



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 10:12 AM EDT

PDB ID : 9PIB / pdb\_00009pib  
Title : Crystal structure of the A/Puerto Rico/8/1934 (H1N1) influenza virus hemagglutinin in complex with fusion inhibitor CLIPS peptide CP121132 (CP-S1)  
Authors : Kadam, R.U.; Wilson, I.A.  
Deposited on : 2025-07-10  
Resolution : 2.35 Å (reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4-5-2 with Phenix2.0  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.010 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.48.1

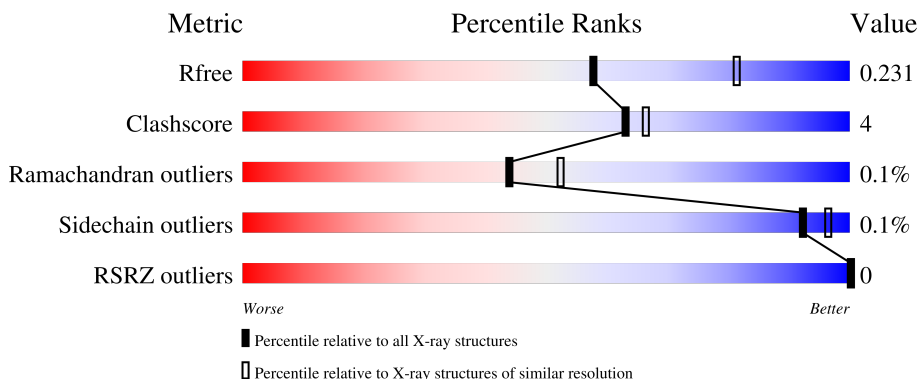
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


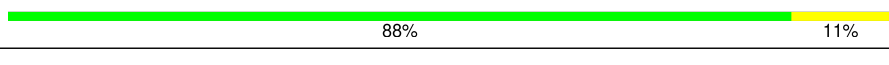
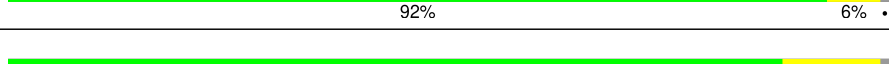
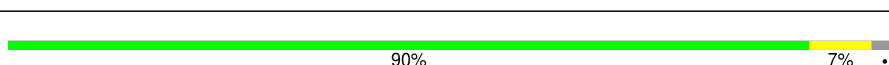

The reported resolution of this entry is 2.35 Å.

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))
$R_{free}$	164625	1460 (2.36-2.36)
Clashscore	180529	1571 (2.36-2.36)
Ramachandran outliers	177936	1559 (2.36-2.36)
Sidechain outliers	177891	1559 (2.36-2.36)
RSRZ outliers	164620	1460 (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  $\geq 3$ , 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	 88% 10%
1	C	326	 88% 11%
1	E	326	 92% 6%
1	G	326	 87% 11%
2	B	176	 90% 7%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	176	 88% 10%
2	F	176	 90% 7%
2	H	176	 88% 10%
3	N	13	 92% 8%
3	O	13	 85% 15%
3	P	13	 69% 23% 8%
3	Q	13	 69% 31%
4	J	3	 67% 33%
5	I	3	 67% 33%
5	L	3	 67% 33%
6	K	2	 100%
6	M	2	 100%

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 17374 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	322	Total 2542	C 1603	N 443	O 483	S 13	0	0	0
1	C	322	Total 2530	C 1595	N 441	O 481	S 13	0	0	0
1	E	322	Total 2536	C 1600	N 440	O 483	S 13	0	0	0
1	G	322	Total 2538	C 1600	N 442	O 483	S 13	0	0	0

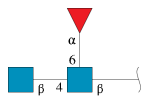
- Molecule 2 is a protein called Hemagglutinin HA2 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	171	Total 1376	C 864	N 235	O 270	S 7	0	0	0
2	D	171	Total 1380	C 866	N 235	O 272	S 7	0	0	0
2	F	171	Total 1380	C 866	N 235	O 272	S 7	0	0	0
2	H	171	Total 1376	C 863	N 234	O 272	S 7	0	0	0

- Molecule 3 is a protein called CLIPS peptide CP121132.

