Full wwPDB X-ray Structure Validation Report (i)

Sep 25, 2023 – 11:16 AM EDT

PDB ID	:	5VMJ
Title	:	Influenza hemagglutinin H1 mutant DH1E in complex with 3'SLN
Authors	:	Ni, F.; Kondrashkina, E.; Wang, Q.
Deposited on	:	2017-04-27
Resolution	:	2.95 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:

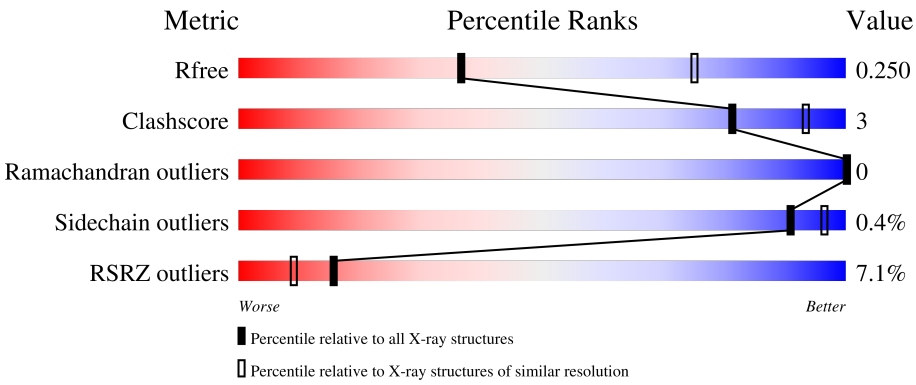
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.95 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	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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	4% 91% 8% •
1	C	326	5% 90% 8% •
1	E	326	% 93% 6% •
2	B	191	20% 80% 5% • 14%
2	D	191	12% 76% 9% • 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	191	6% 83% 14%
3	G	2	50% 50%
3	I	2	100%
3	J	2	100%
4	H	3	33% 33% 33%
5	K	2	50% 50%

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	806	-	-	-	X
6	NAG	C	803	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 22666 atoms, of which 11072 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	4857	1561	2382	426	477	11	0	0	0
1	C	321	4864	1562	2386	426	479	11	0	0	0
1	E	321	4908	1572	2419	426	480	11	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LYS	deletion	UNP Q9WFX4
A	186	GLU	ASP	engineered mutation	UNP Q9WFX4
A	222	LEU	GLN	engineered mutation	UNP Q9WFX4
A	224	SER	GLY	engineered mutation	UNP Q9WFX4
C	?	-	LYS	deletion	UNP Q9WFX4
C	186	GLU	ASP	engineered mutation	UNP Q9WFX4
C	222	LEU	GLN	engineered mutation	UNP Q9WFX4
C	224	SER	GLY	engineered mutation	UNP Q9WFX4
E	?	-	LYS	deletion	UNP Q9WFX4
E	186	GLU	ASP	engineered mutation	UNP Q9WFX4
E	222	LEU	GLN	engineered mutation	UNP Q9WFX4
E	224	SER	GLY	engineered mutation	UNP Q9WFX4

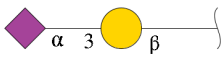
- Molecule 2 is a protein called Hemagglutinin HA2.

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	2551	825	1234	225	261	6	0	0	0
2	F	164	2552	825	1235	225	261	6	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

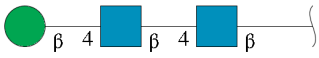
Chain	Residue	Modelled	Actual	Comment	Reference
B	186	GLY	-	expression tag	UNP Q9WFX3
B	187	ALA	-	expression tag	UNP Q9WFX3
B	188	LEU	-	expression tag	UNP Q9WFX3
B	189	VAL	-	expression tag	UNP Q9WFX3
B	190	PRO	-	expression tag	UNP Q9WFX3
B	191	ARG	-	expression tag	UNP Q9WFX3
D	186	GLY	-	expression tag	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 N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galacto pyranose.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
3	G	2	57	17	26	1	13	0	0	0
			Total	C	H	N	O			
3	I	2	57	17	26	1	13	0	0	0
			Total	C	H	N	O			
3	J	2	57	17	26	1	13	0	0	0

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

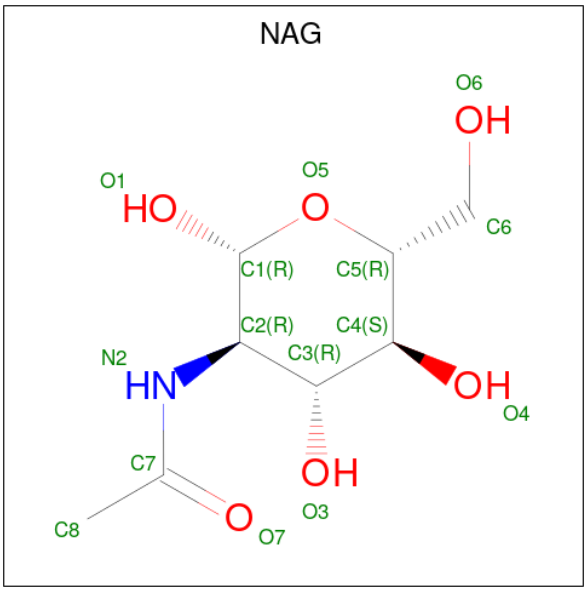


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
4	H	3	75	22	36	2	15	0	0	0

- 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	K	2	52	16	25	2	9	0	0	0

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



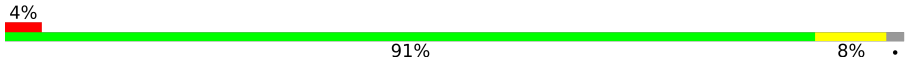
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
6	A	1	28	8	14	1	5	0	0
			Total	C	H	N	O		
6	C	1	28	8	14	1	5	0	0
			Total	C	H	N	O		
6	C	1	28	8	14	1	5	0	0

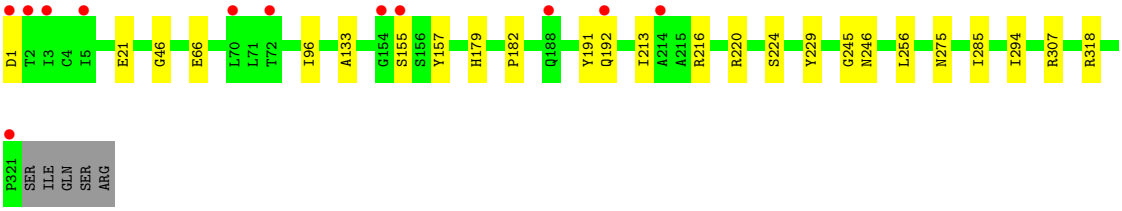
3 Residue-property plots (i)

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

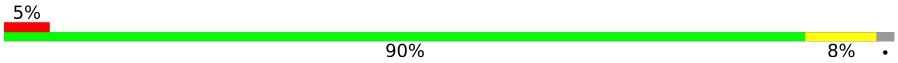
Chain A:

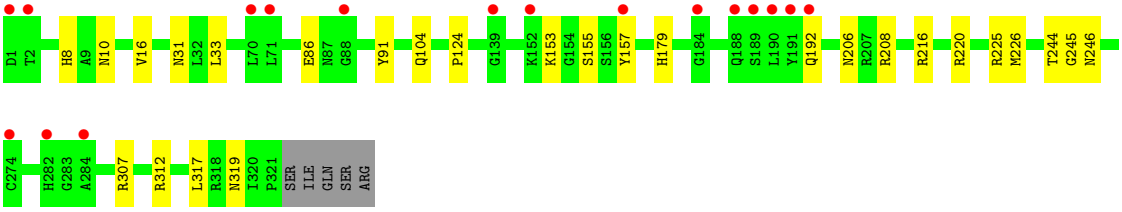




• Molecule 1: Hemagglutinin HA1

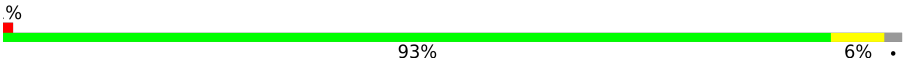
Chain C:

