



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2020 – 03:23 PM BST

PDB ID : 5VMC  
Title : Influenza hemagglutinin H1 mutant DH1 in complex with 6'SLN  
Authors : Ni, F.; Kondrashkina, E.; Wang, Q.  
Deposited on : 2017-04-27  
Resolution : 2.15 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

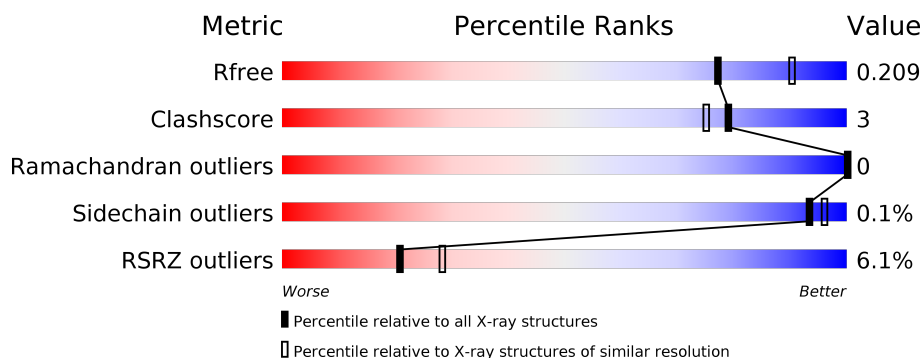
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	326	<div> <div>3%</div> <div> <div></div> <div>93%</div> <div>6%</div> <div></div> </div> </div>
1	C	326	<div> <div>3%</div> <div> <div></div> <div>93%</div> <div>6%</div> <div></div> </div> </div>
1	E	326	<div> <div>2%</div> <div> <div></div> <div>92%</div> <div>6%</div> <div></div> </div> </div>
2	B	191	<div> <div>15%</div> <div> <div></div> <div>84%</div> <div>14%</div> <div></div> </div> </div>
2	D	191	<div> <div>9%</div> <div> <div></div> <div>77%</div> <div>8%</div> <div>14%</div> </div> </div>
2	F	191	<div> <div>7%</div> <div> <div></div> <div>83%</div> <div>15%</div> <div></div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	3	 33%67%
4	H	3	 67%33%
4	J	3	 67%33%
4	L	3	 67%33%
5	I	2	 100%
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	1004	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 24123 atoms, of which 11134 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
1	A	321	Total	C	H	N	O	S	0	0	0
			4878	1565	2393	427	482	11			
1	C	321	Total	C	H	N	O	S	0	0	0
			4878	1565	2393	427	482	11			
1	E	321	Total	C	H	N	O	S	0	0	0
			4904	1571	2413	427	482	11			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	LYS	deletion	UNP Q9WFX3
C	?	-	LYS	deletion	UNP Q9WFX3
E	?	-	LYS	deletion	UNP Q9WFX3

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	164	Total	C	H	N	O	S	0	0	0
			2552	825	1235	225	261	6			
2	D	164	Total	C	H	N	O	S	0	0	0
			2551	825	1234	225	261	6			
2	F	163	Total	C	H	N	O	S	0	0	0
			2536	820	1228	224	258	6			

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

*Continued on next page...*

Continued from previous page...

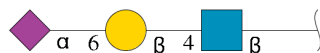
Chain	Residue	Modelled	Actual	Comment	Reference
B	191	ARG	-	expression tag	UNP Q9WFX3
D	186	GLY	-	expression tag	UNP Q9WFX3
D	187	ALA	-	expression tag	UNP Q9WFX3
D	188	LEU	-	expression tag	UNP Q9WFX3
D	189	VAL	-	expression tag	UNP Q9WFX3
D	190	PRO	-	expression tag	UNP Q9WFX3
D	191	ARG	-	expression tag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



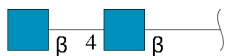
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	3	Total	C	H	N	O	0	0	0
			75	22	36	2	15			

- 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
4	H	3	Total	C	H	N	O	0	0	0
			87	25	41	2	19			
4	J	3	Total	C	H	N	O	0	0	0
			87	25	41	2	19			
4	L	3	Total	C	H	N	O	0	0	0
			87	25	41	2	19			

- 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
5	I	2	Total	C	H	N	O	0	0	0
			54	16	26	2	10			
5	K	2	Total	C	H	N	O	0	0	0
			52	16	25	2	9			

- 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
6	A	1	Total	C	H	N	O	0	0
			28	8	14	1	5		
6	C	1	Total	C	H	N	O	0	0
			28	8	14	1	5		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	329	Total	O	0	0
			329	329		
7	B	104	Total	O	0	0
			104	104		
7	C	327	Total	O	0	0
			327	327		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	111	Total 111	O 111	0	0
7	E	332	Total 332	O 332	0	0
7	F	123	Total 123	O 123	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



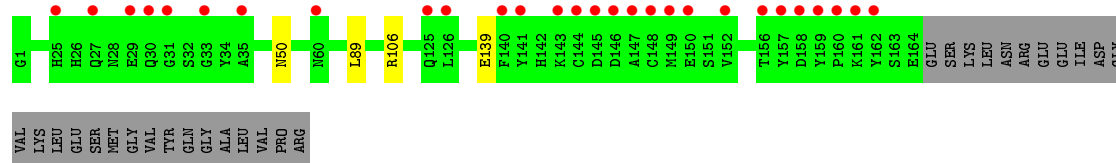
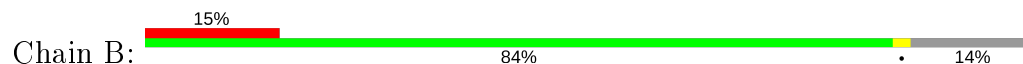
- Molecule 1: Hemagglutinin HA1



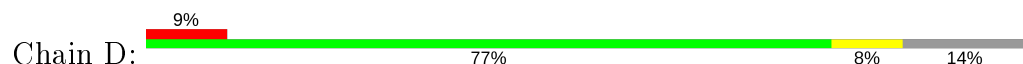
- Molecule 1: Hemagglutinin HA1



- Molecule 2: Hemagglutinin HA2




- Molecule 2: Hemagglutinin HA2





LYS  
LEU  
GLU  
SER  
MET  
GLY  
VAL  
TYR  
GLN  
GLY  
ALA  
LEU  
VAL  
PRO  
ARG


- Molecule 2: Hemagglutinin HA2

Chain F: 



PRO  
ARG

- 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: 

MAG1  
MAG2  
BMA3

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

Chain H: 

MAG1  
GAL2  
SLA3

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

Chain J: 

MAG1  
GAL2  
SLA3

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

Chain L: 

MAG1  
GAL2  
SLA3

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

Chain I: 

MAG1  
MAG2

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

Chain K:  50% 50%

 MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	95.68Å 81.82Å 121.46Å 90.00° 90.79° 90.00°	Depositor
Resolution (Å)	44.72 – 2.15 44.72 – 2.15	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.72-2.15) 100.0 (44.72-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 2.16Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.176 , 0.205 0.182 , 0.209	Depositor DCC
$R_{free}$ test set	5107 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.8	Xtriage
Anisotropy	0.162	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 51.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.021 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	24123	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SIA, GAL, 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	0.34	0/2549	0.57	0/3474
1	C	0.33	0/2549	0.55	1/3474 (0.0%)
1	E	0.34	0/2555	0.57	1/3482 (0.0%)
2	B	0.30	0/1344	0.46	0/1811
2	D	0.32	0/1344	0.49	0/1811
2	F	0.31	0/1335	0.45	0/1799
All	All	0.33	0/11676	0.53	2/15851 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	317	LEU	CA-CB-CG	5.64	128.26	115.30
1	C	69	LEU	N-CA-C	-5.17	97.06	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	2485	2393	2394	17	0
1	C	2485	2393	2394	14	0
1	E	2491	2413	2413	19	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	5	0
2	D	1317	1234	1234	14	0
2	F	1308	1228	1228	7	0
3	G	39	36	34	0	0
4	H	46	41	40	0	0
4	J	46	41	40	0	0
4	L	46	41	40	0	0
5	I	28	26	25	0	0
5	K	27	25	23	1	0
6	A	14	14	13	2	0
6	C	14	14	13	0	0
7	A	329	0	0	6	2
7	B	104	0	0	1	0
7	C	327	0	0	5	1
7	D	111	0	0	6	0
7	E	332	0	0	9	2
7	F	123	0	0	1	0
All	All	12989	11134	11125	64	3

