Full wwPDB X-ray Structure Validation Report (i)

Sep 25, 2023 – 10:20 AM EDT

PDB ID	:	5VMF
Title	:	Influenza hemagglutinin H1 mutant DH1D in complex with 6'SLN
Authors	:	Ni, F.; Kondrashkina, E.; Wang, Q.
Deposited on	:	2017-04-27
Resolution	:	2.35 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

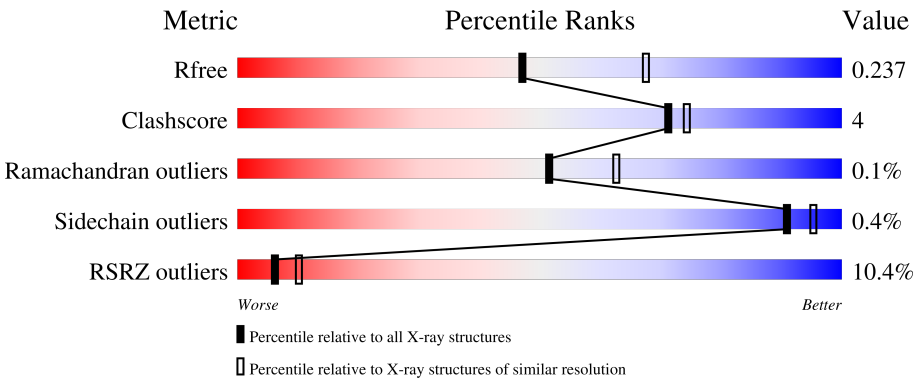
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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 >=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 <=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	326	6% 92% 6% •
1	C	326	3% 93% 6% •
1	E	326	2% 90% 8% •
2	B	191	24% 80% 5% • 14%
2	D	191	20% 81% 5% • 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	191	17% 77% 8% • 14%
3	G	3	33% 67%
4	H	3	33% 67%
4	I	3	67% 33%
4	K	3	67% 33%
5	J	2	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
6	NAG	A	1004	-	-	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 23671 atoms, of which 11123 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 Hemagglutinin HA1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	321	4873	1565	2392	426	479	11	0	0	0
1	C	321	4861	1562	2383	426	479	11	0	0	0
1	E	321	4910	1573	2418	426	482	11	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	222	LEU	GLN	engineered mutation	UNP Q9WFX3
A	224	SER	GLY	engineered mutation	UNP Q9WFX3
C	222	LEU	GLN	engineered mutation	UNP Q9WFX3
C	224	SER	GLY	engineered mutation	UNP Q9WFX3
E	222	LEU	GLN	engineered mutation	UNP Q9WFX3
E	224	SER	GLY	engineered mutation	UNP Q9WFX3

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	164	2552	825	1235	225	261	6	0	0	0
2	D	164	2552	825	1235	225	261	6	0	0	0
2	F	164	2551	825	1234	225	261	6	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

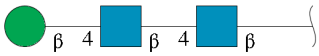
Chain	Residue	Modelled	Actual	Comment	Reference
B	186	GLY	-	expression tag	UNP Q9WFX3
B	187	ALA	-	expression tag	UNP Q9WFX3

Continued on next page...

Continued from previous page...

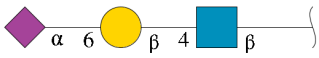
Chain	Residue	Modelled	Actual	Comment	Reference
B	188	LEU	-	expression tag	UNP Q9WFX3
B	189	VAL	-	expression tag	UNP Q9WFX3
B	190	PRO	-	expression tag	UNP Q9WFX3
B	191	ARG	-	expression tag	UNP Q9WFX3
D	186	GLY	-	expression tag	UNP Q9WFX3
D	187	ALA	-	expression tag	UNP Q9WFX3
D	188	LEU	-	expression tag	UNP Q9WFX3
D	189	VAL	-	expression tag	UNP Q9WFX3
D	190	PRO	-	expression tag	UNP Q9WFX3
D	191	ARG	-	expression tag	UNP Q9WFX3
F	186	GLY	-	expression tag	UNP Q9WFX3
F	187	ALA	-	expression tag	UNP Q9WFX3
F	188	LEU	-	expression tag	UNP Q9WFX3
F	189	VAL	-	expression tag	UNP Q9WFX3
F	190	PRO	-	expression tag	UNP Q9WFX3
F	191	ARG	-	expression tag	UNP Q9WFX3

- Molecule 3 is an oligosaccharide called beta-D-mannopyranose-(1-4)-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	H	N	O			
3	G	3	75	22	36	2	15	0	0	0

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

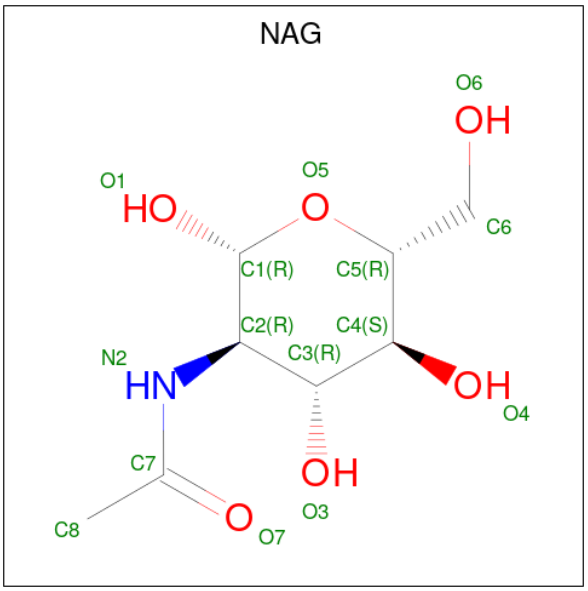


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
4	H	3	87	25	41	2	19	0	0	0
4	I	3	87	25	41	2	19	0	0	0
4	K	3	87	25	41	2	19	0	0	0

- Molecule 5 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	H	N	O			
5	J	2	52	16	25	2	9	0	0	0

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
6	A	1	28	8	14	1	5	0	0
6	C	1	28	8	14	1	5	0	0
6	C	1	28	8	14	1	5	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	220	220	220	0	0
7	B	63	63	63	0	0

Continued on next page...

Continued from previous page...

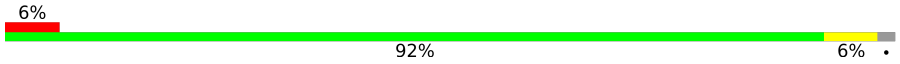
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	C	195	195	195	0	0
			Total	O		
7	D	82	82	82	0	0
			Total	O		
7	E	257	257	257	0	0
			Total	O		
7	F	83	83	83	0	0

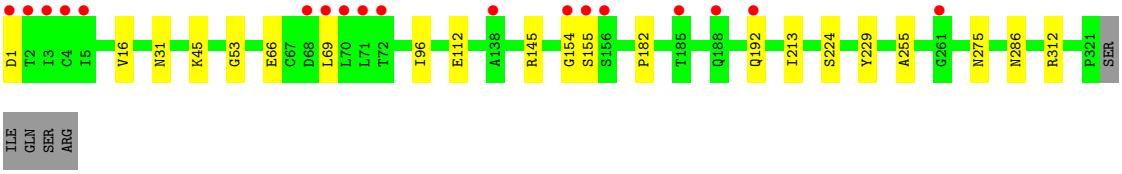
3 Residue-property plots (i)

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: Hemagglutinin HA1

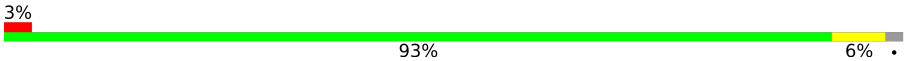
Chain A:

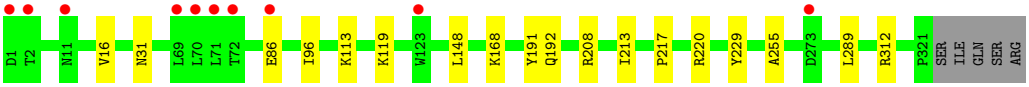




• Molecule 1: Hemagglutinin HA1

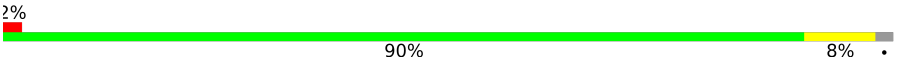
Chain C:





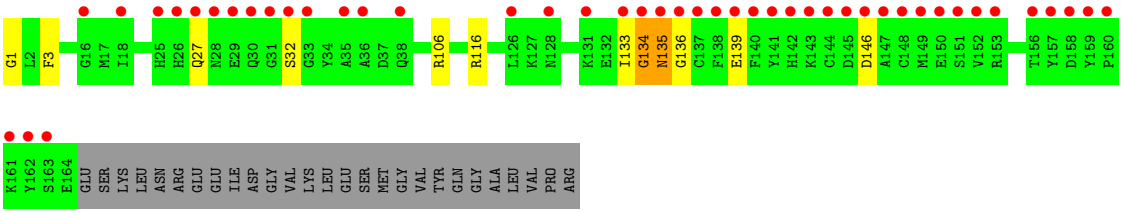
• Molecule 1: Hemagglutinin HA1

Chain E:



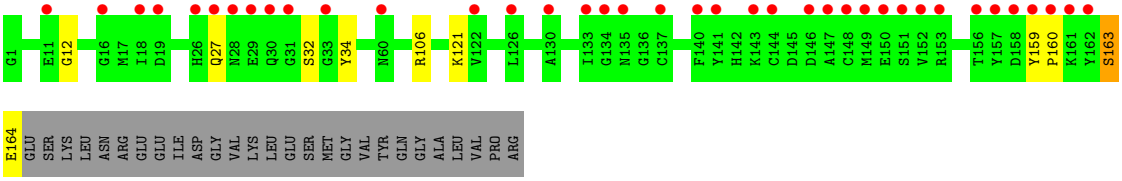
• Molecule 2: Hemagglutinin HA2

Chain B:



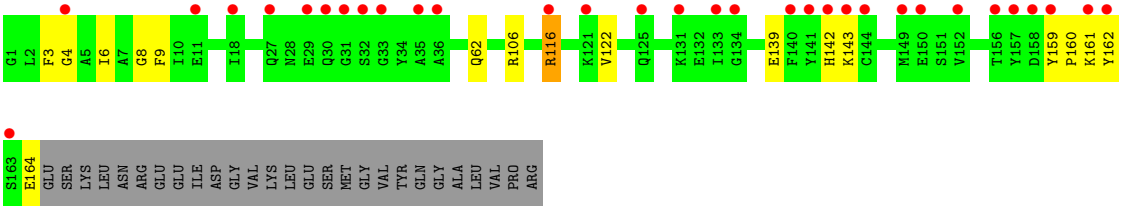
• Molecule 2: Hemagglutinin HA2

Chain D:



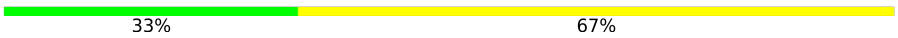
• Molecule 2: Hemagglutinin HA2

Chain F:



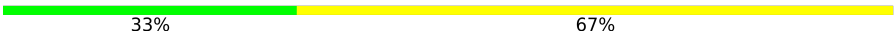
• Molecule 3: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:



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

Chain H:



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

Chain I:

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

Chain K:

• Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain J: 100%

NAG1
NAG2

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	95.75Å 80.26Å 121.23Å 90.00° 91.35° 90.00°	Depositor
Resolution (Å)	44.88 – 2.35 48.36 – 2.35	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.88-2.35) 100.0 (48.36-2.35)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.45 (at 2.34Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.193 , 0.232 0.199 , 0.237	Depositor DCC
R_{free} test set	3854 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	25.3	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 52.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.021 for h,-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	23671	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.46% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: GAL, SIA, BMA, NAG

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	0.27	0/2545	0.51	0/3469
1	C	0.27	0/2542	0.50	0/3465
1	E	0.28	0/2556	0.53	0/3484
2	B	0.29	0/1344	0.46	1/1811 (0.1%)
2	D	0.27	0/1344	0.44	0/1811
2	F	0.29	0/1344	0.56	2/1811 (0.1%)
All	All	0.28	0/11675	0.50	3/15851 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	116	ARG	NE-CZ-NH2	-9.96	115.32	120.30
2	F	116	ARG	NE-CZ-NH1	9.03	124.82	120.30
2	B	134	GLY	N-CA-C	-5.80	98.60	113.10

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	2481	2392	2392	16	0
1	C	2478	2383	2383	14	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2492	2418	2418	23	0
2	B	1317	1235	1234	9	0
2	D	1317	1235	1234	8	0
2	F	1317	1234	1234	13	0
3	G	39	36	34	1	0
4	H	46	41	40	0	0
4	I	46	41	40	0	0
4	K	46	41	40	0	0
5	J	27	25	23	0	0
6	A	14	14	13	1	0
6	C	28	28	26	4	0
7	A	220	0	0	10	0
7	B	63	0	0	4	0
7	C	195	0	0	4	0
7	D	82	0	0	5	0
7	E	257	0	0	13	2
7	F	83	0	0	4	1
All	All	12548	11123	11111	80	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 (80) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:286:ASN:ND2	7:E:1101:HOH:O	1.96	0.96
1:A:1:ASP:N	7:A:1101:HOH:O	1.93	0.89
1:A:69:LEU:O	7:A:1102:HOH:O	1.96	0.82
1:E:109:SER:O	7:E:1102:HOH:O	1.98	0.79
2:D:12:GLY:O	7:D:201:HOH:O	2.02	0.77
2:F:3:PHE:O	2:F:116:ARG:NH2	2.18	0.76
1:A:31:ASN:OD1	7:A:1103:HOH:O	2.05	0.74
1:E:68:ASP:OD1	7:E:1103:HOH:O	2.08	0.72
1:C:31:ASN:OD1	7:C:1101:HOH:O	2.09	0.71
1:A:275:ASN:OD1	7:A:1104:HOH:O	2.07	0.70
1:E:70:LEU:O	7:E:1104:HOH:O	2.09	0.70
1:A:192:GLN:NE2	7:A:1111:HOH:O	2.26	0.68
1:E:192:GLN:OE1	7:E:1106:HOH:O	2.13	0.67
7:A:1105:HOH:O	3:G:3:BMA:O6	2.12	0.66
1:E:72:THR:OG1	7:E:1108:HOH:O	2.14	0.65
1:E:141:SER:O	7:E:1107:HOH:O	2.13	0.65