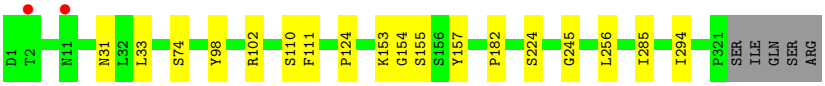




• Molecule 1: Hemagglutinin HA1

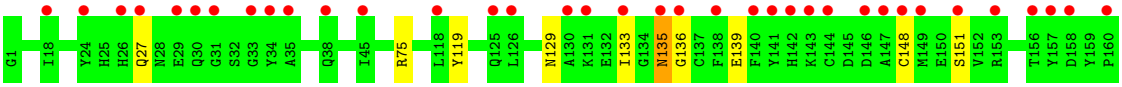
Chain E:

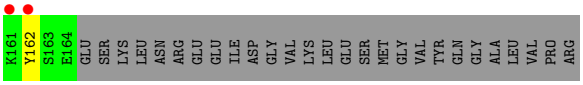




• Molecule 2: Hemagglutinin HA2

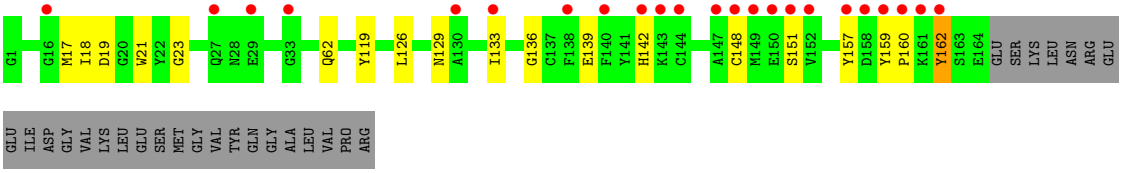
Chain B:





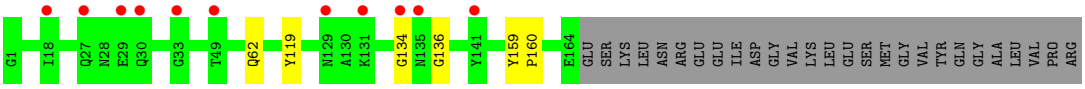
• Molecule 2: Hemagglutinin HA2

Chain D:



• Molecule 2: Hemagglutinin HA2

Chain F:

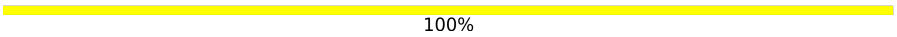


• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose

Chain G:

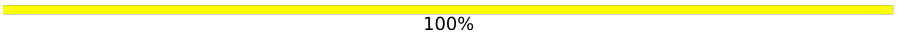
• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose

Chain I:



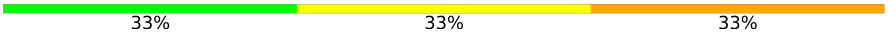
• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose

Chain J:



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

Chain H:



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

Chain K:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	95.94Å 81.05Å 120.95Å 90.00° 91.13° 90.00°	Depositor
Resolution (Å)	54.83 – 2.95 54.82 – 2.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (54.83-2.95) 100.0 (54.82-2.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.07 (at 2.96Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.208 , 0.248 0.213 , 0.250	Depositor DCC
R_{free} test set	2013 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	38.8	Xtriage
Anisotropy	0.194	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 41.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.029 for h,-k,-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	22666	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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: NAG, BMA, GAL, SIA

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	0.24	0/2539	0.46	0/3460
1	C	0.24	0/2542	0.45	0/3464
1	E	0.26	0/2553	0.48	0/3479
2	B	0.25	0/1344	0.40	0/1811
2	D	0.28	0/1344	0.48	1/1811 (0.1%)
2	F	0.25	0/1344	0.44	1/1811 (0.1%)
All	All	0.25	0/11666	0.46	2/15836 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	134	GLY	N-CA-C	-5.67	98.92	113.10
2	D	162	TYR	N-CA-C	5.09	124.75	111.00

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	2475	2382	2384	16	0
1	C	2478	2386	2386	16	0
1	E	2489	2419	2419	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1317	1235	1234	7	0
2	D	1317	1234	1234	12	0
2	F	1317	1235	1234	3	0
3	G	31	26	26	1	0
3	I	31	26	26	0	0
3	J	31	26	26	0	0
4	H	39	36	34	3	0
5	K	27	25	23	0	0
6	A	14	14	13	0	0
6	C	28	28	26	2	0
All	All	11594	11072	11065	58	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 3.