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 (64) 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:192:GLN:OE1	7:C:1101:HOH:O	1.84	0.94
1:A:11:ASN:ND2	7:A:1102:HOH:O	2.03	0.91
1:C:1:ASP:N	7:C:1102:HOH:O	1.90	0.90
1:A:15:THR:OG1	7:A:1101:HOH:O	1.89	0.89
1:C:318:ARG:NH1	7:C:1105:HOH:O	2.21	0.72
1:C:310:LYS:O	7:C:1104:HOH:O	2.08	0.72
1:E:168:LYS:O	7:E:1101:HOH:O	2.12	0.68
1:E:2:THR:HG22	2:F:139:GLU:HG2	1.77	0.67
1:A:307:ARG:NH2	2:D:62:GLN:OE1	2.28	0.66
1:A:280:THR:HG22	1:A:298:THR:HG22	1.78	0.66
2:D:163:SER:OG	7:D:201:HOH:O	2.15	0.64
1:C:307:ARG:NH2	2:F:62:GLN:OE1	2.31	0.64
2:D:72:ASN:OD1	7:D:202:HOH:O	2.16	0.61
2:D:61:THR:HA	7:D:246:HOH:O	2.01	0.60
1:C:280:THR:HG22	1:C:298:THR:HG22	1.85	0.58
1:E:212:GLU:O	1:E:216:ARG:NH2	2.37	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:212:GLU:O	1:C:216:ARG:NH2	2.37	0.56
1:A:304:LYS:NZ	7:A:1114:HOH:O	2.38	0.56
2:D:28:ASN:OD1	7:D:203:HOH:O	2.18	0.55
2:B:50:ASN:OD1	1:E:22:LYS:NZ	2.39	0.55
1:C:304:LYS:HE2	7:D:205:HOH:O	2.07	0.54
1:E:188:GLN:NE2	7:E:1118:HOH:O	2.41	0.53
1:E:280:THR:HG22	1:E:298:THR:HG22	1.90	0.53
1:A:2:THR:HG22	2:B:139:GLU:HG2	1.90	0.53
2:F:106:ARG:HD3	7:F:225:HOH:O	2.08	0.53
1:A:96:ILE:HG13	1:A:229:TYR:CE1	2.45	0.52
1:A:266:THR:O	7:A:1104:HOH:O	2.19	0.52
1:A:170:LYS:NZ	7:A:1113:HOH:O	2.37	0.51
1:E:286:ASN:ND2	7:E:1123:HOH:O	2.43	0.50
1:E:68:ASP:OD1	7:E:1103:HOH:O	2.19	0.50
2:F:159:TYR:HB3	2:F:160:PRO:HD3	1.94	0.49
2:D:106:ARG:NH2	7:D:207:HOH:O	2.44	0.49
7:E:1104:HOH:O	5:K:2:NAG:O6	2.20	0.48
1:E:1:ASP:N	7:E:1124:HOH:O	2.44	0.48
1:C:51:GLN:HG3	7:C:1103:HOH:O	2.13	0.48
2:D:133:ILE:HD11	2:D:139:GLU:HG3	1.97	0.47
1:E:2:THR:CG2	2:F:139:GLU:HG2	2.44	0.47
1:A:307:ARG:HD2	2:D:60:ASN:ND2	2.30	0.47
2:D:86:ASP:HB3	2:F:62:GLN:NE2	2.30	0.47
1:C:285:ILE:HG21	1:C:294:ILE:HG13	1.96	0.46
1:A:286:ASN:HD21	6:A:1004:NAG:C6	2.28	0.46
1:E:23:ASN:ND2	7:E:1130:HOH:O	2.49	0.46
1:E:292:GLN:HG3	1:E:292:GLN:O	2.16	0.46
1:A:305:TYR:CD2	2:B:89:LEU:HD13	2.51	0.45
1:A:216:ARG:HG2	1:C:201:GLY:HA3	1.98	0.45
1:A:35:ASP:C	1:A:294:ILE:HD11	2.37	0.44
2:D:161:LYS:NZ	2:D:162:TYR:OH	2.47	0.44
1:E:96:ILE:HG13	1:E:229:TYR:CE1	2.52	0.44
1:E:2:THR:HG22	2:F:139:GLU:CG	2.46	0.44
2:D:26:HIS:HB2	2:D:149:MET:SD	2.58	0.44
1:A:286:ASN:ND2	6:A:1004:NAG:O6	2.50	0.43
2:D:133:ILE:HD11	2:D:139:GLU:N	2.34	0.43
2:B:106:ARG:HD3	7:B:211:HOH:O	2.18	0.43
1:E:317:LEU:HD22	1:E:318:ARG:O	2.20	0.42
2:D:9:PHE:O	2:D:135:ASN:HA	2.20	0.42
1:C:2:THR:HG22	2:D:139:GLU:HG2	2.02	0.41
1:E:23:ASN:ND2	7:E:1117:HOH:O	2.41	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:1213:HOH:O	1:C:208:ARG:HD3	2.21	0.41
1:A:307:ARG:HG3	2:B:89:LEU:HD11	2.03	0.41
1:E:318:ARG:NE	7:E:1102:HOH:O	2.17	0.41
1:C:317:LEU:HD12	1:C:318:ARG:O	2.21	0.41
1:A:64:ASN:HB3	1:A:67:CYS:SG	2.61	0.41
1:E:75:SER:HB3	1:E:109:SER:O	2.21	0.40
1:E:35:ASP:C	1:E:294:ILE:HD11	2.41	0.40

All (3) 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:A:1362:HOH:O	7:C:1340:HOH:O[2_757]	1.95	0.25
7:A:1294:HOH:O	7:E:1384:HOH:O[2_857]	2.00	0.20
7:E:1399:HOH:O	7:E:1432:HOH:O[2_847]	2.19	0.01

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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%)	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%)	158 (98%)	4 (2%)	0	100	100
2	F	161/191 (84%)	159 (99%)	2 (1%)	0	100	100
All	All	1442/1551 (93%)	1415 (98%)	27 (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	275/282 (98%)	275 (100%)	0	100	100
1	C	275/282 (98%)	275 (100%)	0	100	100
1	E	277/282 (98%)	276 (100%)	1 (0%)	91	93
2	B	139/162 (86%)	139 (100%)	0	100	100
2	D	139/162 (86%)	139 (100%)	0	100	100
2	F	138/162 (85%)	138 (100%)	0	100	100
All	All	1243/1332 (93%)	1242 (100%)	1 (0%)	93	96

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	317	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	193	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

16 monosaccharides are modelled in this entry.