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:305:TYR:OH	7:E:1105:HOH:O	2.13	0.64
1:A:66:GLU:OE2	7:A:1106:HOH:O	2.16	0.62
2:D:160:PRO:O	2:D:163:SER:OG	2.18	0.61
6:C:1001:NAG:O7	6:C:1001:NAG:O3	2.15	0.61
1:A:53:GLY:O	7:A:1107:HOH:O	2.16	0.59
2:B:1:GLY:N	7:B:203:HOH:O	2.33	0.58
2:B:146:ASP:OD1	7:B:201:HOH:O	2.17	0.58
1:C:86:GLU:HG2	6:C:1001:NAG:H83	1.85	0.58
1:E:321:PRO:O	7:E:1109:HOH:O	2.16	0.57
1:E:167:ASN:OD1	7:E:1110:HOH:O	2.18	0.56
1:E:188:GLN:HG2	7:E:1303:HOH:O	2.06	0.56
2:B:106:ARG:HD3	7:B:232:HOH:O	2.07	0.55
2:B:134:GLY:O	2:B:136:GLY:N	2.40	0.55
2:D:106:ARG:NH2	7:D:207:HOH:O	2.41	0.54
2:B:135:ASN:HB2	7:B:214:HOH:O	2.07	0.54
2:F:6:ILE:O	7:F:201:HOH:O	2.19	0.54
1:E:112:GLU:HB3	7:E:1318:HOH:O	2.08	0.53
2:D:121:LYS:O	7:D:202:HOH:O	2.19	0.52
1:C:96:ILE:HG13	1:C:229:TYR:CE2	2.45	0.51
1:E:110:SER:O	1:E:256:LEU:HA	2.11	0.51
1:A:45:LYS:N	7:A:1104:HOH:O	2.44	0.50
1:A:96:ILE:HG13	1:A:229:TYR:CE1	2.48	0.49
1:E:263:GLY:HA3	7:F:260:HOH:O	2.14	0.48
2:F:62:GLN:NE2	7:F:202:HOH:O	2.22	0.48
1:C:119:LYS:HG3	1:C:148:LEU:HD11	1.96	0.47
1:C:168:LYS:HD3	1:C:255:ALA:HB2	1.96	0.46
2:B:3:PHE:O	2:B:116:ARG:NE	2.39	0.46
2:D:34:TYR:O	7:D:203:HOH:O	2.21	0.46
1:C:213:ILE:O	1:E:208:ARG:NH1	2.46	0.46
1:C:16:VAL:HG12	1:C:312:ARG:HG2	1.97	0.46
2:F:4:GLY:O	2:F:8:GLY:HA3	2.15	0.46
1:C:289:LEU:HD11	6:C:1002:NAG:H82	1.98	0.46
2:B:133:ILE:HD11	2:B:139:GLU:HG3	1.98	0.45
1:A:16:VAL:HG12	1:A:312:ARG:HG2	1.99	0.45
1:E:2:THR:HG22	2:F:139:GLU:HG2	1.98	0.45
1:A:154:GLY:O	1:A:155:SER:OG	2.30	0.45
1:C:113:LYS:NZ	7:C:1110:HOH:O	2.43	0.45
2:D:106:ARG:HD3	7:D:208:HOH:O	2.16	0.44
2:F:4:GLY:C	2:F:116:ARG:NH2	2.70	0.44
1:A:145:ARG:NH2	7:A:1112:HOH:O	2.27	0.44
1:E:156:SER:HA	1:E:192:GLN:OE1	2.17	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:217:PRO:HG3	1:E:238:THR:HB	2.00	0.43
1:A:213:ILE:H	1:C:208:ARG:NH1	2.15	0.43
1:A:182:PRO:HD3	1:A:224:SER:HB3	2.00	0.43
1:A:286:ASN:HD21	6:A:1004:NAG:C6	2.32	0.43
2:F:143:LYS:O	2:F:143:LYS:HG3	2.17	0.43
2:F:159:TYR:HB3	2:F:160:PRO:HD3	1.99	0.43
2:F:3:PHE:O	2:F:116:ARG:NE	2.52	0.42
2:F:106:ARG:HD3	7:F:237:HOH:O	2.19	0.42
2:B:134:GLY:O	2:B:135:ASN:OD1	2.38	0.42
2:D:27:GLN:CD	2:D:32:SER:HB3	2.40	0.42
1:E:5:ILE:HD11	2:F:122:VAL:HG21	2.01	0.42
2:F:9:PHE:HE2	2:F:116:ARG:NH2	2.18	0.42
2:B:27:GLN:NE2	2:B:32:SER:HB3	2.35	0.41
1:E:51:GLN:NE2	7:E:1116:HOH:O	2.33	0.41
2:D:159:TYR:N	2:D:160:PRO:HD2	2.36	0.41
7:C:1159:HOH:O	1:E:208:ARG:HD3	2.21	0.41
1:A:112:GLU:CG	1:A:255:ALA:HB3	2.50	0.41
1:C:191:TYR:O	1:C:192:GLN:HB3	2.21	0.41
2:F:142:HIS:CG	2:F:162:TYR:HB3	2.56	0.41
1:C:208:ARG:NE	7:C:1119:HOH:O	2.53	0.40
1:E:182:PRO:HD3	1:E:224:SER:HB3	2.03	0.40
1:C:220:ARG:NH2	6:C:1001:NAG:H62	2.37	0.40
1:E:96:ILE:HG13	1:E:229:TYR:CE2	2.57	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)
7:E:1333:HOH:O	7:E:1350:HOH:O[2_1059]	1.89	0.31
7:E:1321:HOH:O	7:F:264:HOH:O[2_1049]	1.97	0.23

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	319/326 (98%)	312 (98%)	7 (2%)	0	100	100
1	C	319/326 (98%)	312 (98%)	7 (2%)	0	100	100
1	E	319/326 (98%)	314 (98%)	5 (2%)	0	100	100
2	B	162/191 (85%)	158 (98%)	3 (2%)	1 (1%)	25	27
2	D	162/191 (85%)	159 (98%)	3 (2%)	0	100	100
2	F	162/191 (85%)	159 (98%)	3 (2%)	0	100	100
All	All	1443/1551 (93%)	1414 (98%)	28 (2%)	1 (0%)	51	63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	135	ASN

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	274/283 (97%)	274 (100%)	0	100	100
1	C	273/283 (96%)	273 (100%)	0	100	100
1	E	278/283 (98%)	277 (100%)	1 (0%)	91	95
2	B	139/162 (86%)	139 (100%)	0	100	100
2	D	139/162 (86%)	137 (99%)	2 (1%)	67	78
2	F	139/162 (86%)	137 (99%)	2 (1%)	67	78
All	All	1242/1335 (93%)	1237 (100%)	5 (0%)	91	95

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

Mol	Chain	Res	Type
2	D	163	SER
2	D	164	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	280	THR
2	F	161	LYS
2	F	164	GLU

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

Mol	Chain	Res	Type
1	A	51	GLN
2	B	154	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)

14 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	NAG	G	1	1,3	14,14,15	0.30	0	17,19,21	0.54	0
3	NAG	G	2	3	14,14,15	0.68	1 (7%)	17,19,21	0.38	0
3	BMA	G	3	3	11,11,12	0.74	0	15,15,17	0.85	0
4	NAG	H	1	4	15,15,15	0.40	0	21,21,21	0.52	0
4	GAL	H	2	4	11,11,12	0.87	0	15,15,17	1.11	1 (6%)
4	SIA	H	3	4	20,20,21	1.24	3 (15%)	24,28,31	1.65	4 (16%)
4	NAG	I	1	4	15,15,15	0.18	0	21,21,21	0.19	0
4	GAL	I	2	4	11,11,12	0.38	0	15,15,17	0.98	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	SIA	I	3	4	20,20,21	1.22	3 (15%)	24,28,31	1.65	4 (16%)
5	NAG	J	1	5,1	14,14,15	0.35	0	17,19,21	0.54	0
5	NAG	J	2	5	13,13,15	0.53	0	16,17,21	0.51	0
4	NAG	K	1	4	15,15,15	0.12	0	21,21,21	0.26	0
4	GAL	K	2	4	11,11,12	0.36	0	15,15,17	0.99	0
4	SIA	K	3	4	20,20,21	1.24	3 (15%)	24,28,31	1.64	5 (20%)

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	NAG	G	1	1,3	-	1/6/23/26	0/1/1/1
3	NAG	G	2	3	-	2/6/23/26	0/1/1/1
3	BMA	G	3	3	-	2/2/19/22	0/1/1/1
4	NAG	H	1	4	-	0/6/26/26	0/1/1/1
4	GAL	H	2	4	-	0/2/19/22	0/1/1/1
4	SIA	H	3	4	-	1/18/34/38	0/1/1/1
4	NAG	I	1	4	-	0/6/26/26	0/1/1/1
4	GAL	I	2	4	-	0/2/19/22	0/1/1/1
4	SIA	I	3	4	-	4/18/34/38	0/1/1/1
5	NAG	J	1	5,1	-	2/6/23/26	0/1/1/1
5	NAG	J	2	5	-	2/6/19/26	0/1/1/1
4	NAG	K	1	4	-	1/6/26/26	0/1/1/1
4	GAL	K	2	4	-	1/2/19/22	0/1/1/1
4	SIA	K	3	4	-	2/18/34/38	0/1/1/1