All (58) 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:124:PRO:O	1:E:153:LYS:NZ	1.99	0.95
1:A:213:ILE:O	1:C:208:ARG:NH1	2.15	0.80
2:B:129:ASN:ND2	2:B:162:TYR:O	2.19	0.76
1:A:216:ARG:NH1	1:C:206:ASN:OD1	2.23	0.70
1:A:155:SER:O	1:A:192:GLN:NE2	2.25	0.69
2:D:133:ILE:HD11	2:D:139:GLU:HG3	1.73	0.68
1:A:307:ARG:NH2	2:D:62:GLN:OE1	2.32	0.63
1:A:1:ASP:N	2:B:27:GLN:O	2.30	0.61
1:C:179:HIS:O	1:C:246:ASN:ND2	2.32	0.61
2:B:135:ASN:O	2:B:135:ASN:ND2	2.31	0.60
1:A:179:HIS:ND1	1:A:191:TYR:OH	2.36	0.58
2:B:148:CYS:O	2:B:151:SER:OG	2.13	0.58
2:D:133:ILE:HD11	2:D:139:GLU:CG	2.34	0.57
2:B:133:ILE:HD11	2:B:139:GLU:HG3	1.91	0.53
1:A:21:GLU:OE1	1:A:318:ARG:NH2	2.42	0.53
1:C:10:ASN:OD1	1:C:319:ASN:ND2	2.41	0.53
1:A:220:ARG:NH2	4:H:1:NAG:O7	2.42	0.52
1:E:154:GLY:O	1:E:155:SER:HB3	2.10	0.51
1:C:155:SER:C	1:C:192:GLN:HE21	2.14	0.50
1:C:124:PRO:O	1:C:153:LYS:NZ	2.44	0.50
2:D:119:TYR:CE1	2:D:136:GLY:HA2	2.48	0.49
1:C:8:HIS:N	2:D:21:TRP:O	2.44	0.48
2:B:119:TYR:CE1	2:B:136:GLY:HA2	2.49	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:31:ASN:ND2	1:E:33:LEU:O	2.47	0.48
2:D:142:HIS:CD2	2:D:162:TYR:HB3	2.48	0.47
1:C:31:ASN:ND2	1:C:33:LEU:O	2.48	0.46
1:E:157:TYR:CZ	1:E:245:GLY:HA2	2.50	0.46
2:D:17:MET:SD	2:D:23:GLY:HA3	2.55	0.46
2:D:148:CYS:O	2:D:151:SER:OG	2.21	0.46
1:E:154:GLY:O	1:E:155:SER:CB	2.65	0.45
2:F:159:TYR:HB3	2:F:160:PRO:HD3	1.98	0.45
2:D:129:ASN:ND2	2:D:162:TYR:HB2	2.31	0.45
1:C:86:GLU:HB2	6:C:803:NAG:H82	2.00	0.43
1:E:182:PRO:HD3	1:E:224:SER:HB3	2.00	0.43
2:D:19:ASP:OD1	2:D:19:ASP:N	2.51	0.43
1:C:16:VAL:HG12	1:C:312:ARG:HG2	2.00	0.42
1:A:133:ALA:N	3:G:2:SIA:O1A	2.50	0.42
4:H:2:NAG:O7	4:H:2:NAG:O3	2.31	0.42
1:C:157:TYR:CZ	1:C:245:GLY:HA2	2.54	0.42
1:C:216:ARG:HD2	1:C:225:ARG:HG2	2.01	0.42
1:A:179:HIS:O	1:A:246:ASN:ND2	2.47	0.42
1:C:307:ARG:HH21	2:F:62:GLN:HG3	1.84	0.41
1:E:285:ILE:HG21	1:E:294:ILE:HG13	2.02	0.41
1:A:96:ILE:HG13	1:A:229:TYR:CE1	2.55	0.41
1:E:98:TYR:CZ	1:E:102:ARG:HD2	2.55	0.41
1:A:182:PRO:HD3	1:A:224:SER:HB3	2.03	0.41
2:D:159:TYR:HB3	2:D:160:PRO:HD3	2.01	0.41
1:A:66:GLU:HG3	4:H:1:NAG:HN2	1.85	0.41
1:A:157:TYR:CZ	1:A:245:GLY:HA2	2.55	0.41
2:B:75:ARG:HD3	1:C:104:GLN:HG2	2.01	0.41
1:C:91:TYR:CD1	1:C:226:MET:HB2	2.56	0.41
1:C:220:ARG:HH22	6:C:803:NAG:H5	1.85	0.41
1:A:46:GLY:HA2	1:A:275:ASN:HD21	1.85	0.41
1:E:110:SER:O	1:E:256:LEU:HA	2.21	0.40
2:F:119:TYR:CE1	2:F:136:GLY:HA2	2.56	0.40
1:A:285:ILE:HG21	1:A:294:ILE:HG13	2.03	0.40
2:D:126:LEU:HD23	2:D:157:TYR:CE2	2.57	0.40
1:E:74:SER:O	1:E:111:PHE:N	2.54	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	319/326 (98%)	315 (99%)	4 (1%)	0	100	100
1	C	319/326 (98%)	313 (98%)	6 (2%)	0	100	100
1	E	319/326 (98%)	314 (98%)	5 (2%)	0	100	100
2	B	162/191 (85%)	159 (98%)	3 (2%)	0	100	100
2	D	162/191 (85%)	160 (99%)	2 (1%)	0	100	100
2	F	162/191 (85%)	159 (98%)	3 (2%)	0	100	100
All	All	1443/1551 (93%)	1420 (98%)	23 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	272/282 (96%)	271 (100%)	1 (0%)	91	96
1	C	273/282 (97%)	271 (99%)	2 (1%)	84	93
1	E	277/282 (98%)	277 (100%)	0	100	100
2	B	139/162 (86%)	138 (99%)	1 (1%)	84	93
2	D	139/162 (86%)	138 (99%)	1 (1%)	84	93
2	F	139/162 (86%)	139 (100%)	0	100	100
All	All	1239/1332 (93%)	1234 (100%)	5 (0%)	91	96

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

Mol	Chain	Res	Type
1	A	256	LEU
2	B	135	ASN
1	C	244	THR
1	C	317	LEU
2	D	18	ILE

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	188	GLN
1	C	192	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)

11 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	GAL	G	1	3	11,11,12	1.45	3 (27%)	15,15,17	1.68	3 (20%)
3	SIA	G	2	3	20,20,21	1.25	3 (15%)	24,28,31	1.59	5 (20%)
4	NAG	H	1	4,1	14,14,15	0.35	0	17,19,21	0.58	0
4	NAG	H	2	4	14,14,15	0.80	1 (7%)	17,19,21	0.75	0
4	BMA	H	3	4	11,11,12	0.60	0	15,15,17	0.80	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	GAL	I	1	3	11,11,12	1.13	1 (9%)	15,15,17	1.36	2 (13%)
3	SIA	I	2	3	20,20,21	1.24	3 (15%)	24,28,31	1.65	5 (20%)
3	GAL	J	1	3	11,11,12	1.31	3 (27%)	15,15,17	1.68	4 (26%)
3	SIA	J	2	3	20,20,21	1.25	3 (15%)	24,28,31	1.76	4 (16%)
5	NAG	K	1	5,1	14,14,15	0.19	0	17,19,21	0.59	0
5	NAG	K	2	5	13,13,15	0.66	1 (7%)	16,17,21	0.51	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
3	GAL	G	1	3	-	0/2/19/22	0/1/1/1
3	SIA	G	2	3	-	5/18/34/38	0/1/1/1
4	NAG	H	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	H	2	4	-	4/6/23/26	0/1/1/1
4	BMA	H	3	4	-	0/2/19/22	0/1/1/1
3	GAL	I	1	3	-	0/2/19/22	0/1/1/1
3	SIA	I	2	3	-	2/18/34/38	0/1/1/1
3	GAL	J	1	3	-	0/2/19/22	0/1/1/1
3	SIA	J	2	3	-	4/18/34/38	0/1/1/1
5	NAG	K	1	5,1	-	0/6/23/26	0/1/1/1
5	NAG	K	2	5	-	2/6/19/26	0/1/1/1

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	1	GAL	O5-C1	-3.00	1.38	1.43
4	H	2	NAG	O5-C1	-2.89	1.39	1.43
3	J	2	SIA	C4-C5	-2.68	1.50	1.53
3	G	2	SIA	C4-C5	-2.66	1.50	1.53
3	J	1	GAL	O5-C1	-2.63	1.39	1.43
3	I	1	GAL	O5-C1	-2.62	1.39	1.43
3	I	2	SIA	C4-C5	-2.57	1.50	1.53
3	G	1	GAL	C2-C3	2.50	1.56	1.52
3	G	1	GAL	C4-C3	2.41	1.58	1.52
3	I	2	SIA	C10-N5	2.35	1.42	1.34
3	J	2	SIA	C10-N5	2.31	1.42	1.34
3	G	2	SIA	C10-N5	2.28	1.42	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	J	1	GAL	C2-C3	2.28	1.55	1.52
3	J	1	GAL	C4-C3	2.20	1.57	1.52
5	K	2	NAG	C1-C2	2.12	1.53	1.51
3	J	2	SIA	C5-N5	-2.10	1.42	1.45
3	G	2	SIA	C5-N5	-2.08	1.42	1.45
3	I	2	SIA	C5-N5	-2.03	1.42	1.45

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	J	2	SIA	O6-C2-C1	4.05	115.65	107.70
3	J	1	GAL	C1-C2-C3	4.03	114.62	109.67
3	J	2	SIA	C3-C4-C5	-4.03	106.59	111.46
3	I	2	SIA	O6-C2-C3	-3.93	105.05	110.46
3	I	2	SIA	O6-C2-C1	3.80	115.15	107.70
3	J	2	SIA	O6-C2-C3	-3.79	105.24	110.46
3	G	2	SIA	O6-C2-C1	3.45	114.46	107.70
3	G	1	GAL	C1-C2-C3	3.43	113.88	109.67
3	G	2	SIA	O6-C2-C3	-3.41	105.76	110.46
3	J	1	GAL	C2-C3-C4	3.29	116.59	110.89
3	G	1	GAL	C2-C3-C4	3.25	116.52	110.89
3	G	2	SIA	C3-C4-C5	-3.12	107.69	111.46
3	I	2	SIA	C3-C4-C5	-3.09	107.72	111.46
3	J	2	SIA	O1B-C1-C2	2.72	120.80	113.03
3	G	1	GAL	C3-C4-C5	2.61	114.90	110.24
3	G	2	SIA	O1B-C1-C2	2.52	120.21	113.03
3	G	2	SIA	C11-C10-N5	2.43	120.22	116.10
3	I	2	SIA	O1B-C1-C2	2.41	119.92	113.03
3	I	1	GAL	O5-C1-C2	-2.40	107.06	110.77
3	I	2	SIA	C11-C10-N5	2.27	119.94	116.10
3	I	1	GAL	C3-C4-C5	2.12	114.02	110.24
3	J	1	GAL	O2-C2-C3	-2.11	105.91	110.14
3	J	1	GAL	O3-C3-C4	-2.01	105.69	110.35

There are no chirality outliers.