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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
3	NAG	G	1	1,3	14,14,15	0.69	1 (7%)	17,19,21	0.71	0
3	NAG	G	2	3	14,14,15	0.24	0	17,19,21	0.52	0
3	BMA	G	3	3	11,11,12	0.97	1 (9%)	15,15,17	0.96	0
4	NAG	H	1	4	15,15,15	0.18	0	21,21,21	0.27	0
4	GAL	H	2	4	11,11,12	0.44	0	15,15,17	0.94	0
4	SIA	H	3	4	17,20,21	1.04	1 (5%)	21,28,31	1.07	2 (9%)
5	NAG	I	1	1,5	14,14,15	0.53	0	17,19,21	0.65	0
5	NAG	I	2	5	14,14,15	0.48	0	17,19,21	0.53	0
4	NAG	J	1	4	15,15,15	0.16	0	21,21,21	0.23	0
4	GAL	J	2	4	11,11,12	0.72	0	15,15,17	0.97	0
4	SIA	J	3	4	17,20,21	1.24	3 (17%)	21,28,31	1.19	3 (14%)
5	NAG	K	1	1,5	14,14,15	0.25	0	17,19,21	0.58	0
5	NAG	K	2	5	13,13,15	0.97	1 (7%)	16,17,21	0.62	0
4	NAG	L	1	4	15,15,15	0.26	0	21,21,21	0.32	0
4	GAL	L	2	4	11,11,12	0.66	0	15,15,17	0.99	0
4	SIA	L	3	4	17,20,21	0.96	1 (5%)	21,28,31	1.15	2 (9%)

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	-	2/6/23/26	0/1/1/1
3	NAG	G	2	3	-	0/6/23/26	0/1/1/1
3	BMA	G	3	3	-	0/2/19/22	0/1/1/1
4	NAG	H	1	4	-	1/6/26/26	0/1/1/1
4	GAL	H	2	4	-	0/2/19/22	0/1/1/1
4	SIA	H	3	4	-	0/14/34/38	0/1/1/1
5	NAG	I	1	1,5	-	2/6/23/26	0/1/1/1
5	NAG	I	2	5	-	0/6/23/26	0/1/1/1
4	NAG	J	1	4	-	0/6/26/26	0/1/1/1
4	GAL	J	2	4	-	0/2/19/22	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	SIA	J	3	4	-	0/14/34/38	0/1/1/1
5	NAG	K	1	1,5	-	2/6/23/26	0/1/1/1
5	NAG	K	2	5	-	0/6/19/26	0/1/1/1
4	NAG	L	1	4	-	0/6/26/26	0/1/1/1
4	GAL	L	2	4	-	0/2/19/22	0/1/1/1
4	SIA	L	3	4	-	0/14/34/38	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	K	2	NAG	C1-C2	3.05	1.54	1.51
3	G	1	NAG	O5-C1	-2.49	1.39	1.43
4	H	3	SIA	C10-N5	2.35	1.42	1.34
3	G	3	BMA	C2-C3	2.29	1.55	1.52
4	J	3	SIA	C4-C5	-2.29	1.51	1.53
4	J	3	SIA	C5-N5	-2.28	1.42	1.45
4	J	3	SIA	C10-N5	2.23	1.42	1.34
4	L	3	SIA	C10-N5	2.19	1.41	1.34

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	J	3	SIA	C3-C4-C5	-2.90	107.95	111.46
4	L	3	SIA	C3-C4-C5	-2.86	108.00	111.46
4	H	3	SIA	C3-C4-C5	-2.41	108.55	111.46
4	J	3	SIA	O6-C2-C3	-2.38	105.70	109.87
4	H	3	SIA	O6-C2-C3	-2.16	106.08	109.87
4	J	3	SIA	C9-C8-C7	-2.12	107.81	112.41
4	L	3	SIA	O6-C2-C3	-2.02	106.32	109.87

There are no chirality outliers.

All (7) torsion outliers are listed below:

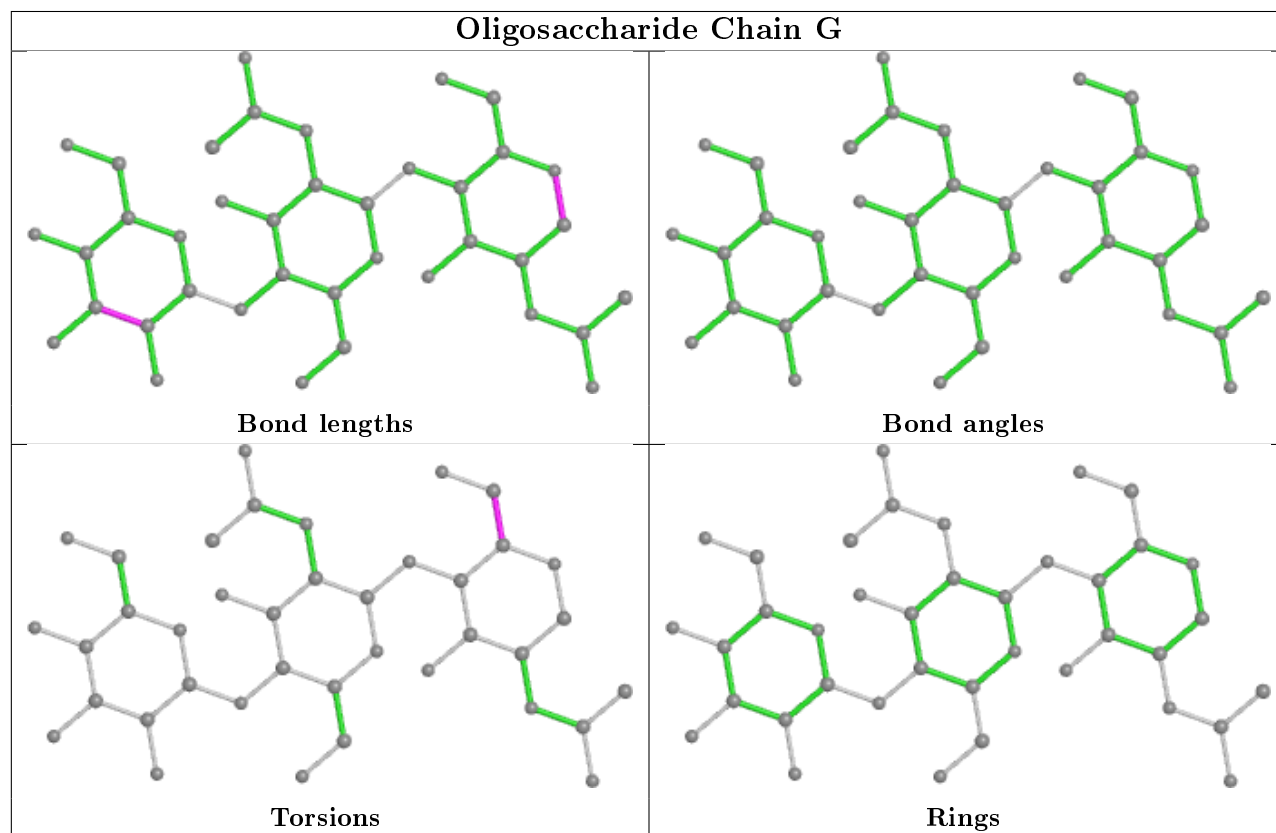
Mol	Chain	Res	Type	Atoms
5	I	1	NAG	O5-C5-C6-O6
5	K	1	NAG	O5-C5-C6-O6
5	I	1	NAG	C4-C5-C6-O6
3	G	1	NAG	C4-C5-C6-O6
3	G	1	NAG	O5-C5-C6-O6
5	K	1	NAG	C4-C5-C6-O6
4	H	1	NAG	C4-C5-C6-O6

There are no ring outliers.