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	3	SIA	C4-C5	-2.56	1.50	1.53
4	I	3	SIA	C4-C5	-2.51	1.51	1.53
3	G	2	NAG	O5-C1	-2.39	1.39	1.43
4	I	3	SIA	C10-N5	2.37	1.42	1.34
4	K	3	SIA	C10-N5	2.35	1.42	1.34
4	H	3	SIA	C5-N5	-2.30	1.42	1.45
4	H	3	SIA	C10-N5	2.30	1.42	1.34
4	K	3	SIA	C5-N5	-2.22	1.42	1.45
4	K	3	SIA	C4-C5	-2.12	1.51	1.53
4	I	3	SIA	C5-N5	-2.04	1.42	1.45

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	K	3	SIA	O6-C2-C3	-4.74	103.94	110.46
4	H	3	SIA	O6-C2-C3	-4.43	104.37	110.46
4	I	3	SIA	O6-C2-C3	-4.16	104.73	110.46
4	I	3	SIA	O6-C2-C1	3.93	115.41	107.70
4	H	3	SIA	O6-C2-C1	3.83	115.21	107.70
4	K	3	SIA	O6-C2-C1	3.75	115.05	107.70
4	H	3	SIA	O1B-C1-C2	2.64	120.55	113.03
4	H	3	SIA	C3-C4-C5	-2.50	108.44	111.46
4	I	3	SIA	O1B-C1-C2	2.36	119.77	113.03
4	K	3	SIA	O1B-C1-C2	2.35	119.74	113.03
4	K	3	SIA	C8-C7-C6	-2.33	108.61	113.03
4	H	2	GAL	O5-C1-C2	-2.24	107.32	110.77
4	I	3	SIA	O1B-C1-O1A	-2.23	119.02	124.09
4	K	3	SIA	C3-C4-C5	-2.23	108.76	111.46

There are no chirality outliers.

All (18) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	K	3	SIA	C7-C8-C9-O9
4	K	3	SIA	O8-C8-C9-O9
5	J	2	NAG	O5-C5-C6-O6
5	J	1	NAG	C4-C5-C6-O6
4	I	3	SIA	O8-C8-C9-O9
3	G	3	BMA	O5-C5-C6-O6
5	J	2	NAG	C4-C5-C6-O6
5	J	1	NAG	O5-C5-C6-O6
4	I	3	SIA	C7-C8-C9-O9
3	G	2	NAG	C4-C5-C6-O6
3	G	3	BMA	C4-C5-C6-O6
3	G	2	NAG	O5-C5-C6-O6
4	K	1	NAG	C4-C5-C6-O6
3	G	1	NAG	C4-C5-C6-O6
4	K	2	GAL	O5-C5-C6-O6
4	H	3	SIA	O1B-C1-C2-O6
4	I	3	SIA	O1A-C1-C2-C3
4	I	3	SIA	O1B-C1-C2-C3

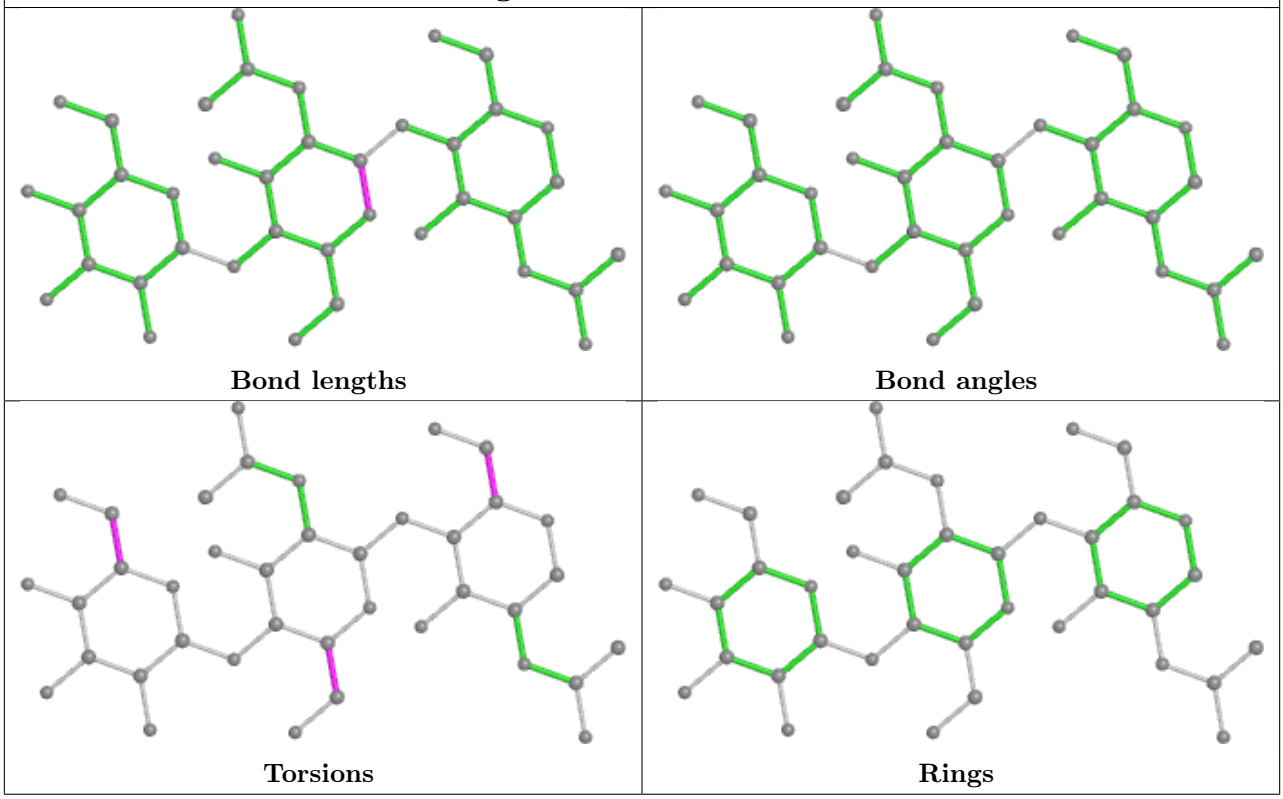
There are no ring outliers.

1 monomer is involved in 1 short contact:

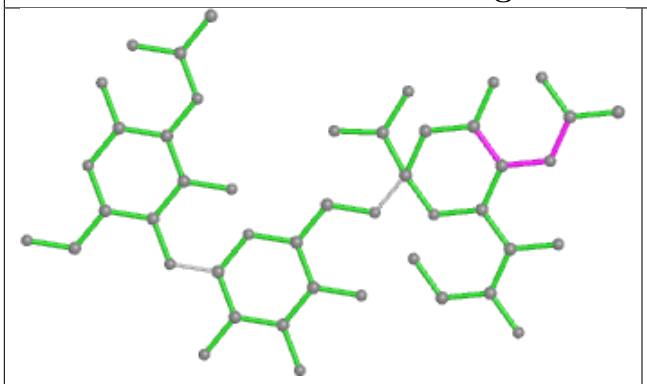
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	3	BMA	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