All (19) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	H	1	NAG	O5-C5-C6-O6
3	G	2	SIA	O8-C8-C9-O9
5	K	2	NAG	C4-C5-C6-O6
4	H	1	NAG	C4-C5-C6-O6

Continued on next page...

Continued from previous page...

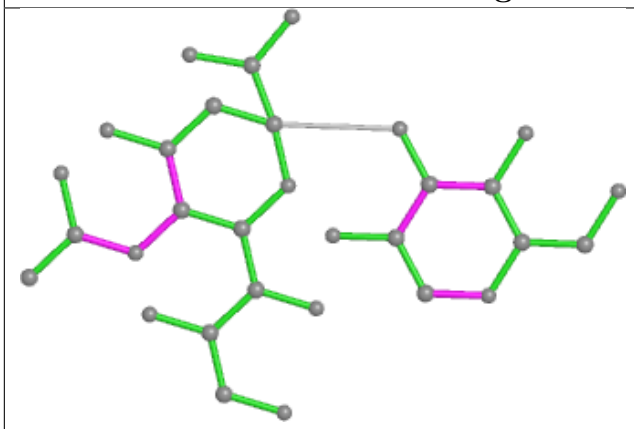
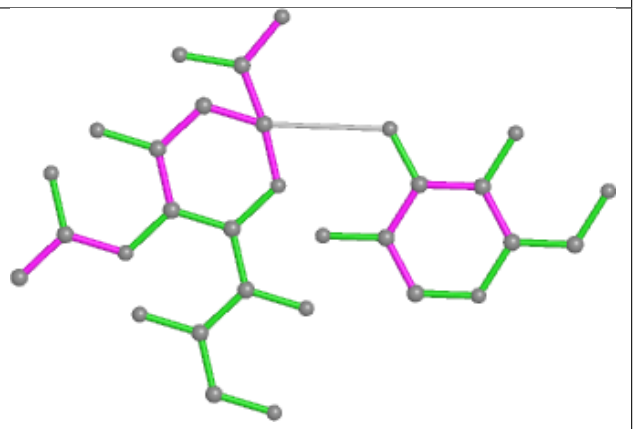
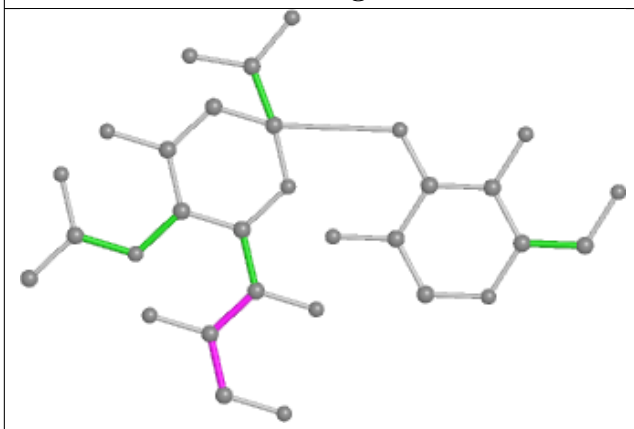
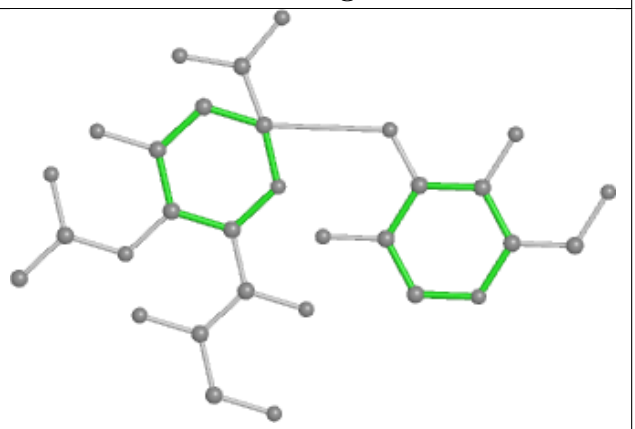
Mol	Chain	Res	Type	Atoms
5	K	2	NAG	O5-C5-C6-O6
3	J	2	SIA	O8-C8-C9-O9
3	G	2	SIA	C7-C8-C9-O9
4	H	2	NAG	C4-C5-C6-O6
3	J	2	SIA	C7-C8-C9-O9
4	H	2	NAG	C1-C2-N2-C7
3	I	2	SIA	O8-C8-C9-O9
4	H	2	NAG	O5-C5-C6-O6
3	G	2	SIA	C6-C7-C8-O8
4	H	2	NAG	C3-C2-N2-C7
3	G	2	SIA	O7-C7-C8-C9
3	J	2	SIA	C6-C7-C8-O8
3	G	2	SIA	O7-C7-C8-O8
3	I	2	SIA	C7-C8-C9-O9
3	J	2	SIA	O7-C7-C8-C9

There are no ring outliers.

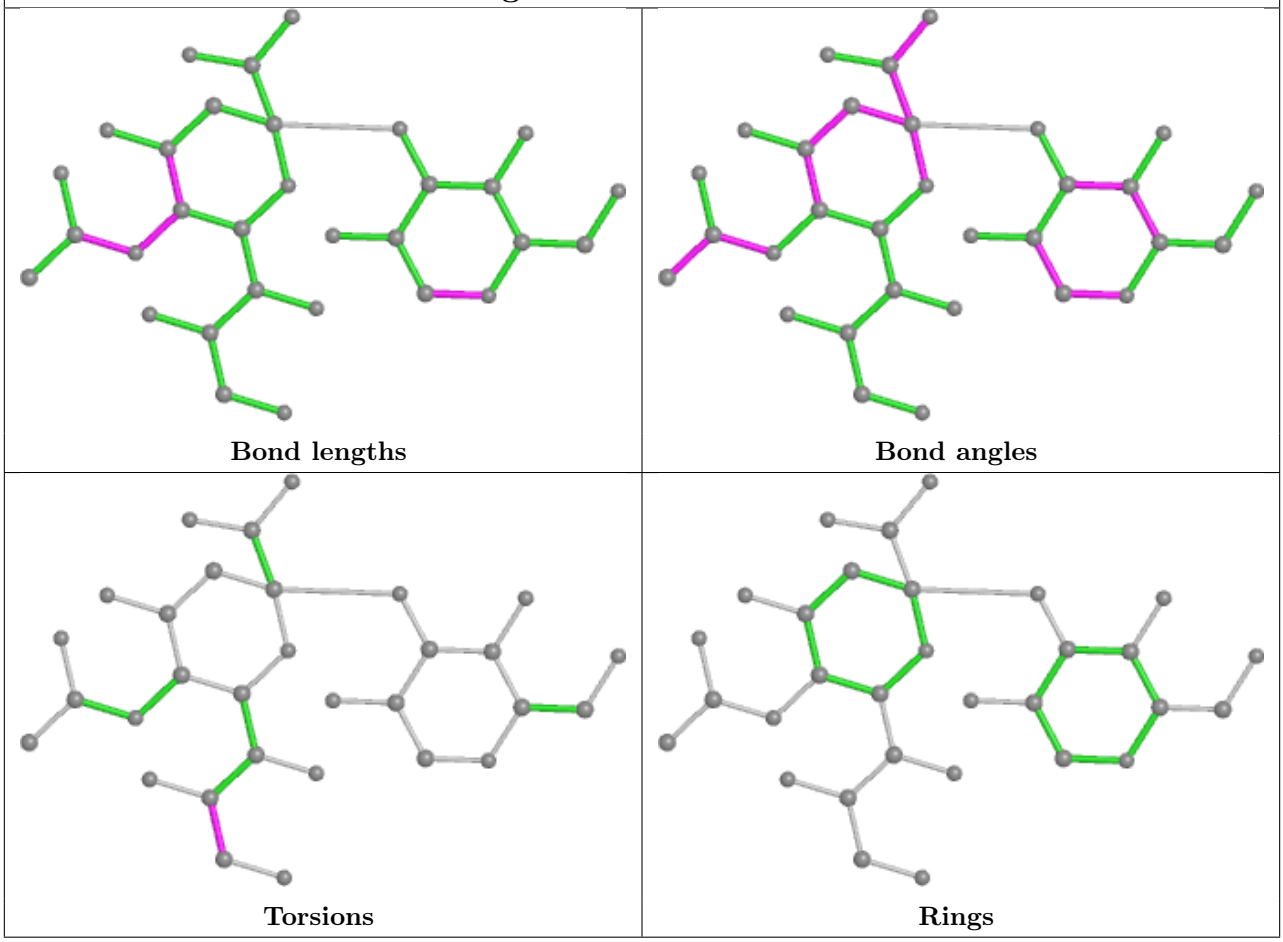
3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	2	NAG	1	0
4	H	1	NAG	2	0
3	G	2	SIA	1	0