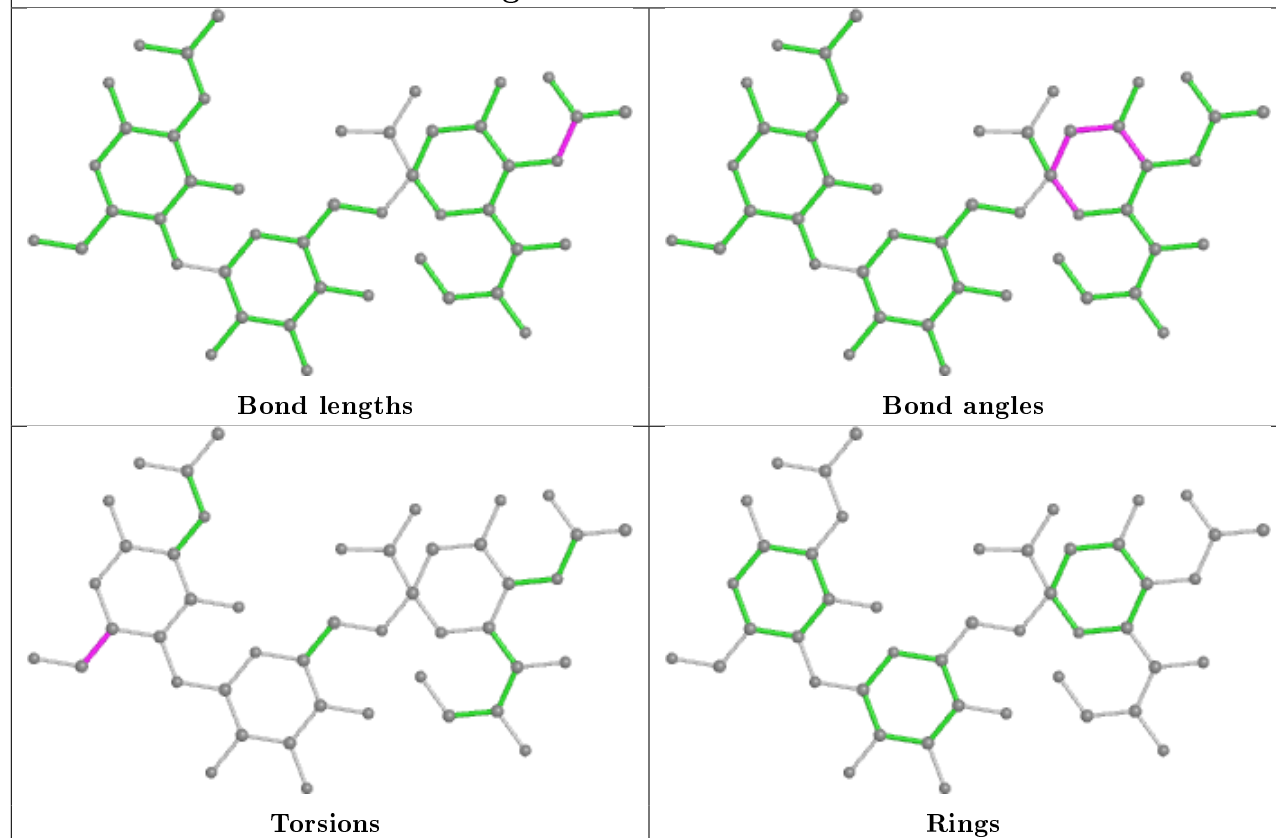
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	K	2	NAG	1	0

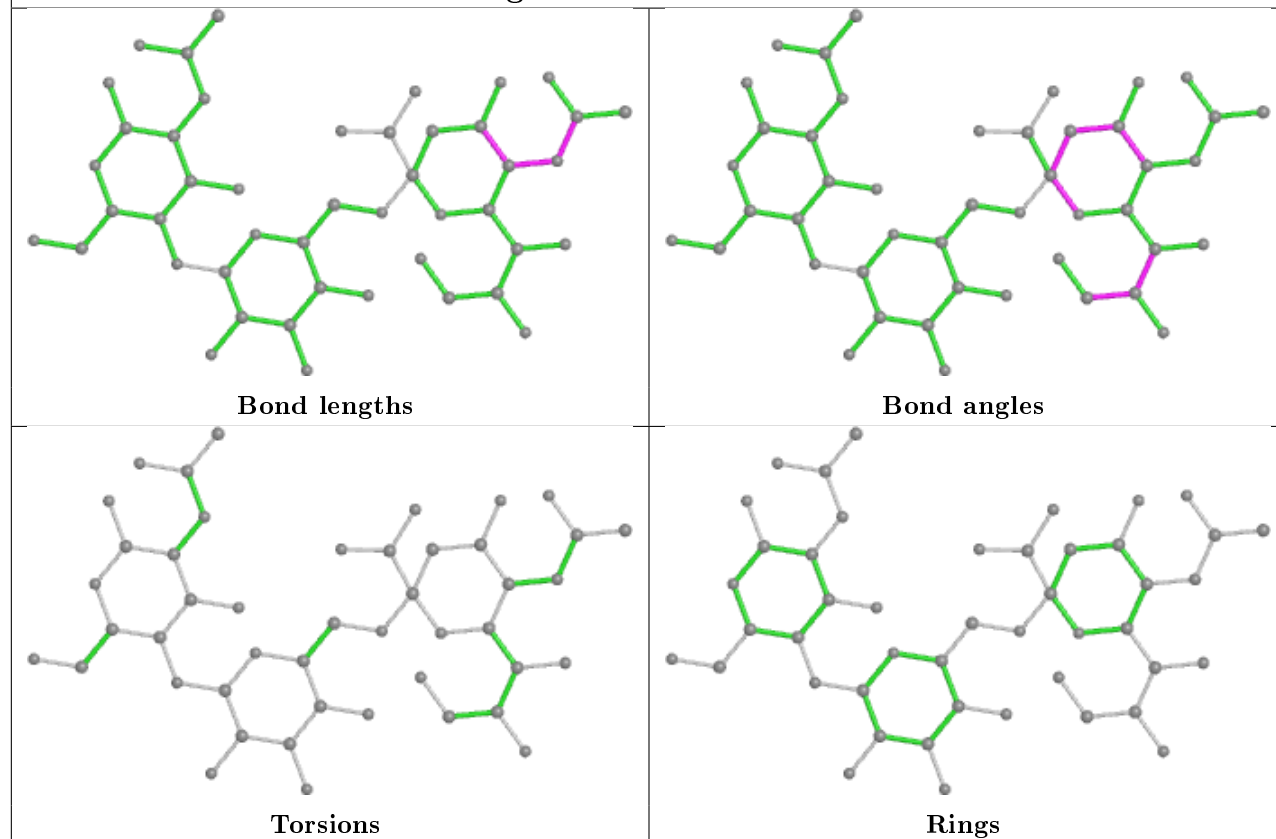
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