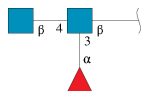
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	N	13	Total 101	C 66	N 16	O 16	S 3	0	0	1
3	O	13	Total 101	C 66	N 16	O 16	S 3	0	0	1
3	P	13	Total 101	C 66	N 16	O 16	S 3	0	0	1
3	Q	13	Total 101	C 66	N 16	O 16	S 3	0	0	1

- Molecule 4 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	J	3	38	22	2	14	0	0	0

- Molecule 5 is an oligosaccharide called alpha-L-fucopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	I	3	38	22	2	14	0	0	0
5	L	3	38	22	2	14	0	0	0

- Molecule 6 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
6	K	2	28	16	2	10	0	0	0
6	M	2	28	16	2	10	0	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



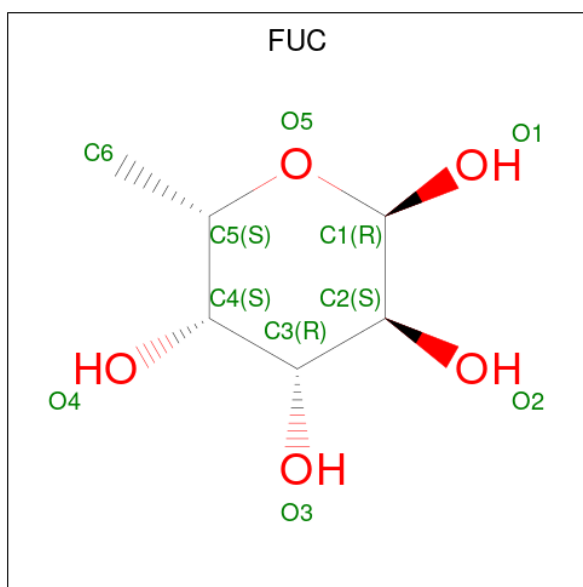
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
7	A	1	Total 14	C 8	N 1	O 5	0	0
7	C	1	Total 14	C 8	N 1	O 5	0	0
7	D	1	Total 14	C 8	N 1	O 5	0	0
7	E	1	Total 14	C 8	N 1	O 5	0	0
7	F	1	Total 14	C 8	N 1	O 5	0	0
7	G	1	Total 14	C 8	N 1	O 5	0	0
7	H	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 8 is DI(HYDROXYETHYL)ETHER (CCD ID: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



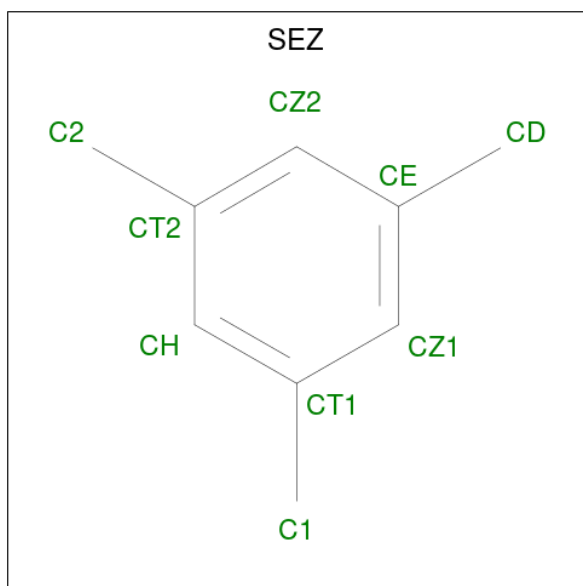
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	1	Total C O 7 4 3	0	0
8	A	1	Total C O 7 4 3	0	0
8	B	1	Total C O 7 4 3	0	0
8	C	1	Total C O 7 4 3	0	0
8	E	1	Total C O 7 4 3	0	0
8	E	1	Total C O 7 4 3	0	0
8	F	1	Total C O 7 4 3	0	0
8	G	1	Total C O 7 4 3	0	0
8	N	1	Total C O 7 4 3	0	0
8	N	1	Total C O 7 4 3	0	0

- Molecule 9 is alpha-L-fucopyranose (CCD ID: FUC) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>5</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	G	1	Total	C O	0	0
			10	6 4		

- Molecule 10 is 1,3,5-trimethylbenzene (CCD ID: SEZ) (formula: C<sub>9</sub>H<sub>12</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	N	1	Total	C	0	0
			9	9		
10	O	1	Total	C	0	0
			9	9		
10	P	1	Total	C	0	0
			9	9		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	Q	1	Total C 9 9	0	0