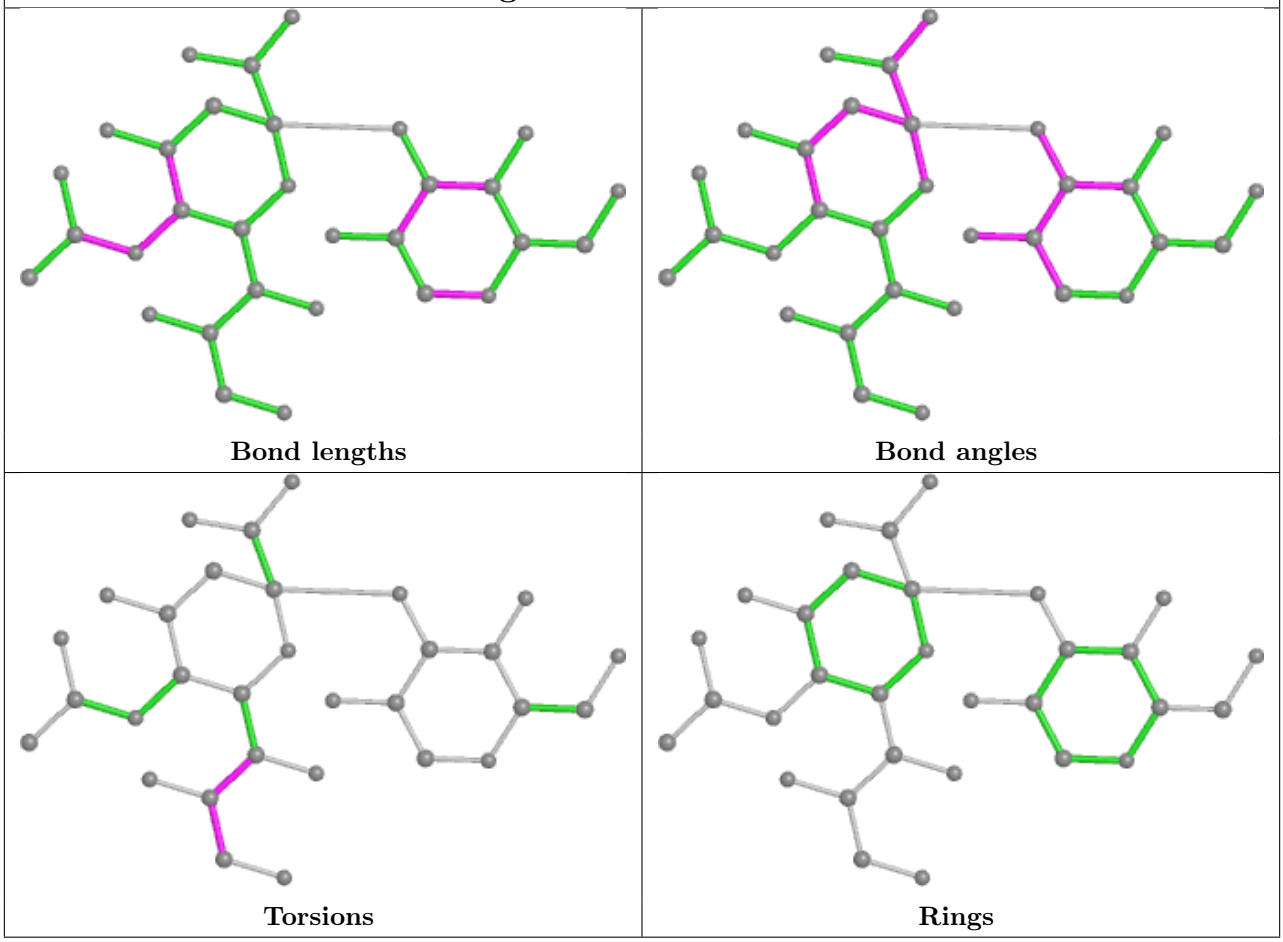
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain G	
 Bond lengths	 Bond angles
 Torsions	 Rings