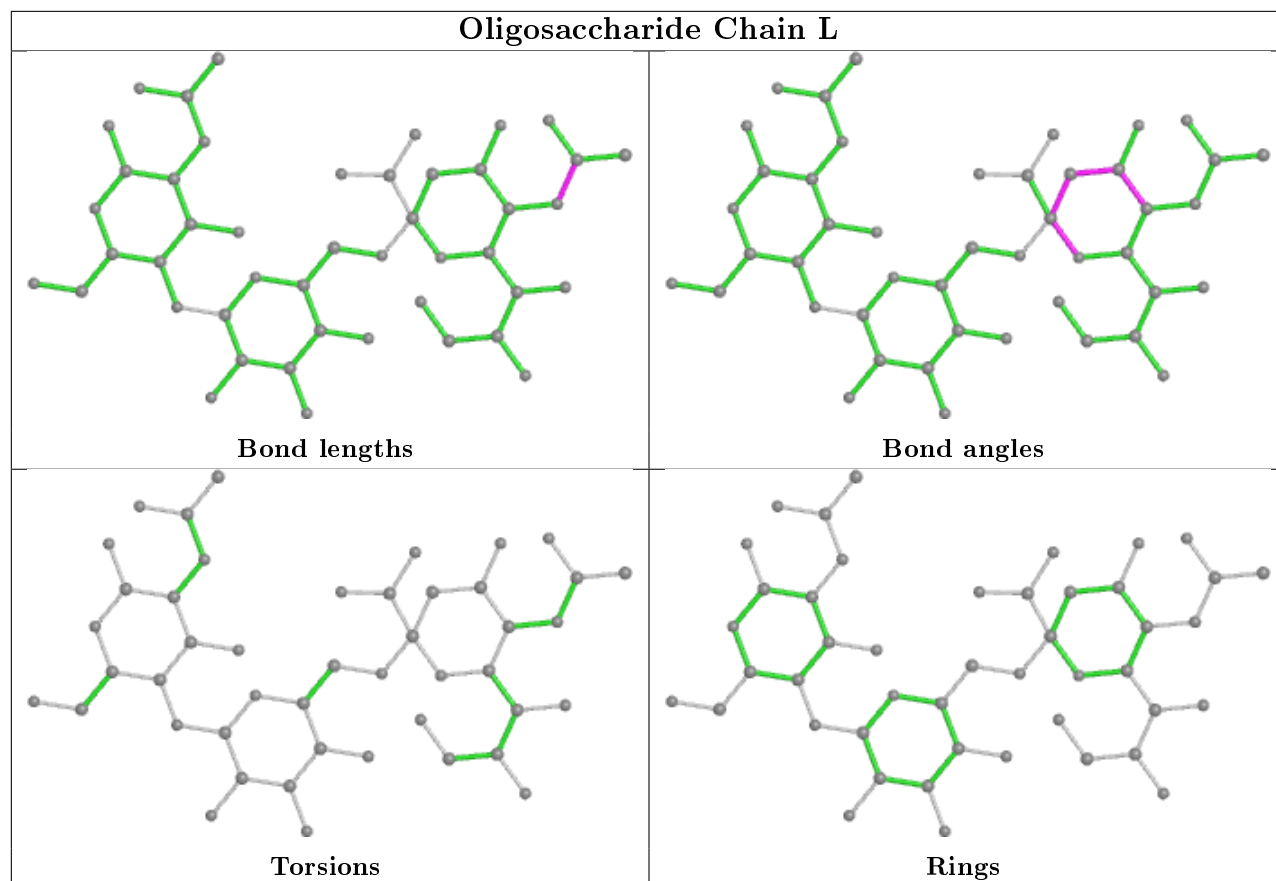


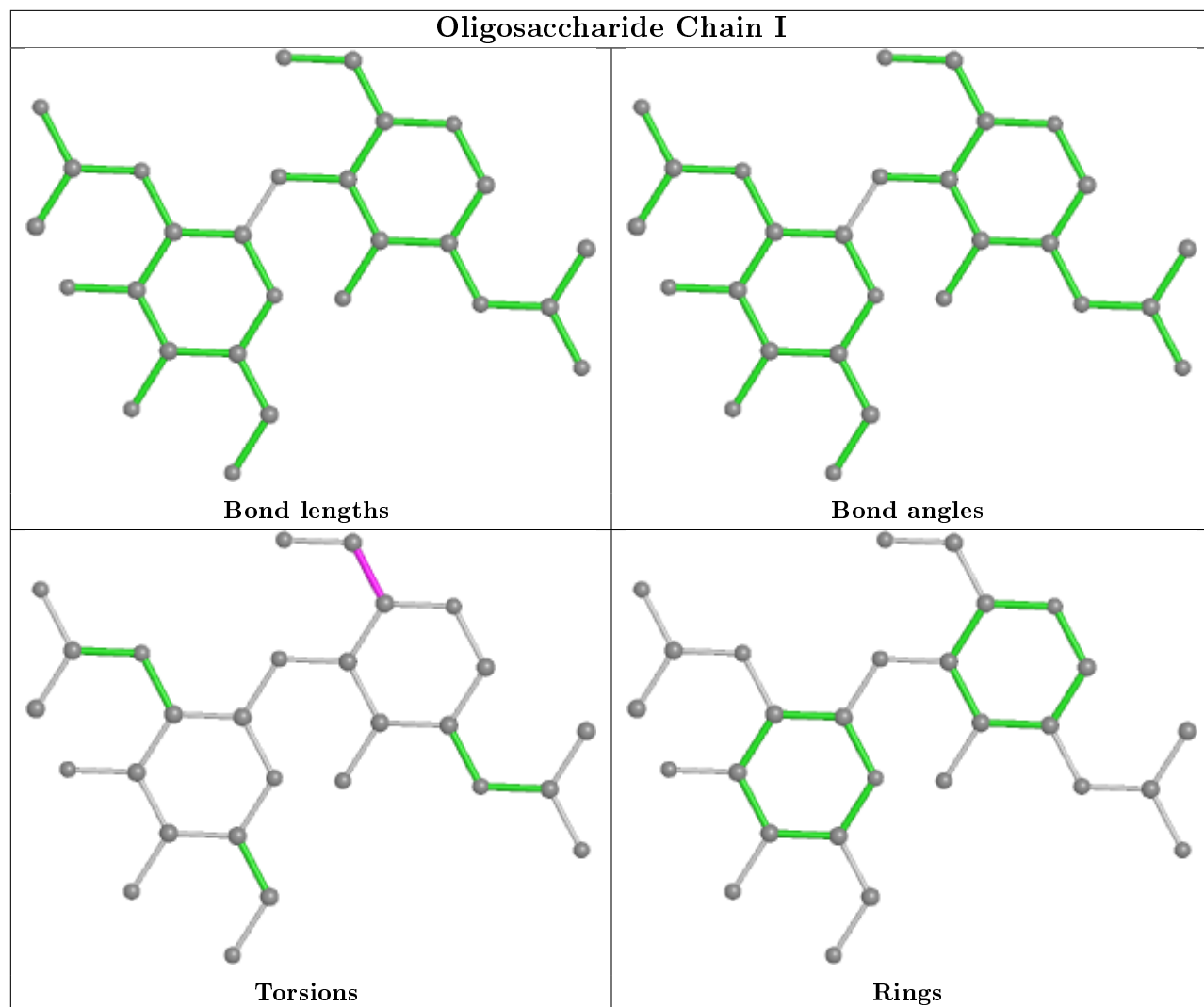
## Oligosaccharide Chain H

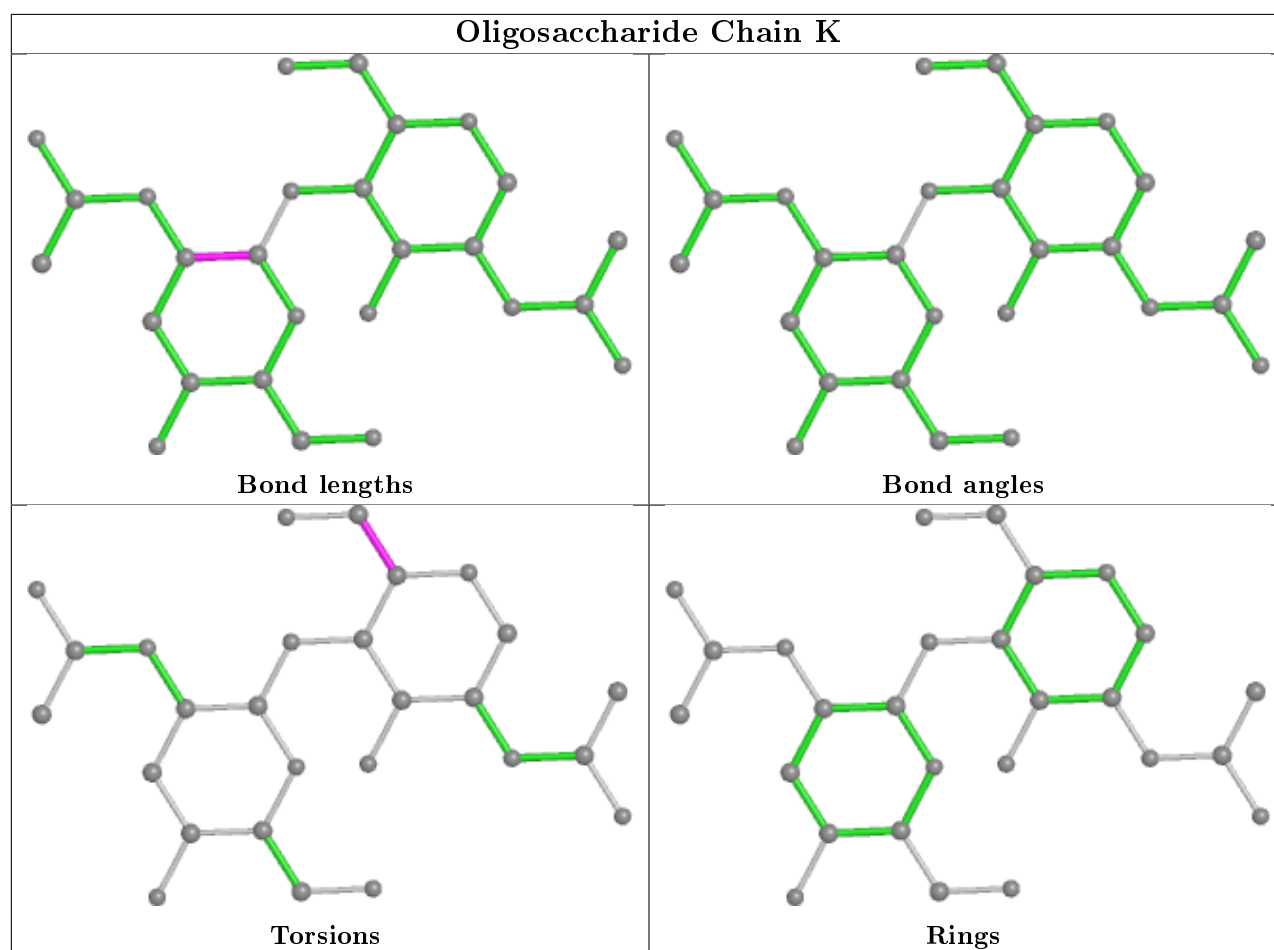


## Oligosaccharide Chain J









## 5.6 Ligand geometry [i](#)

2 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	1003	1	14,14,15	0.64	0	17,19,21	0.82	1 (5%)
6	NAG	A	1004	1	14,14,15	0.73	1 (7%)	17,19,21	0.74	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	1003	1	-	4/6/23/26	0/1/1/1
6	NAG	A	1004	1	-	1/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1004	NAG	O5-C1	2.24	1.47	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	1003	NAG	C1-O5-C5	3.10	116.39	112.19
6	A	1004	NAG	C1-O5-C5	2.73	115.89	112.19

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	1003	NAG	C4-C5-C6-O6
6	C	1003	NAG	C1-C2-N2-C7
6	C	1003	NAG	O5-C5-C6-O6
6	A	1004	NAG	C1-C2-N2-C7
6	C	1003	NAG	C3-C2-N2-C7

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1004	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

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	321/326 (98%)	-0.04	11 (3%) 45 53	13, 24, 47, 82	0
1	C	321/326 (98%)	-0.06	11 (3%) 45 53	15, 27, 50, 77	0
1	E	321/326 (98%)	-0.18	6 (1%) 66 74	14, 23, 43, 67	0
2	B	164/191 (85%)	0.79	28 (17%) 1 1	13, 41, 73, 87	0
2	D	164/191 (85%)	0.56	18 (10%) 5 8	13, 39, 66, 80	0
2	F	163/191 (85%)	0.57	14 (8%) 10 15	13, 39, 63, 75	0
All	All	1454/1551 (93%)	0.16	88 (6%) 21 28	13, 28, 62, 87	0

All (88) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	71	LEU	6.8
1	C	69	LEU	5.6
2	B	144	CYS	5.5
1	C	68	ASP	5.3
2	B	29	GLU	5.0
2	B	162	TYR	4.9
1	A	154	GLY	4.8
2	B	147	ALA	4.8