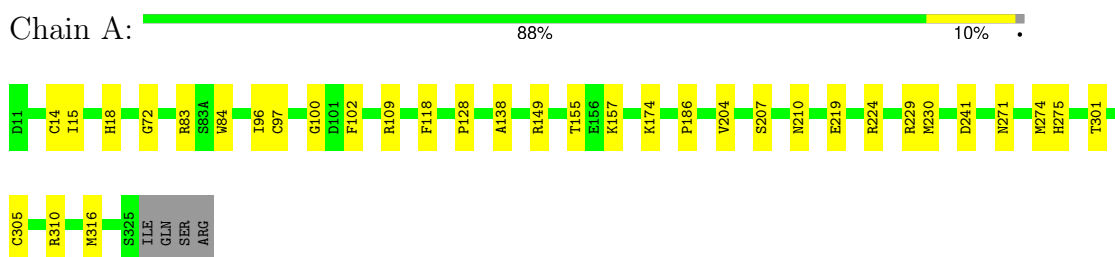
- Molecule 11 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	182	Total O 182 182	0	0
11	B	120	Total O 120 120	0	0
11	C	124	Total O 124 124	0	0
11	D	50	Total O 50 50	0	0
11	E	179	Total O 179 179	0	0
11	F	108	Total O 108 108	0	0
11	G	95	Total O 95 95	0	0
11	H	56	Total O 56 56	0	0
11	N	5	Total O 5 5	0	0
11	O	7	Total O 7 7	0	0
11	Q	2	Total O 2 2	0	0

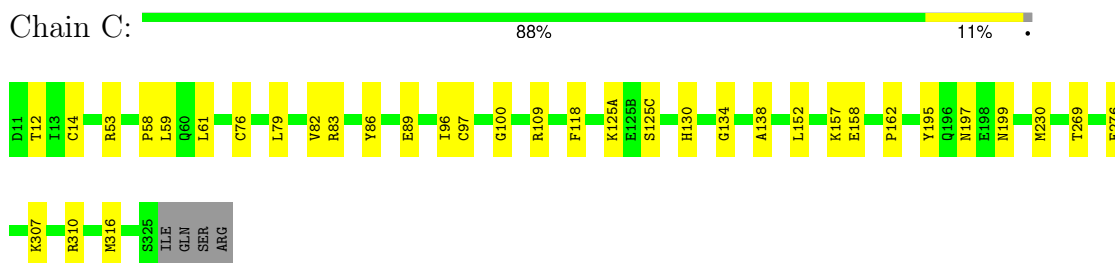
### 3 Residue-property plots [i](#)

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.

