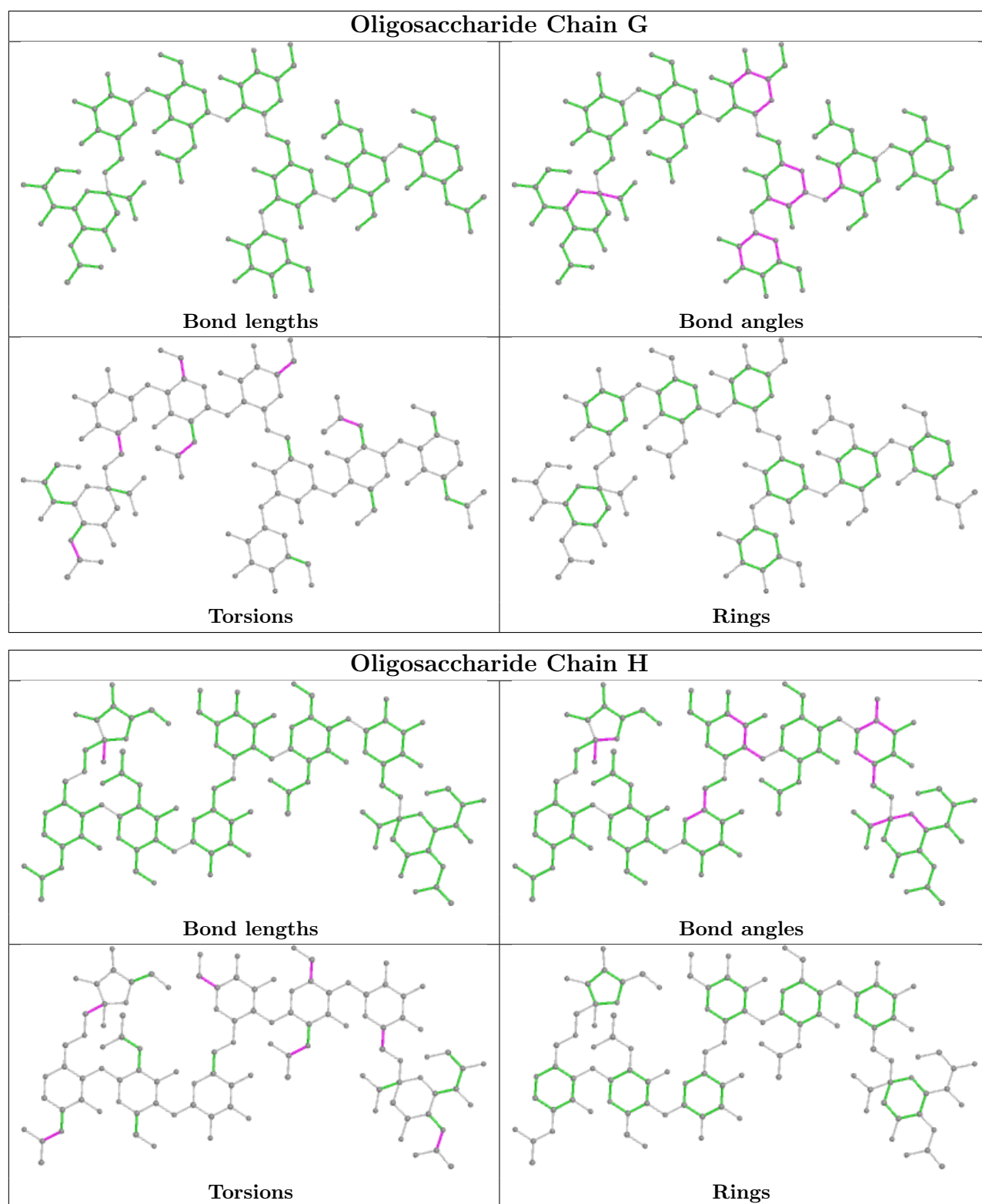
There are no ring outliers.