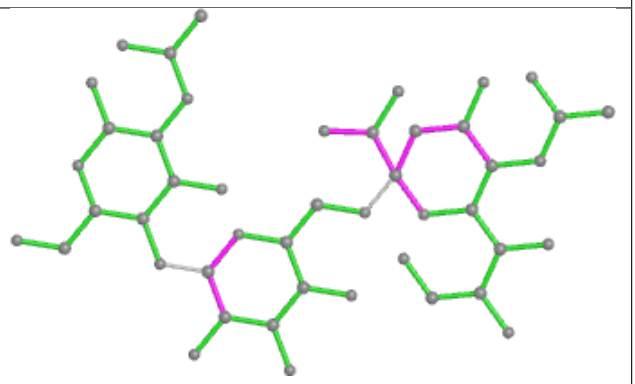
Oligosaccharide Chain G



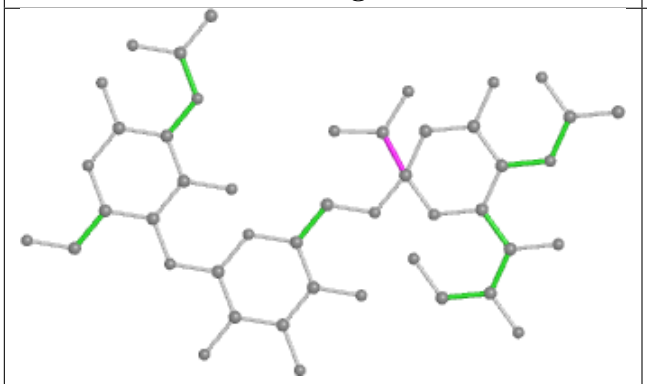
Oligosaccharide Chain H



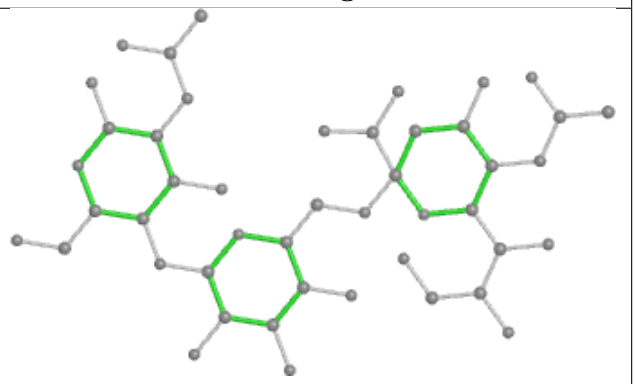
Bond lengths



Bond angles

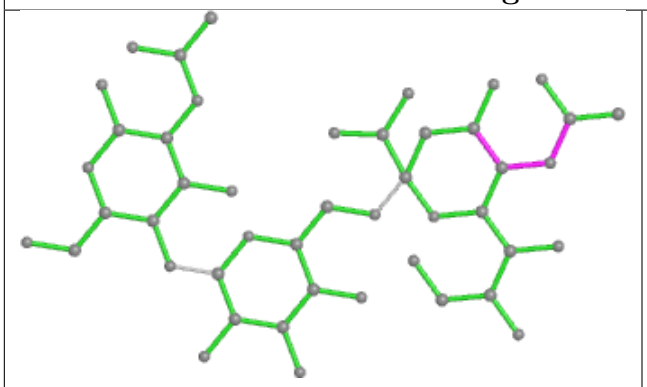


Torsions

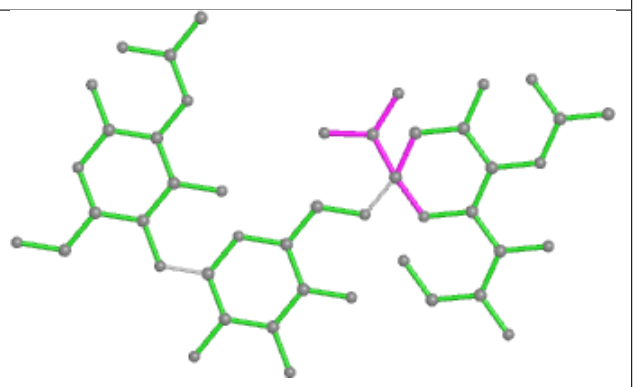


Rings

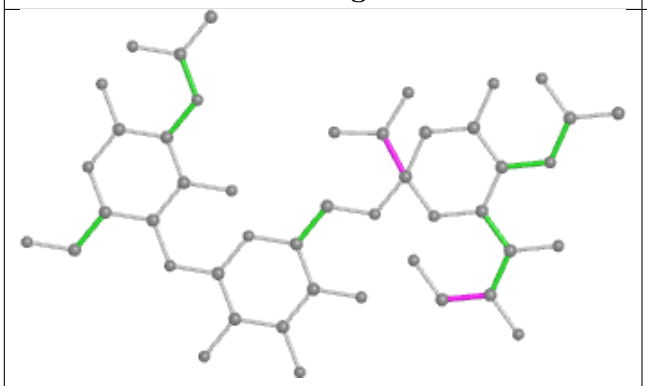
Oligosaccharide Chain I



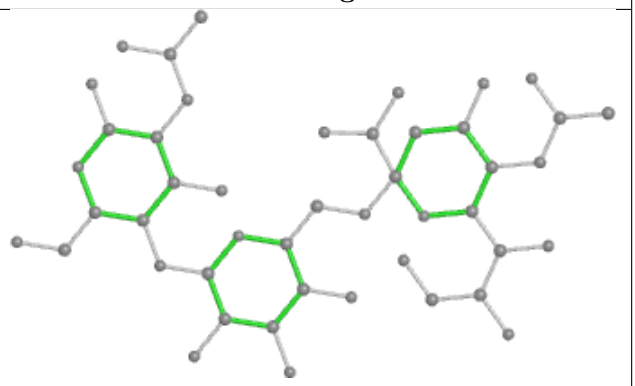
Bond lengths



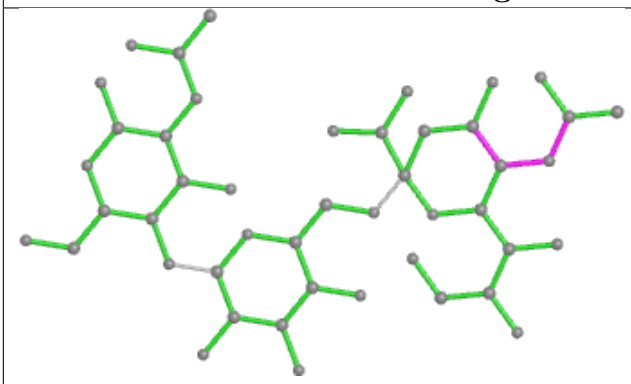
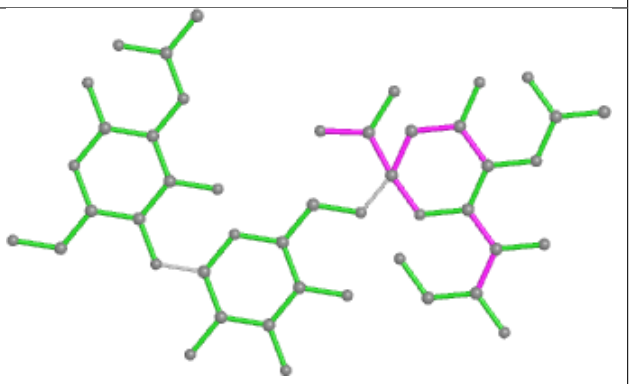
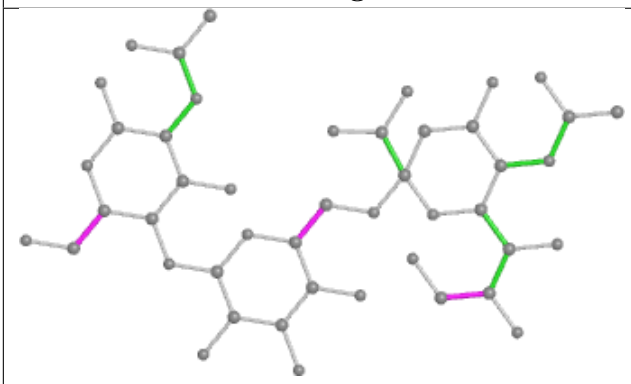
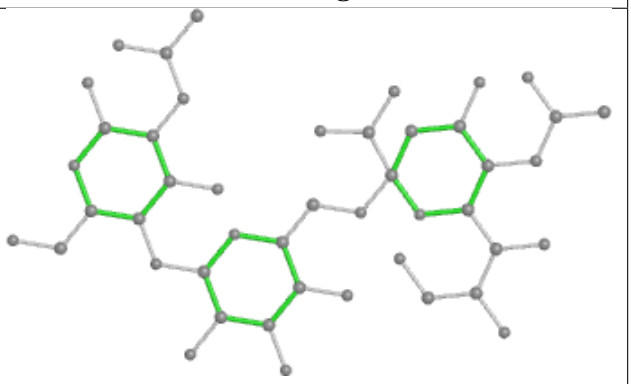
Bond angles