2	D	160	PRO	4.8
2	B	140	PHE	4.4
1	A	69	LEU	4.2
2	B	148	CYS	4.1
2	B	141	TYR	4.1
2	B	143	LYS	4.1
1	A	70	LEU	4.0
2	F	18	ILE	3.9
2	D	163	SER	3.7
2	B	31	GLY	3.7
2	D	162	TYR	3.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	160	PRO	3.5
1	E	70	LEU	3.5
2	B	145	ASP	3.5
2	F	134	GLY	3.4
2	B	156	THR	3.4
2	D	161	LYS	3.4
2	B	159	TYR	3.3
2	B	33	GLY	3.3
2	D	35	ALA	3.3
1	A	2	THR	3.2
2	D	18	ILE	3.2
1	A	261	GLY	3.2
2	F	31	GLY	3.2
2	F	32	SER	3.2
1	C	70	LEU	3.1
1	C	154	GLY	3.1
2	D	29	GLU	3.1
1	C	84	ASN	3.0
2	B	27	GLN	3.0
2	B	158	ASP	3.0
2	B	35	ALA	3.0
2	D	156	THR	2.9
2	F	139	GLU	2.9
2	D	158	ASP	2.9
1	E	71	LEU	2.9
2	F	147	ALA	2.8
2	B	126	LEU	2.8
2	F	133	ILE	2.8
2	F	27	GLN	2.8
1	C	53	GLY	2.8
2	B	30	GLN	2.8
1	A	4	CYS	2.8
2	B	152	VAL	2.7
2	F	144	CYS	2.7
1	A	13	THR	2.7
2	D	31	GLY	2.7
2	B	157	TYR	2.6
2	D	149	MET	2.6
2	B	161	LYS	2.6
2	D	140	PHE	2.6
2	D	147	ALA	2.6
1	A	68	ASP	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	F	34	TYR	2.5
2	D	157	TYR	2.5
1	C	52	LEU	2.5
2	B	149	MET	2.4
1	C	137	TYR	2.4
2	B	25	HIS	2.4
2	D	148	CYS	2.3
2	F	153	ARG	2.3
1	C	273	ASP	2.3
2	F	36	ALA	2.3
2	B	146	ASP	2.3
1	E	110	SER	2.3
2	B	150	GLU	2.2
2	F	33	GLY	2.2
1	E	69	LEU	2.2
1	A	155	SER	2.2
2	F	141	TYR	2.2
2	D	38	GLN	2.2
2	B	125	GLN	2.1
1	E	84	ASN	2.1
1	C	185	THR	2.1
2	D	134	GLY	2.1
2	D	32	SER	2.0
1	A	1	ASP	2.0
1	E	11	ASN	2.0
2	B	60	ASN	2.0
1	A	71	LEU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

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

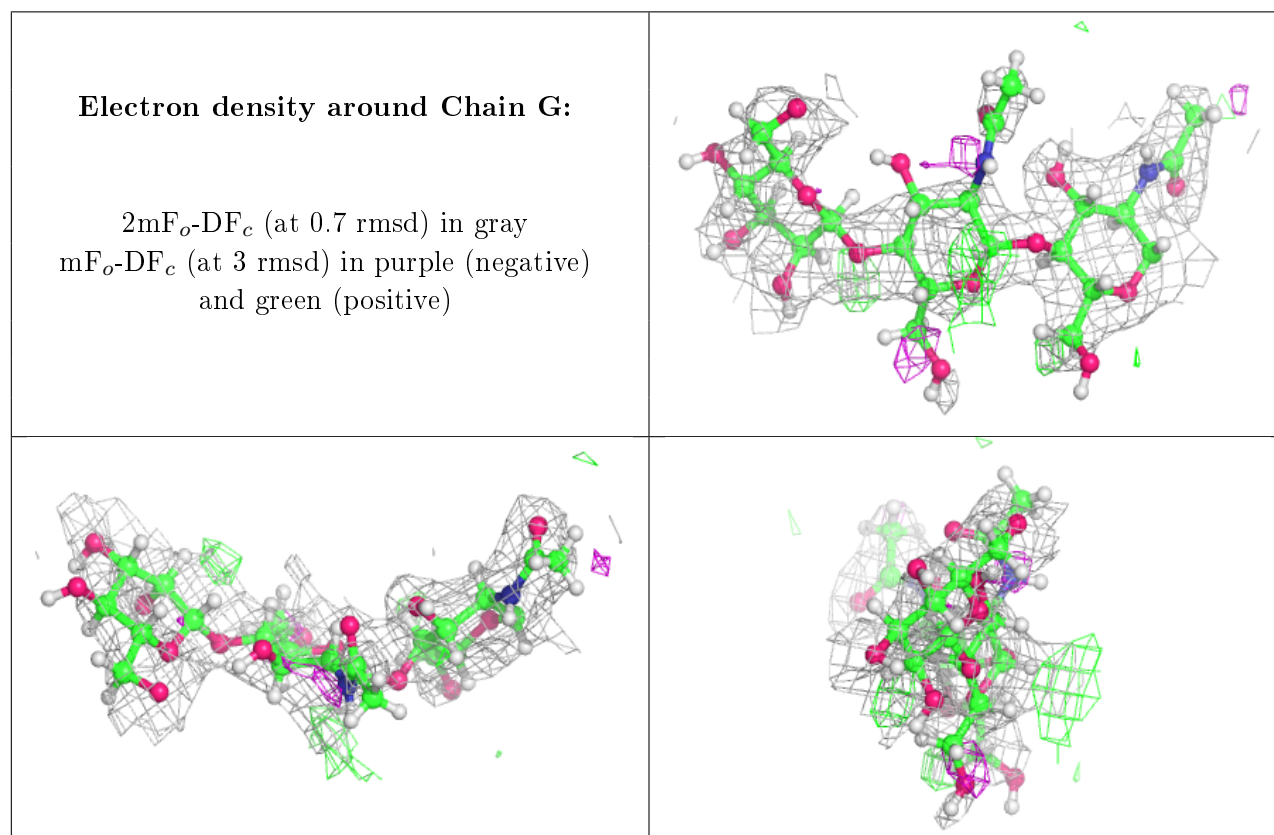
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	BMA	G	3	11/12	0.67	0.29	66,79,92,102	0