9 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	7	SIA	0	2
5	G	5	NDG	1	0
6	H	5	NDG	2	0
6	H	6	GAL	1	0
5	G	6	GLA	1	0
6	H	4	BMA	1	0
5	G	2	NDG	2	0
6	H	3	BMA	1	0
5	G	4	MAN	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](#)

6 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
7	PO4	B	301	-	4,4,4	0.91	0	6,6,6	0.43	0
7	PO4	D	301	-	4,4,4	0.90	0	6,6,6	0.43	0
7	PO4	L	301	-	4,4,4	0.90	0	6,6,6	0.45	0
8	NAG	C	301	1	14,14,15	0.23	0	17,19,21	0.45	0
8	NAG	L	302	3	14,14,15	0.94	2 (14%)	17,19,21	0.94	1 (5%)
8	NAG	E	301	3	14,14,15	0.33	0	17,19,21	0.52	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	NAG	C	301	1	-	2/6/23/26	0/1/1/1
8	NAG	L	302	3	-	1/6/23/26	0/1/1/1
8	NAG	E	301	3	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	L	302	NAG	O5-C1	2.74	1.48	1.43
8	L	302	NAG	C1-C2	2.08	1.55	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	L	302	NAG	C1-O5-C5	3.68	117.18	112.19

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	C	301	NAG	C8-C7-N2-C2
8	C	301	NAG	O7-C7-N2-C2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
8	L	302	NAG	O5-C5-C6-O6

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	224/230 (97%)	-0.13	0	100	100	48, 73, 141, 178	0
1	C	230/230 (100%)	-0.13	0	100	100	46, 78, 126, 152	0
2	B	215/215 (100%)	-0.12	0	100	100	40, 59, 102, 149	0
2	D	215/215 (100%)	-0.17	0	100	100	39, 61, 92, 144	0
3	E	202/219 (92%)	-0.03	1 (0%)	91	75	65, 115, 160, 176	0
3	L	208/219 (94%)	-0.07	1 (0%)	91	75	62, 113, 159, 186	0
All	All	1294/1328 (97%)	-0.11	2 (0%)	95	87	39, 79, 145, 186	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	E	245	ASN	2.8
3	L	245	ASN	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 [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.