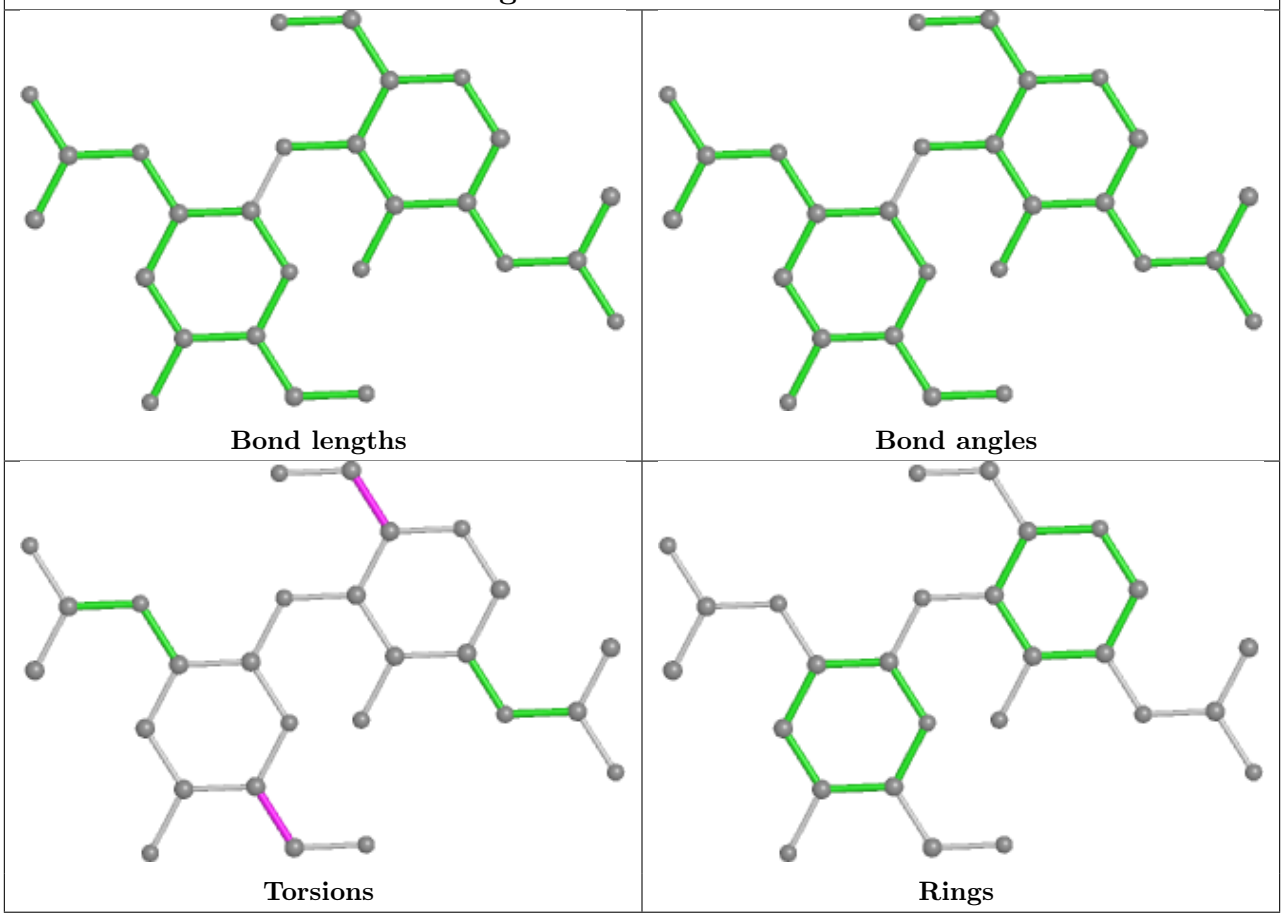
Torsions



Rings

Oligosaccharide Chain K	
 Bond lengths	 Bond angles
 Torsions	 Rings

Oligosaccharide Chain J



5.6 Ligand geometry (i)

3 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	$\#\|Z\| > 2$
6	NAG	C	1002	1	14,14,15	0.61	0	17,19,21	0.82	1 (5%)
6	NAG	C	1001	1	14,14,15	1.05	1 (7%)	17,19,21	0.64	0
6	NAG	A	1004	1	14,14,15	0.72	1 (7%)	17,19,21	0.75	1 (5%)

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
6	NAG	C	1002	1	-	3/6/23/26	0/1/1/1
6	NAG	C	1001	1	-	3/6/23/26	0/1/1/1
6	NAG	A	1004	1	-	3/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	1001	NAG	O5-C1	-3.85	1.37	1.43
6	A	1004	NAG	O5-C1	2.16	1.47	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	C	1002	NAG	C1-O5-C5	3.04	116.31	112.19
6	A	1004	NAG	C1-O5-C5	2.72	115.88	112.19

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	1001	NAG	C1-C2-N2-C7
6	C	1002	NAG	C4-C5-C6-O6
6	C	1002	NAG	O5-C5-C6-O6
6	C	1001	NAG	O5-C5-C6-O6
6	C	1001	NAG	C3-C2-N2-C7
6	A	1004	NAG	C1-C2-N2-C7
6	A	1004	NAG	C4-C5-C6-O6
6	C	1002	NAG	C1-C2-N2-C7
6	A	1004	NAG	O5-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	1002	NAG	1	0
6	C	1001	NAG	3	0
6	A	1004	NAG	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th 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^2)	Q<0.9
1	A	321/326 (98%)	0.30	18 (5%)	24	35	13, 26, 52, 95	0
1	C	321/326 (98%)	0.17	10 (3%)	49	61	18, 29, 52, 88	0
1	E	321/326 (98%)	0.09	7 (2%)	62	72	11, 22, 47, 75	0
2	B	164/191 (85%)	1.41	46 (28%)	0	0	14, 49, 92, 108	0
2	D	164/191 (85%)	1.23	38 (23%)	0	1	15, 54, 98, 107	0
2	F	164/191 (85%)	1.03	32 (19%)	1	2	13, 47, 83, 96	0
All	All	1455/1551 (93%)	0.54	151 (10%)	6	10	11, 30, 83, 108	0