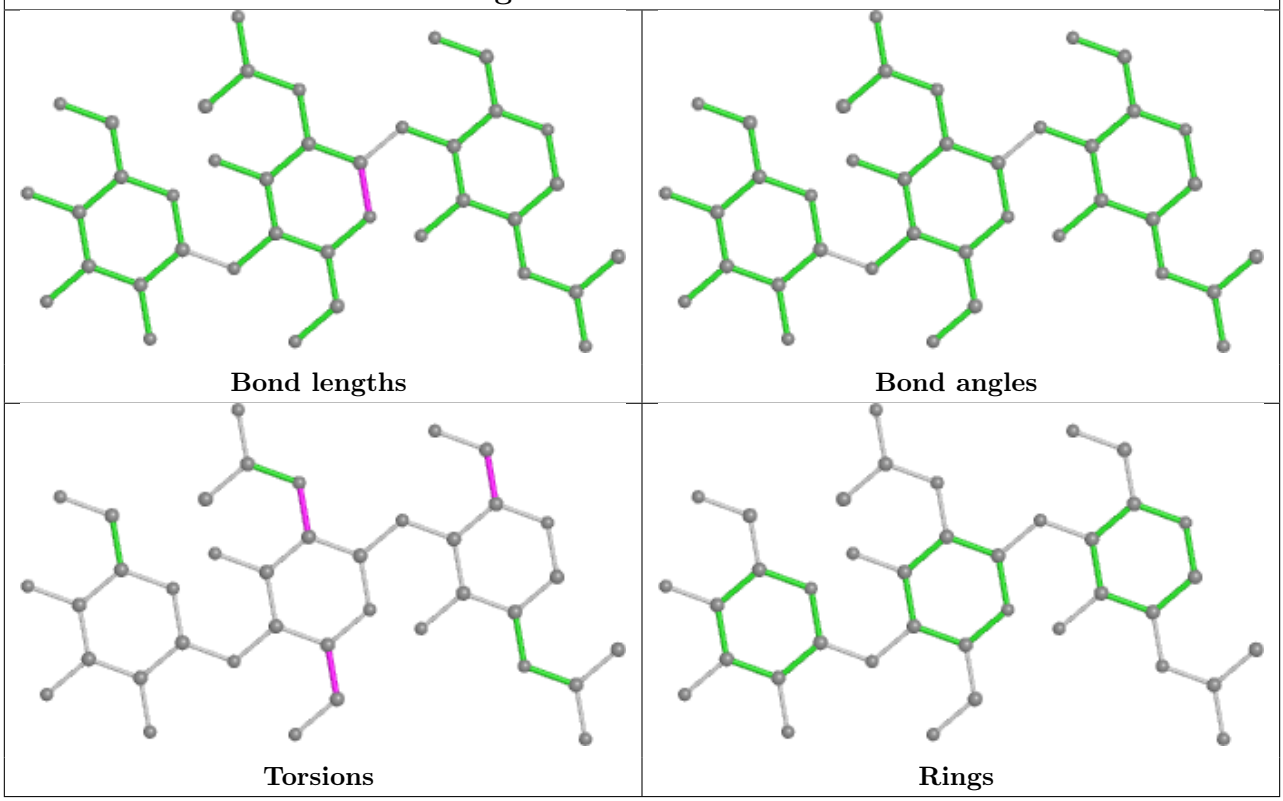
Oligosaccharide Chain I



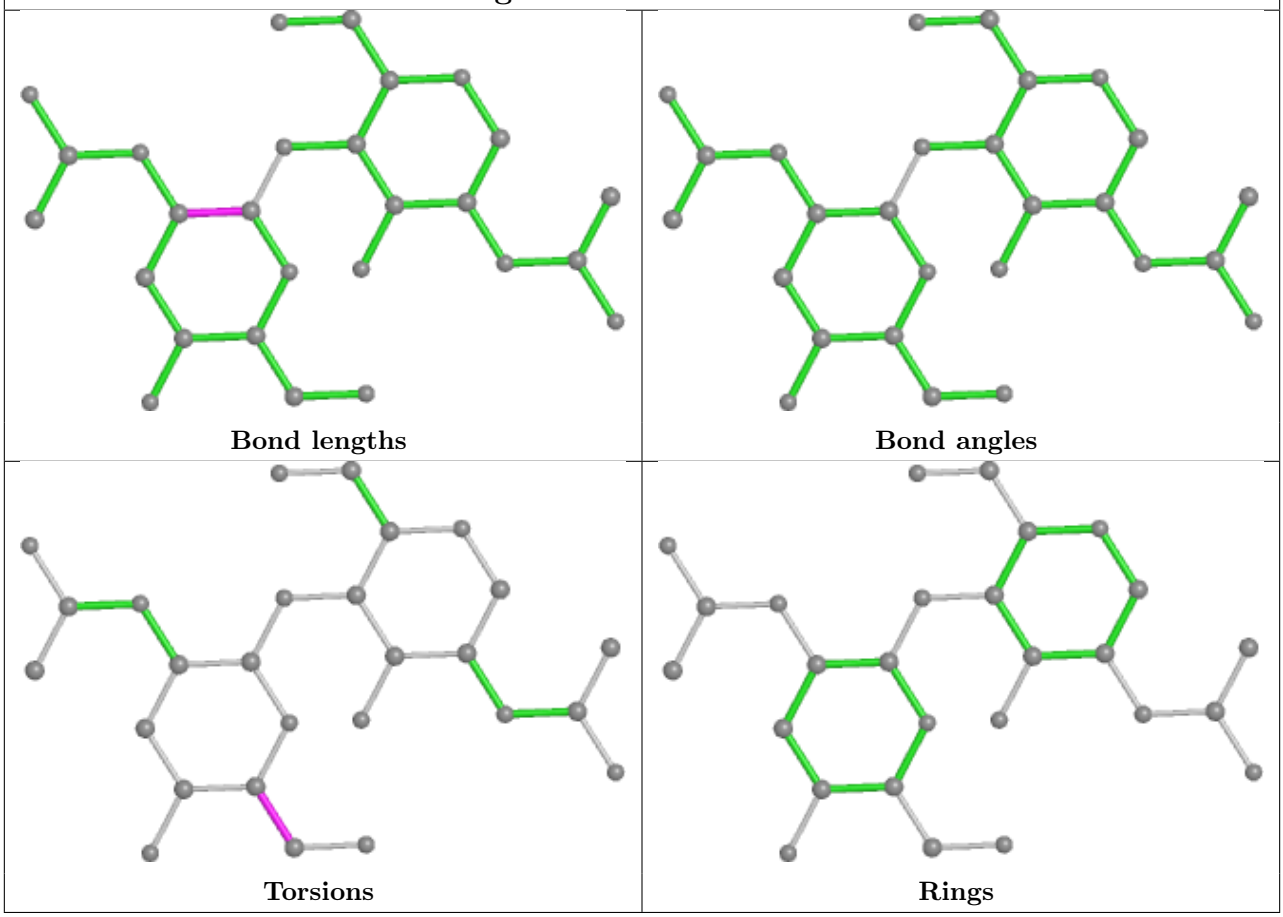
Oligosaccharide Chain J



Oligosaccharide Chain H



Oligosaccharide Chain K



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	804	1	14,14,15	0.48	0	17,19,21	0.71	1 (5%)
6	NAG	C	803	1	14,14,15	0.48	0	17,19,21	0.42	0
6	NAG	A	806	1	14,14,15	0.73	1 (7%)	17,19,21	0.88	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	804	1	-	2/6/23/26	0/1/1/1
6	NAG	C	803	1	-	4/6/23/26	0/1/1/1
6	NAG	A	806	1	-	3/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	806	NAG	O5-C1	2.37	1.47	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	A	806	NAG	C1-O5-C5	3.27	116.63	112.19
6	C	804	NAG	C1-O5-C5	2.59	115.70	112.19

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	804	NAG	O5-C5-C6-O6
6	C	804	NAG	C4-C5-C6-O6
6	C	803	NAG	O5-C5-C6-O6
6	C	803	NAG	C4-C5-C6-O6
6	A	806	NAG	C4-C5-C6-O6
6	C	803	NAG	C1-C2-N2-C7
6	A	806	NAG	O5-C5-C6-O6
6	C	803	NAG	C3-C2-N2-C7
6	A	806	NAG	C1-C2-N2-C7

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	803	NAG	2	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%)	0.24	12 (3%) 41 27	12, 28, 52, 97	0
1	C	321/326 (98%)	0.57	17 (5%) 26 16	19, 38, 63, 89	0
1	E	321/326 (98%)	0.10	2 (0%) 89 78	12, 22, 49, 69	0
2	B	164/191 (85%)	1.06	38 (23%) 0 0	12, 46, 90, 105	0
2	D	164/191 (85%)	0.78	23 (14%) 2 1	13, 55, 93, 102	0
2	F	164/191 (85%)	0.57	11 (6%) 17 10	10, 48, 77, 97	0
All	All	1455/1551 (93%)	0.47	103 (7%) 16 9	10, 34, 81, 105	0