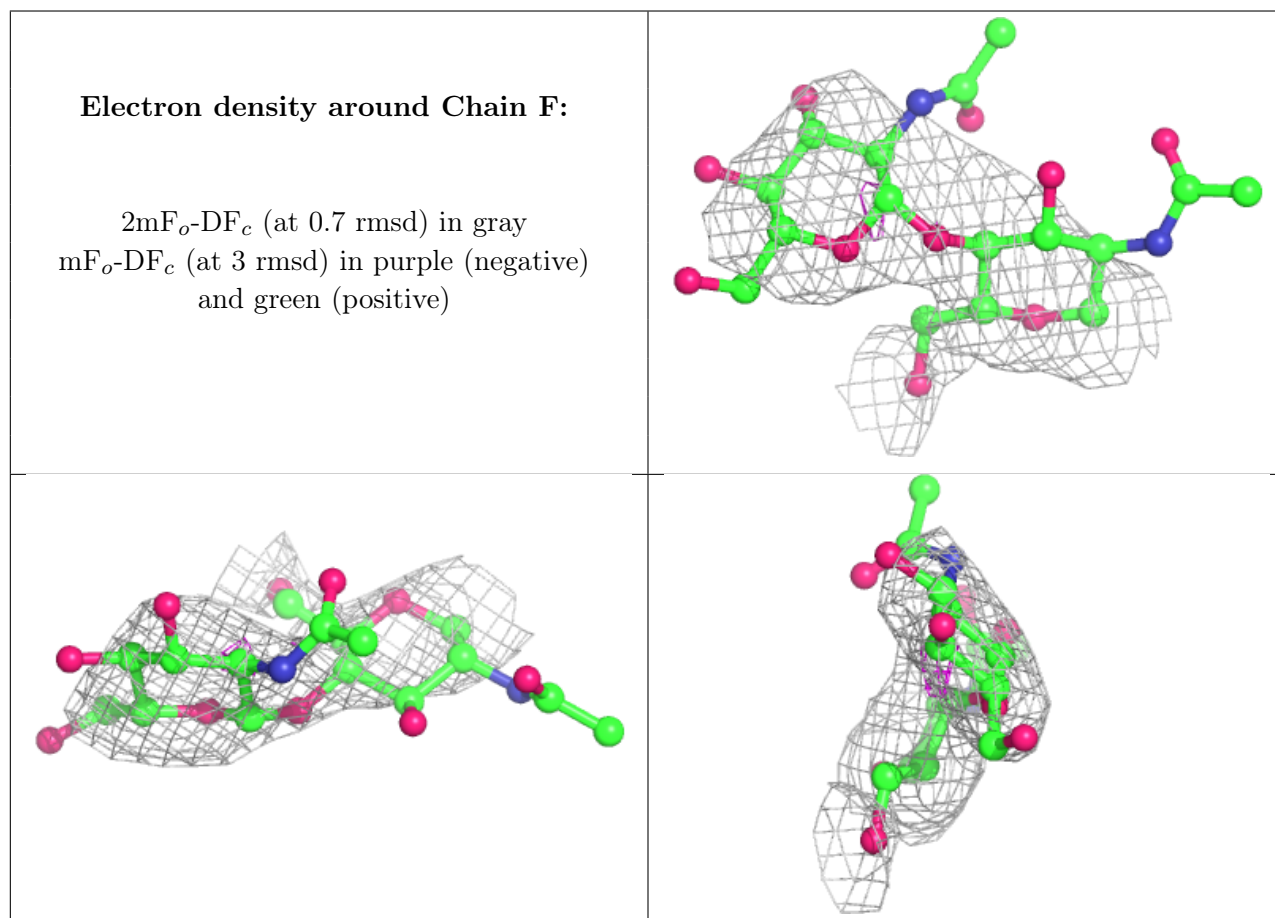
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	MAN	G	8	11/12	0.51	0.40	168,181,188,194	0
4	NAG	F	1	14/15	0.60	0.55	191,215,223,224	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	F	2	14/15	0.73	0.51	170,192,202,205	0
6	BMA	H	3	11/12	0.79	0.29	131,141,150,152	0
5	MAN	G	4	11/12	0.81	0.30	127,170,177,189	0
5	NDG	G	2	14/15	0.84	0.35	129,139,144,156	0
5	NAG	G	1	14/15	0.84	0.24	120,143,151,155	0
5	MAN	G	3	11/12	0.85	0.33	162,167,173,179	0
6	NAG	H	2	14/15	0.88	0.25	123,130,140,140	0
6	Z9N	H	8	12/12	0.90	0.50	158,174,182,182	0
6	GAL	H	6	11/12	0.91	0.18	114,116,146,148	0
6	SIA	H	7	20/21	0.91	0.29	113,131,142,143	0
5	NDG	G	5	14/15	0.91	0.21	131,142,156,157	0
6	NAG	H	1	13/15	0.92	0.22	135,138,154,156	0
5	GLA	G	6	11/12	0.92	0.16	115,126,142,144	0
6	BMA	H	4	11/12	0.93	0.20	128,132,145,153	0
5	SIA	G	7	20/21	0.93	0.26	108,119,154,157	0
6	NDG	H	5	14/15	0.95	0.29	111,137,148,157	0