All (151) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	69	LEU	7.9
2	D	147	ALA	7.7
1	A	71	LEU	7.6
2	B	29	GLU	7.0
1	A	70	LEU	6.6
2	B	158	ASP	6.4
2	B	144	CYS	6.3
2	D	149	MET	6.2
2	D	144	CYS	6.2
2	F	134	GLY	6.2
1	C	71	LEU	6.2
2	B	162	TYR	5.9
2	D	33	GLY	5.8
2	B	160	PRO	5.7
2	F	31	GLY	5.7
2	B	157	TYR	5.6
2	B	27	GLN	5.6
2	D	156	THR	5.5
2	D	150	GLU	5.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	69	LEU	5.5
1	A	154	GLY	5.2
2	D	157	TYR	5.1
2	D	158	ASP	5.0
2	B	143	LYS	5.0
2	B	141	TYR	4.8
2	D	31	GLY	4.7
2	B	140	PHE	4.7
2	B	161	LYS	4.6
2	B	131	LYS	4.6
2	B	147	ALA	4.5
1	A	3	ILE	4.4
1	A	1	ASP	4.4
2	B	151	SER	4.4
2	D	152	VAL	4.4
2	B	18	ILE	4.3
2	F	131	LYS	4.3
2	B	38	GLN	4.3
2	B	133	ILE	4.3
2	B	31	GLY	4.2
2	F	158	ASP	4.0
2	D	151	SER	4.0
2	B	33	GLY	3.9
2	B	156	THR	3.9
2	D	27	GLN	3.9
1	E	11	ASN	3.8
2	B	35	ALA	3.8
2	F	33	GLY	3.8
2	D	28	ASN	3.8
2	F	143	LYS	3.8
2	D	146	ASP	3.7
1	C	1	ASP	3.7
2	F	4	GLY	3.7
2	B	148	CYS	3.7
1	E	71	LEU	3.7
2	F	142	HIS	3.7
1	C	70	LEU	3.6
2	B	134	GLY	3.6
2	D	140	PHE	3.6
2	D	160	PRO	3.6
2	D	130	ALA	3.6
2	B	152	VAL	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	142	HIS	3.4
2	B	159	TYR	3.4
2	B	138	PHE	3.4
2	B	146	ASP	3.4
2	F	18	ILE	3.4
1	A	2	THR	3.4
2	D	141	TYR	3.3
2	F	156	THR	3.3
1	A	138	ALA	3.3
2	D	29	GLU	3.3
1	C	2	THR	3.3
1	A	4	CYS	3.3
2	B	153	ARG	3.2
2	B	26	HIS	3.2
2	B	150	GLU	3.2
2	B	136	GLY	3.2
1	A	192	GLN	3.2
2	D	16	GLY	3.2
2	D	26	HIS	3.2
2	D	143	LYS	3.2
2	B	126	LEU	3.1
2	B	25	HIS	3.1
2	B	163	SER	3.1
2	D	153	ARG	3.1
2	F	30	GLN	3.1
2	F	140	PHE	3.1
2	F	116	ARG	3.1
2	B	30	GLN	3.0
2	F	141	TYR	3.0
2	D	133	ILE	3.0
2	D	19	ASP	2.9
2	F	27	GLN	2.9
2	D	126	LEU	2.9
1	A	68	ASP	2.9
2	B	149	MET	2.9
2	F	29	GLU	2.9
2	D	134	GLY	2.8
2	F	149	MET	2.8
1	A	155	SER	2.8
1	A	188	GLN	2.8
2	B	128	ASN	2.8
2	B	137	CYS	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	145	ASP	2.7
1	A	5	ILE	2.7
2	F	32	SER	2.7
2	D	137	CYS	2.6
2	D	162	TYR	2.6
2	B	36	ALA	2.6
2	D	18	ILE	2.6
2	D	30	GLN	2.6
2	D	148	CYS	2.6
1	C	11	ASN	2.5
1	E	2	THR	2.5
2	F	144	CYS	2.5
2	B	135	ASN	2.5
2	D	122	VAL	2.5
2	F	162	TYR	2.5
1	A	72	THR	2.5
1	C	123	TRP	2.4
2	F	161	LYS	2.4
2	D	161	LYS	2.4
2	F	163	SER	2.4
2	D	159	TYR	2.4
2	F	125	GLN	2.4
2	D	60	ASN	2.4
2	D	135	ASN	2.4
2	B	16	GLY	2.3
1	E	70	LEU	2.3
2	F	157	TYR	2.3
1	C	273	ASP	2.3
2	F	152	VAL	2.2
2	F	159	TYR	2.2
1	E	12	SER	2.2
1	A	185	THR	2.2
1	C	86	GLU	2.2
2	F	11	GLU	2.2
1	E	321	PRO	2.2
2	B	32	SER	2.2
1	C	72	THR	2.2
1	E	13	THR	2.2
2	F	133	ILE	2.2
2	B	28	ASN	2.1
1	A	156	SER	2.1
1	A	261	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	150	GLU	2.1
2	F	121	LYS	2.1
2	F	35	ALA	2.0
2	F	36	ALA	2.0
2	D	11	GLU	2.0
2	B	139	GLU	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)

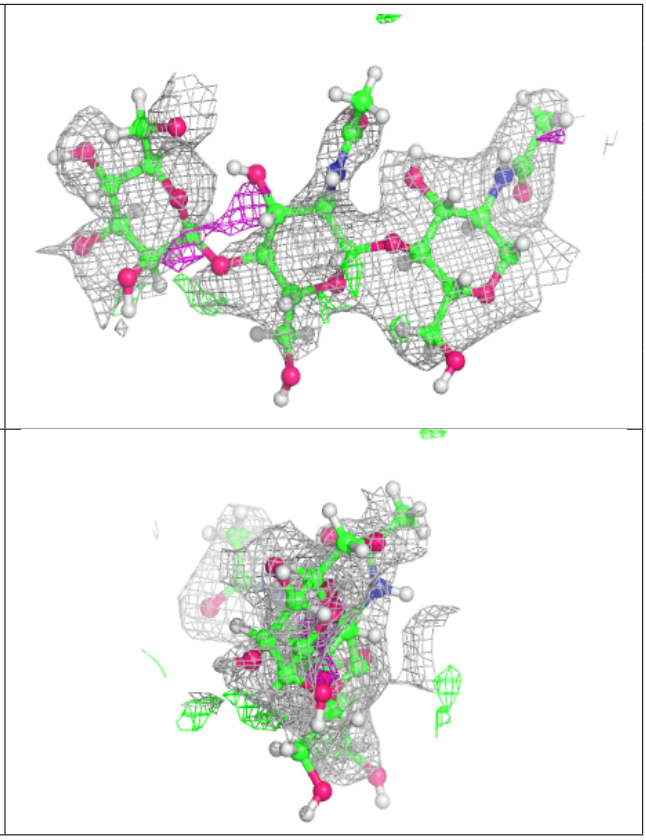
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	BMA	G	3	11/12	0.62	0.31	62,72,82,86	0
4	NAG	H	1	15/15	0.70	0.38	53,64,74,77	0
3	NAG	G	2	14/15	0.78	0.39	58,71,85,86	0
5	NAG	J	2	13/15	0.79	0.33	49,61,73,74	0
4	NAG	I	1	15/15	0.84	0.31	49,59,68,71	0
4	NAG	K	1	15/15	0.86	0.24	38,48,58,70	0
4	GAL	H	2	11/12	0.89	0.18	34,45,60,76	0
4	GAL	K	2	11/12	0.92	0.13	26,37,50,60	0
4	SIA	I	3	20/21	0.93	0.16	29,35,43,52	0
5	NAG	J	1	14/15	0.94	0.14	24,38,51,51	0
3	NAG	G	1	14/15	0.94	0.17	28,38,57,71	0
4	GAL	I	2	11/12	0.95	0.11	32,40,51,61	0
4	SIA	K	3	20/21	0.95	0.12	14,21,28,32	0
4	SIA	H	3	20/21	0.96	0.11	25,31,43,50	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.

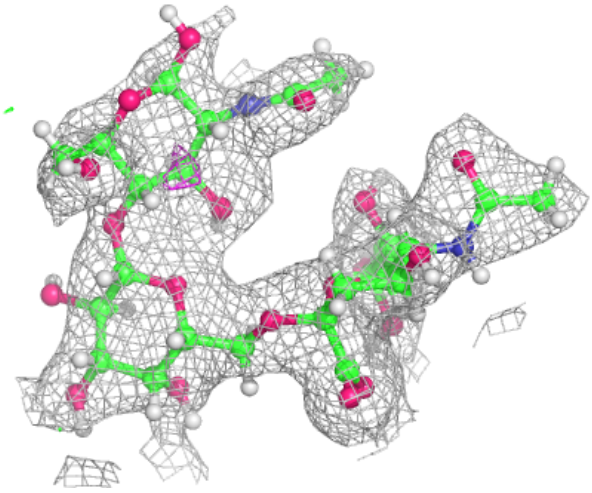
Electron density around Chain G:

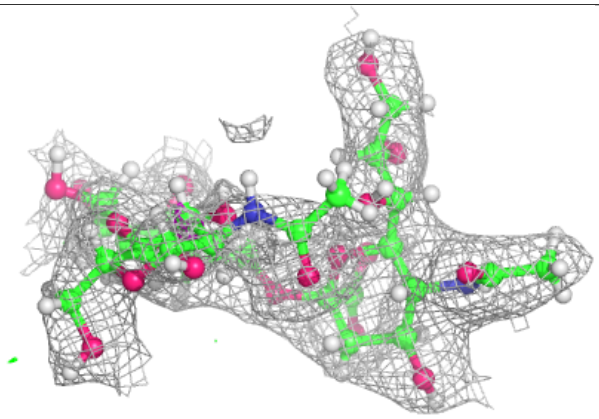
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