*Continued on next page...*

*Continued from previous page...*

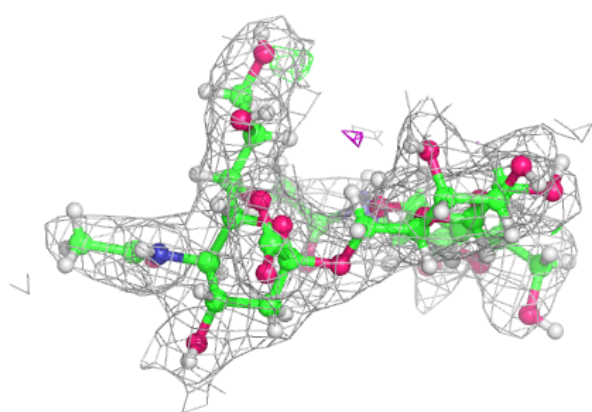
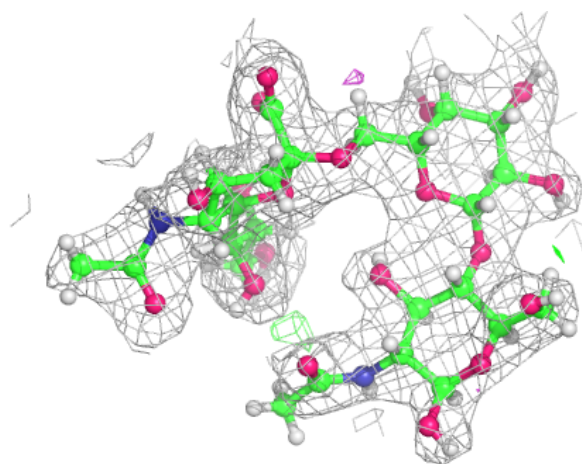
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	G	2	14/15	0.81	0.37	60,75,104,109	0
4	NAG	H	1	15/15	0.83	0.24	40,52,63,72	0
5	NAG	I	2	14/15	0.86	0.21	48,61,74,87	0
5	NAG	K	2	13/15	0.90	0.14	34,46,53,61	0
4	NAG	J	1	15/15	0.90	0.15	39,49,63,65	0
4	GAL	L	2	11/12	0.94	0.10	22,31,38,39	0
5	NAG	I	1	14/15	0.94	0.15	24,35,51,67	0
4	GAL	J	2	11/12	0.94	0.08	29,36,44,52	0
5	NAG	K	1	14/15	0.95	0.09	18,27,42,59	0
4	NAG	L	1	15/15	0.95	0.11	30,38,48,65	0
3	NAG	G	1	14/15	0.95	0.13	24,37,63,85	0
4	SIA	J	3	20/21	0.97	0.10	21,31,37,39	0
4	GAL	H	2	11/12	0.97	0.10	26,32,39,46	0
4	SIA	L	3	20/21	0.98	0.07	15,20,24,28	0
4	SIA	H	3	20/21	0.98	0.08	19,26,33,34	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.



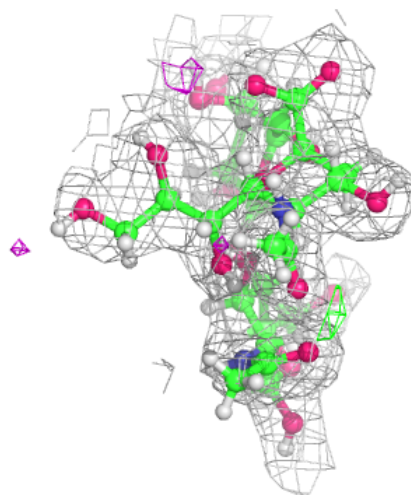
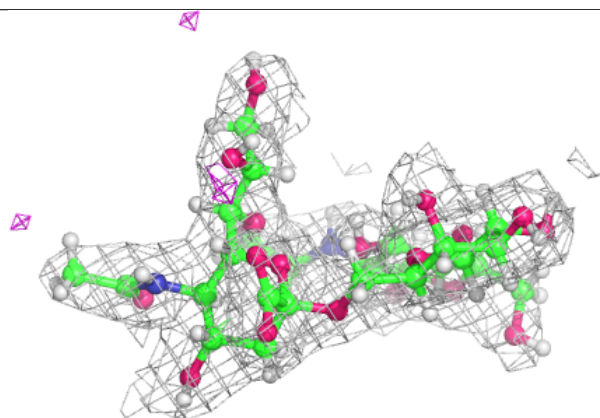
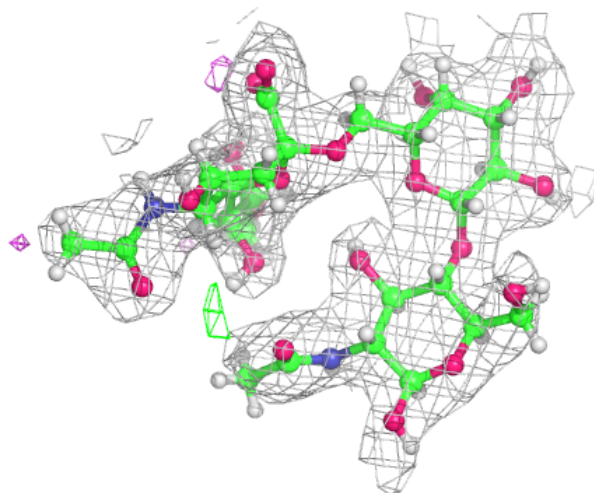
**Electron density around Chain H:**