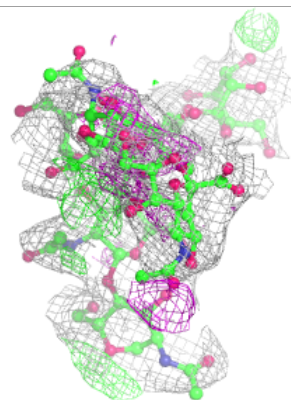
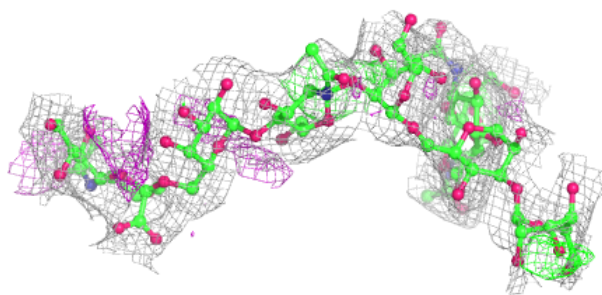
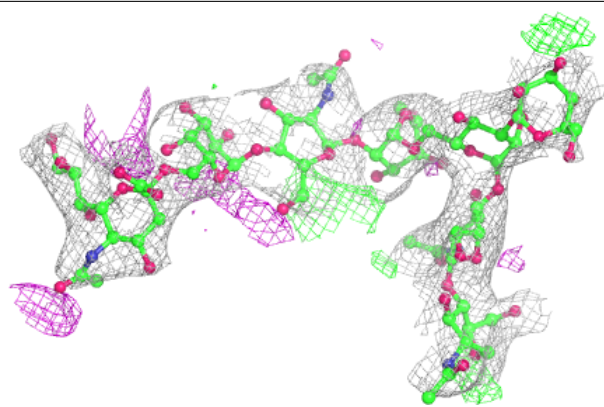
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



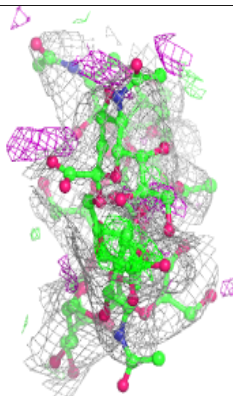
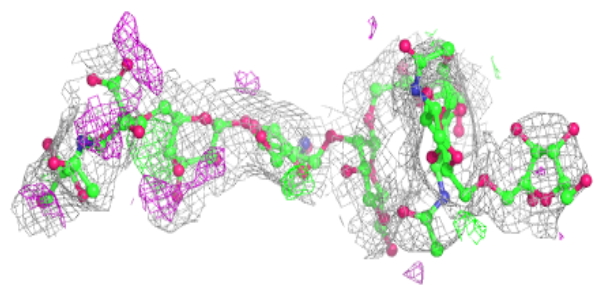
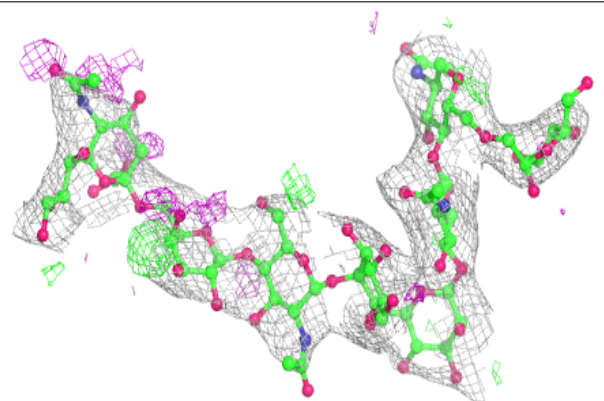


**Electron density around Chain G:**

$2mF_o-DF_c$  (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
8	NAG	C	301	14/15	0.64	0.30	186,195,201,201	0
8	NAG	L	302	14/15	0.66	0.38	167,181,189,190	0
8	NAG	E	301	14/15	0.77	0.22	145,160,163,175	0
7	PO4	B	301	5/5	0.78	0.30	177,180,182,184	0
7	PO4	D	301	5/5	0.82	0.53	154,159,165,175	0
7	PO4	L	301	5/5	0.91	0.21	128,142,144,150	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.