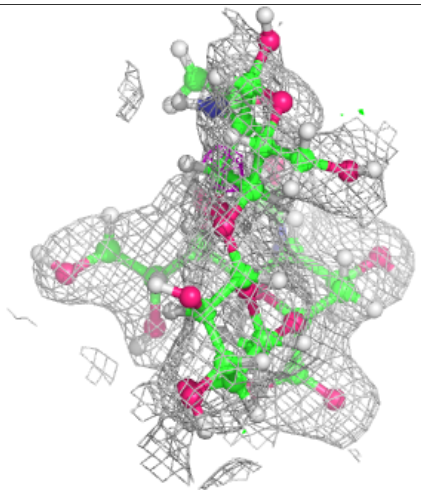


Electron density around Chain H:

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

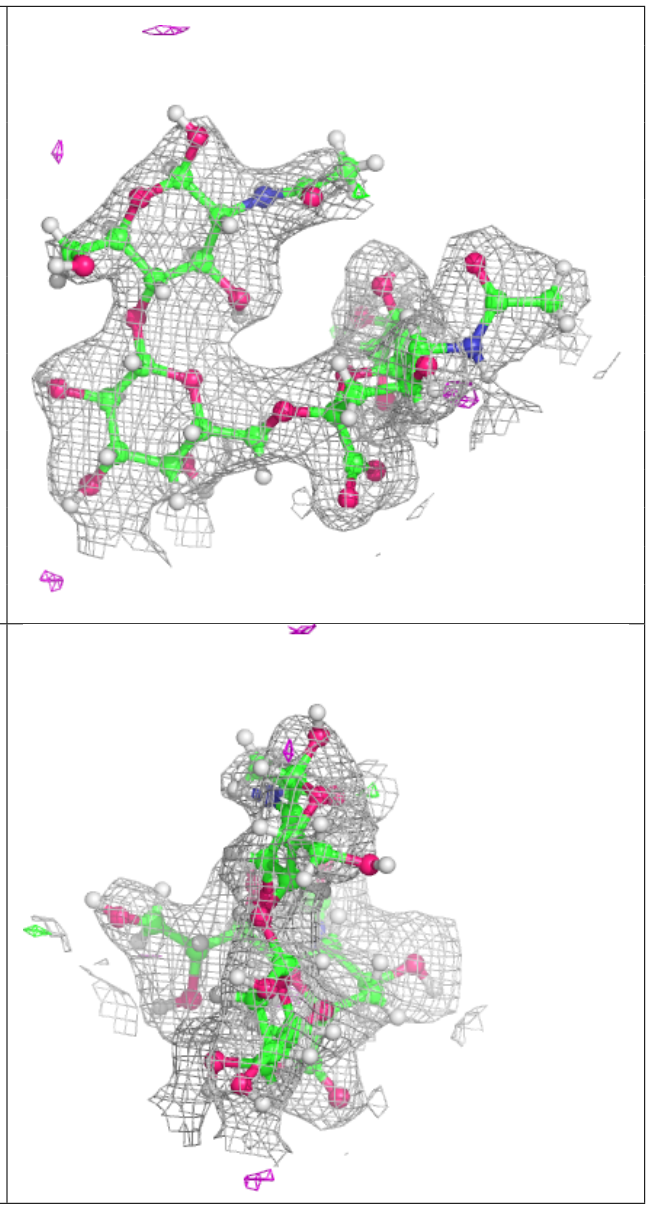






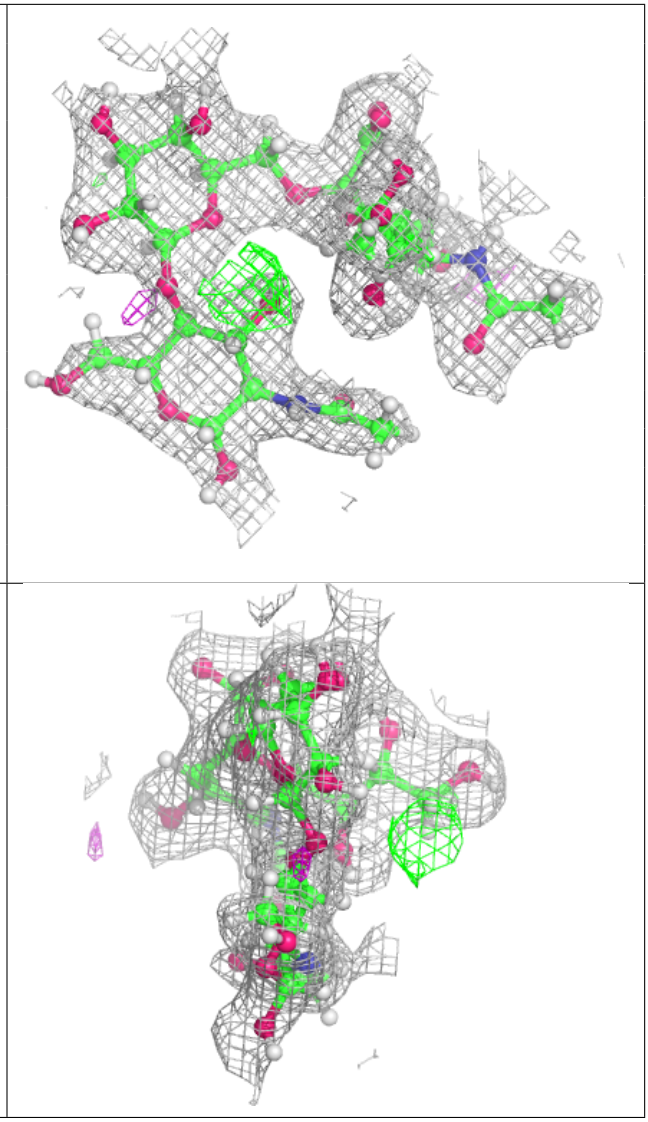
Electron density around Chain I:

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



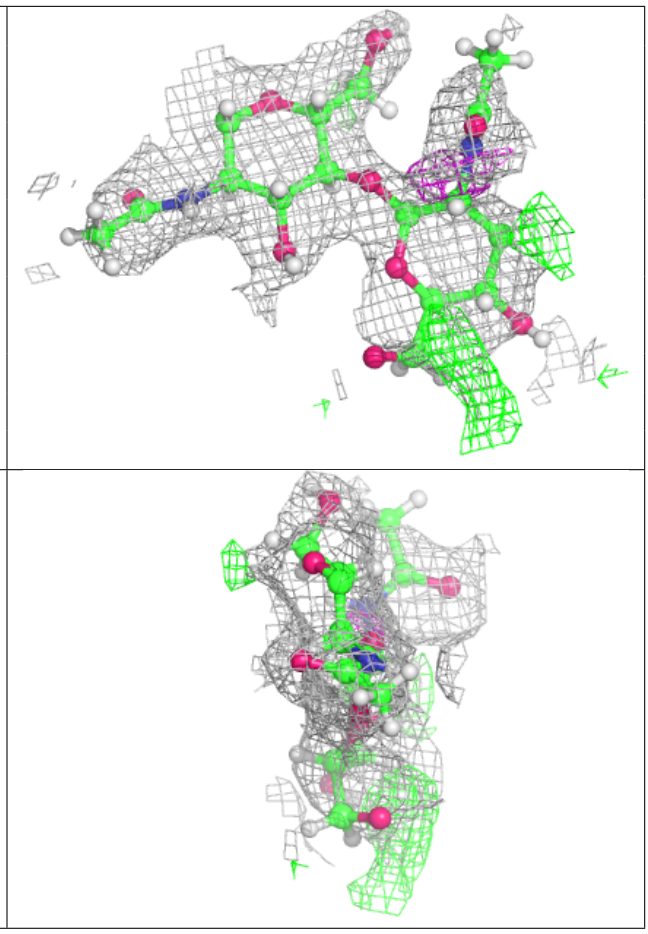
Electron density around Chain K:

$2mF_o$-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 J:

$2mF_o$-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, 95th 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^2)	Q<0.9
6	NAG	A	1004	14/15	0.53	0.58	59,71,85,87	0
6	NAG	C	1002	14/15	0.78	0.28	46,58,70,80	0
6	NAG	C	1001	14/15	0.80	0.34	39,52,65,65	0

6.5 Other polymers (i)

There are no such residues in this entry.