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



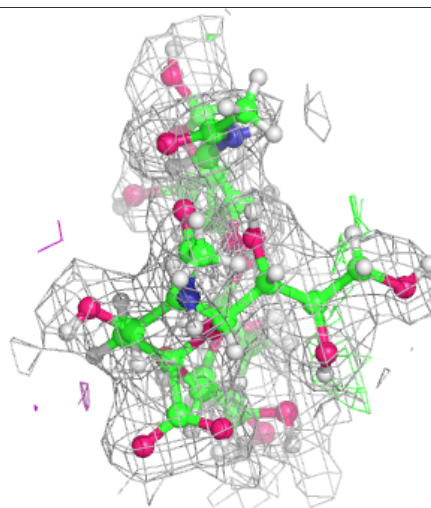
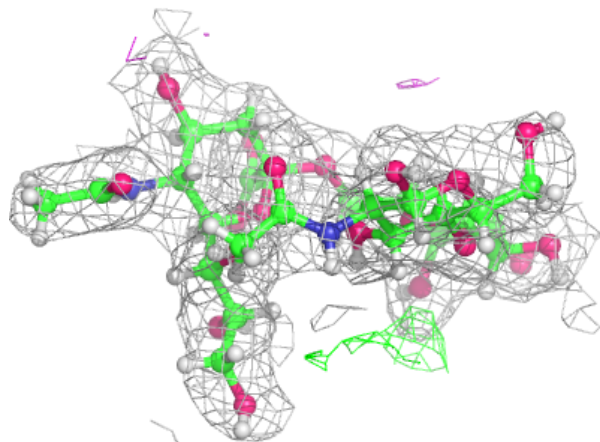
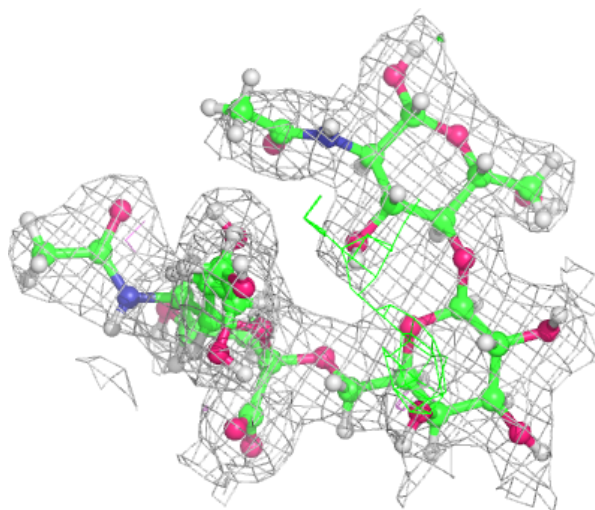
**Electron density around Chain J:**

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



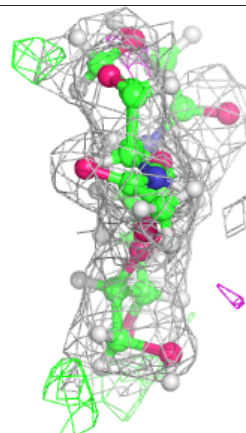
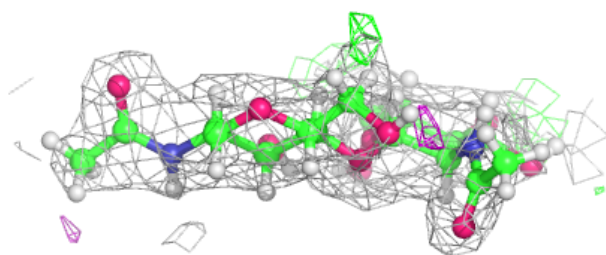
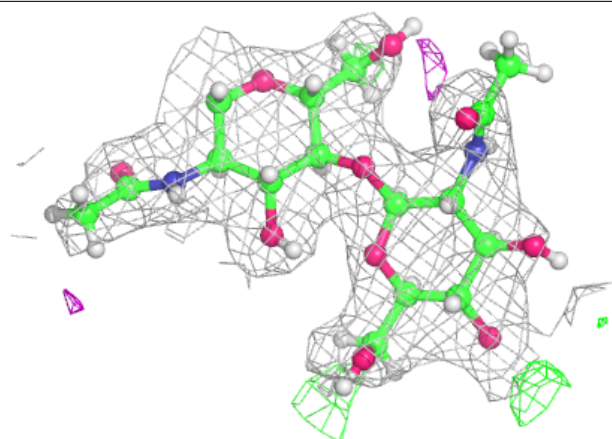
**Electron density around Chain L:**

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

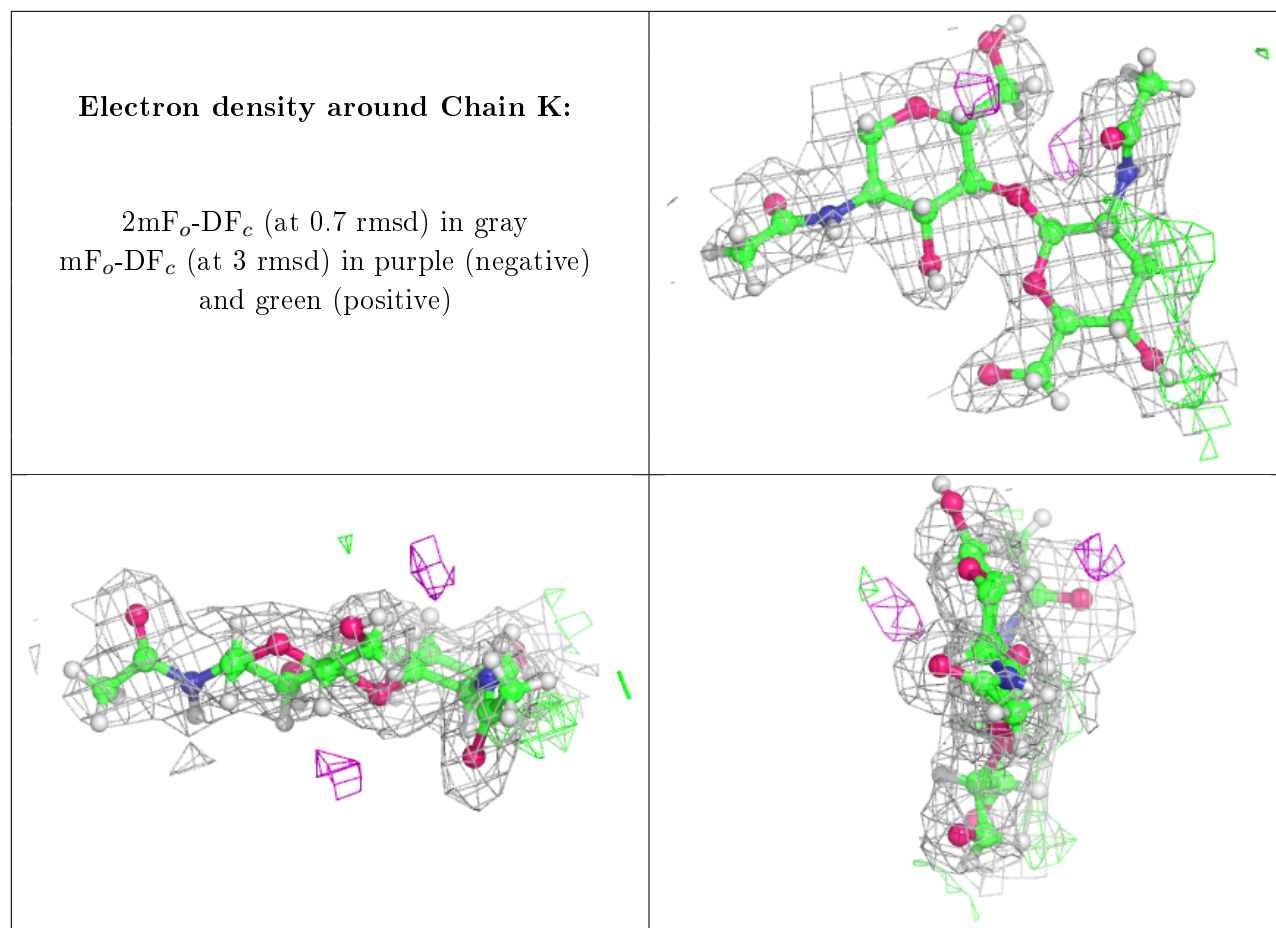


**Electron density around Chain I:**

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







## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	NAG	A	1004	14/15	0.52	0.43	68,87,103,114	0
6	NAG	C	1003	14/15	0.81	0.22	47,58,70,96	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.