All (103) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	151	SER	5.5
2	B	35	ALA	5.2
2	B	143	LYS	4.9
2	B	140	PHE	4.7
1	A	70	LEU	4.6
2	B	160	PRO	4.5
1	C	188	GLN	4.5
2	B	29	GLU	4.4
1	C	2	THR	4.3
2	D	144	CYS	4.2
2	D	138	PHE	3.9
2	F	33	GLY	3.8
1	C	70	LEU	3.8
2	B	27	GLN	3.8
2	D	157	TYR	3.7
2	D	149	MET	3.6
2	B	162	TYR	3.6
2	B	144	CYS	3.6
2	D	143	LYS	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	146	ASP	3.4
2	B	161	LYS	3.4
2	B	131	LYS	3.3
2	B	156	THR	3.3
2	D	158	ASP	3.2
1	A	72	THR	3.2
2	B	24	TYR	3.1
2	F	134	GLY	3.1
2	B	138	PHE	3.1
1	C	184	GLY	3.1
2	B	133	ILE	3.0
2	B	153	ARG	2.9
1	C	190	LEU	2.9
2	B	147	ALA	2.9
2	B	18	ILE	2.9
2	B	126	LEU	2.9
2	D	147	ALA	2.9
2	D	29	GLU	2.9
2	D	160	PRO	2.8
2	D	33	GLY	2.8
1	A	1	ASP	2.8
1	C	192	GLN	2.8
2	B	158	ASP	2.8
2	B	141	TYR	2.8
2	D	140	PHE	2.7
1	C	189	SER	2.7
2	D	159	TYR	2.7
1	A	3	ILE	2.7
1	A	192	GLN	2.6
2	B	157	TYR	2.6
2	F	30	GLN	2.6
2	D	27	GLN	2.6
1	C	1	ASP	2.6
1	C	157	TYR	2.6
1	C	191	TYR	2.6
1	E	11	ASN	2.6
1	C	282	HIS	2.6
2	B	38	GLN	2.5
2	B	135	ASN	2.5
2	D	130	ALA	2.5
2	B	34	TYR	2.5
2	B	125	GLN	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	129	ASN	2.4
2	D	142	HIS	2.4
2	D	162	TYR	2.4
2	D	148	CYS	2.4
1	A	5	ILE	2.4
2	B	45	ILE	2.4
2	F	29	GLU	2.4
2	B	33	GLY	2.4
2	B	149	MET	2.3
2	B	130	ALA	2.3
1	A	2	THR	2.3
2	D	133	ILE	2.3
1	C	139	GLY	2.3
2	B	148	CYS	2.3
2	B	136	GLY	2.3
2	F	27	GLN	2.3
1	A	154	GLY	2.3
1	C	71	LEU	2.3
2	D	161	LYS	2.3
1	C	284	ALA	2.2
1	A	188	GLN	2.2
2	B	30	GLN	2.2
2	B	31	GLY	2.2
2	D	150	GLU	2.2
2	B	142	HIS	2.2
2	F	18	ILE	2.2
2	D	16	GLY	2.1
2	D	152	VAL	2.1
1	C	88	GLY	2.1
2	B	118	LEU	2.1
2	D	151	SER	2.1
2	F	141	TYR	2.1
1	C	152	LYS	2.0
1	A	214	ALA	2.0
1	E	2	THR	2.0
1	C	274	CYS	2.0
1	A	155	SER	2.0
2	F	131	LYS	2.0
1	A	321	PRO	2.0
2	F	135	ASN	2.0
2	B	26	HIS	2.0
2	F	49	THR	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)

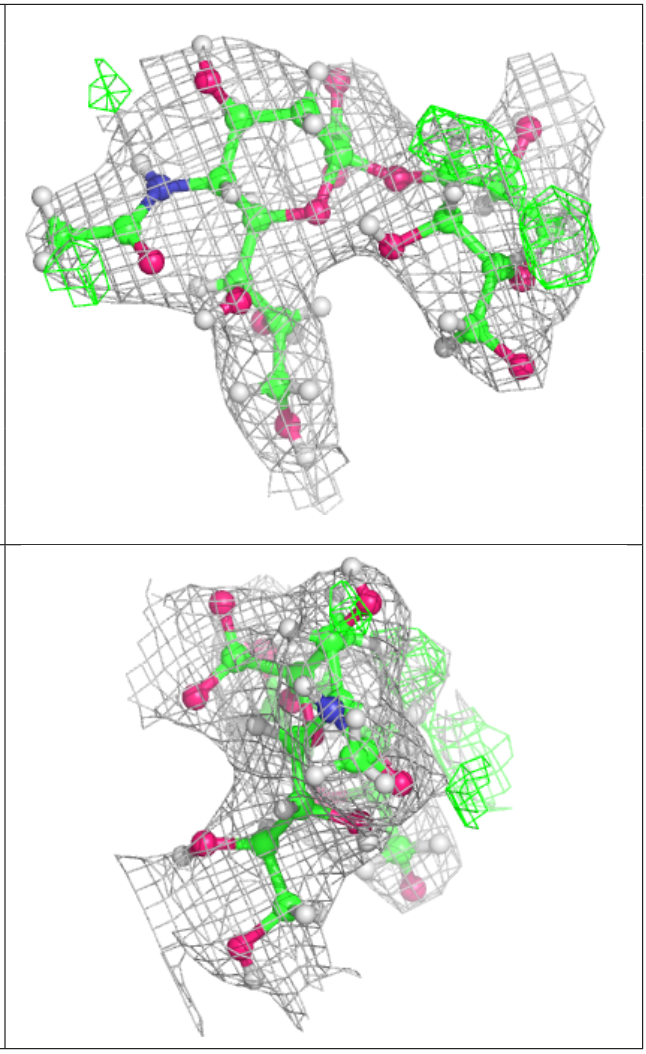
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	BMA	H	3	11/12	0.63	0.28	57,70,84,92	0
4	NAG	H	2	14/15	0.78	0.32	56,74,88,90	0
5	NAG	K	2	13/15	0.79	0.29	27,42,54,74	0
4	NAG	H	1	14/15	0.86	0.35	36,52,70,77	0
3	GAL	G	1	11/12	0.92	0.15	37,49,60,61	0
3	GAL	J	1	11/12	0.92	0.19	28,37,44,48	0
3	GAL	I	1	11/12	0.93	0.21	43,55,66,70	0
5	NAG	K	1	14/15	0.94	0.18	22,28,36,66	0
3	SIA	I	2	20/21	0.94	0.24	33,46,55,56	0
3	SIA	J	2	20/21	0.96	0.17	13,18,24,28	0
3	SIA	G	2	20/21	0.96	0.20	24,33,43,48	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.

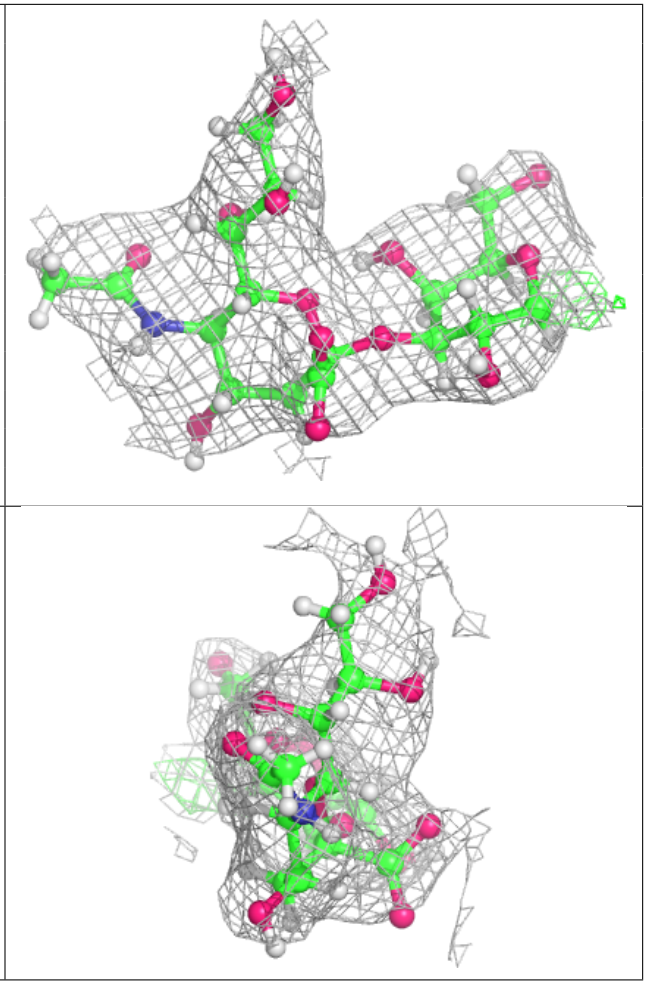
Electron density around Chain G:

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



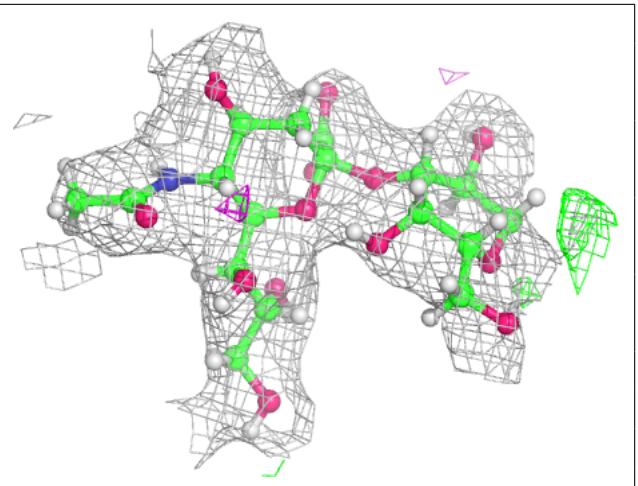
Electron density around Chain I:

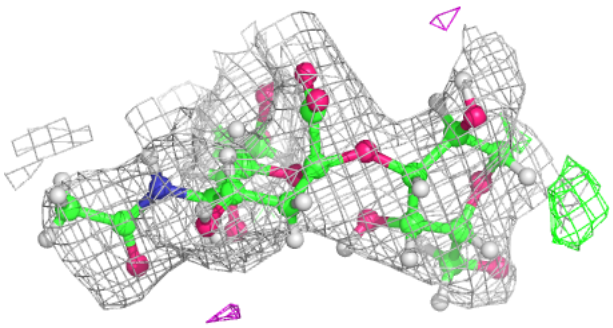
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

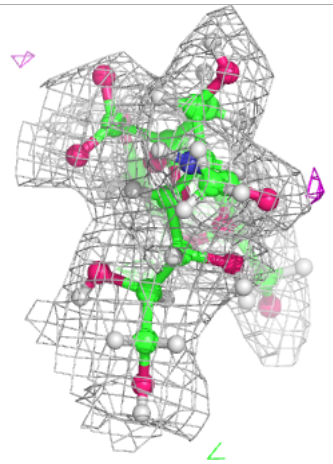


Electron density around Chain J:

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

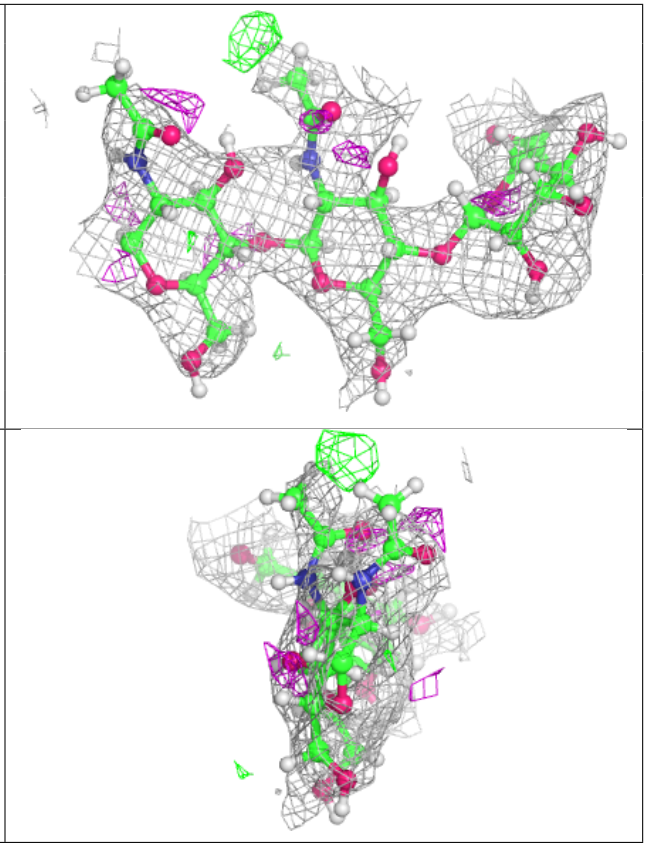






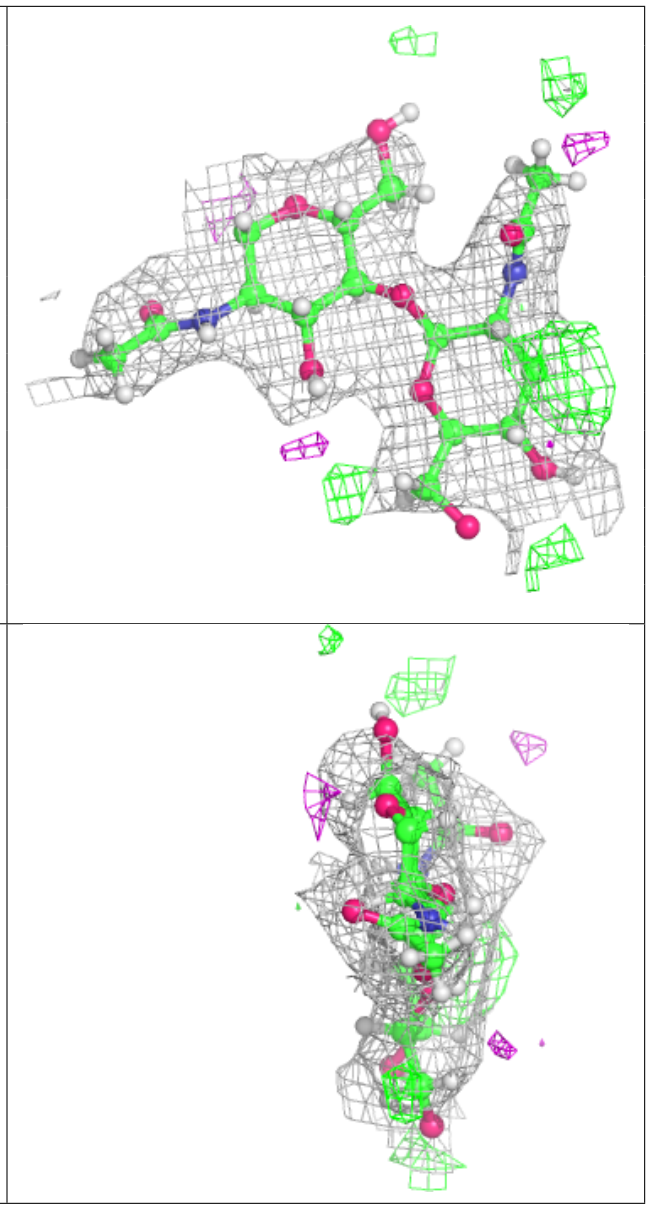
Electron density around Chain H:

$2mF_o$-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 K:

$2mF_o$-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	806	14/15	0.51	0.45	57,78,98,107	0
6	NAG	C	803	14/15	0.69	0.44	54,70,82,86	0
6	NAG	C	804	14/15	0.78	0.29	45,61,74,79	0

6.5 Other polymers (i)

There are no such residues in this entry.