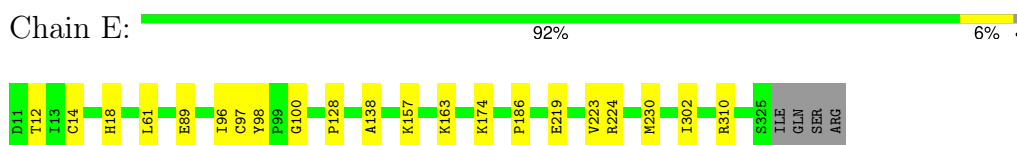
- Molecule 1: Hemagglutinin



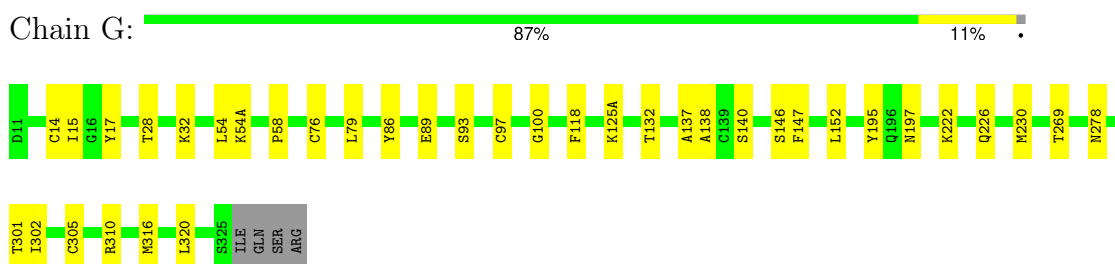
- Molecule 1: Hemagglutinin




- Molecule 1: Hemagglutinin

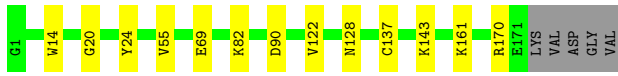


- Molecule 1: Hemagglutinin




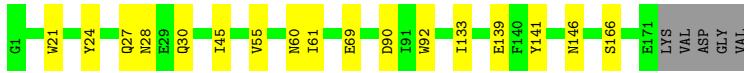
- Molecule 2: Hemagglutinin HA2 chain

Chain B:  90% 7%




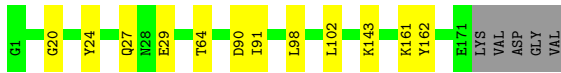
- Molecule 2: Hemagglutinin HA2 chain

Chain D:  88% 10%




- Molecule 2: Hemagglutinin HA2 chain

Chain F:  90% 7%



- Molecule 2: Hemagglutinin HA2 chain

Chain H:  88% 10%




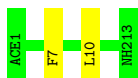
- Molecule 3: CLIPS peptide CP121132

Chain N:  92% 8%



- Molecule 3: CLIPS peptide CP121132

Chain O:  85% 15%



- Molecule 3: CLIPS peptide CP121132

Chain P:  69% 23% 8%



- Molecule 3: CLIPS peptide CP121132

Chain Q:  69% 31%



- Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-[alpha-L-fucopyranose-(1-6)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain J:  67% 33%



- Molecule 5: alpha-L-fucopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain I:  67% 33%



- Molecule 5: alpha-L-fucopyranose-(1-3)-[2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)]2-acetamido-2-deoxy-beta-D-glucopyranose

Chain L:  67% 33%



- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K:  100%



- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain M:  100%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	93.62Å 93.62Å 279.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.81 – 2.35 46.81 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.8 (46.81-2.35) 99.8 (46.81-2.35)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.87 (at 2.34Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660, PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.189 , 0.231 0.190 , 0.231	Depositor DCC
$R_{free}$ test set	5700 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.5	Xtrriage
Anisotropy	0.632	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 20.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtrriage
Estimated twinning fraction	0.015 for -h,-k,l 0.467 for h,-h-k,-l 0.017 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	17374	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% 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: FUC, NAG, ACE, PEG, SEZ, NH2

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	0.15	0/2606	0.40	0/3544
1	C	0.15	0/2594	0.40	0/3531
1	E	0.15	0/2600	0.41	0/3537
1	G	0.14	0/2602	0.38	0/3540
2	B	0.17	0/1403	0.40	0/1886
2	D	0.18	0/1407	0.42	0/1891
2	F	0.17	0/1407	0.38	0/1891
2	H	0.18	0/1403	0.41	0/1887
3	N	0.28	0/101	0.48	0/136
3	O	0.27	0/101	0.44	0/136
3	P	0.28	0/101	0.49	0/136
3	Q	0.29	0/101	0.47	0/136
All	All	0.16	0/16426	0.40	0/22251

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2542	0	2473	24	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2530	0	2445	24	0
1	E	2536	0	2462	16	0
1	G	2538	0	2460	22	0
2	B	1376	0	1305	14	0
2	D	1380	0	1309	14	0
2	F	1380	0	1309	11	0
2	H	1376	0	1298	15	0
3	N	101	0	89	1	0
3	O	101	0	89	1	0
3	P	101	0	89	4	0
3	Q	101	0	89	4	0
4	J	38	0	34	0	0
5	I	38	0	34	0	0
5	L	38	0	34	0	0
6	K	28	0	25	0	0
6	M	28	0	25	0	0
7	A	14	0	13	0	0
7	C	14	0	13	0	0
7	D	14	0	13	0	0
7	E	14	0	13	0	0
7	F	14	0	13	0	0
7	G	14	0	13	0	0
7	H	14	0	13	0	0
8	A	14	0	20	2	0
8	B	7	0	10	3	0
8	C	7	0	10	1	0
8	E	14	0	20	0	0
8	F	7	0	10	0	0
8	G	7	0	10	0	0
8	N	14	0	20	1	0
9	G	10	0	10	0	0
10	N	9	0	0	0	0
10	O	9	0	0	0	0
10	P	9	0	0	0	0
10	Q	9	0	0	0	0
11	A	182	0	0	2	0
11	B	120	0	0	3	0
11	C	124	0	0	0	0
11	D	50	0	0	0	0
11	E	179	0	0	0	0
11	F	108	0	0	2	0
11	G	95	0	0	1	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	H	56	0	0	3	0
11	N	5	0	0	0	0
11	O	7	0	0	0	0
11	Q	2	0	0	1	0
All	All	17374	0	15770	125	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 4.

The worst 5 of 125 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:G:54:LEU:HG	1:G:54(A):LYS:HE3	1.61	0.80
2:B:82:LYS:NZ	11:B:301:HOH:O	2.25	0.68
3:N:7:PHE:HA	8:N:103:PEG:H31	1.76	0.67
2:D:28:ASN:ND2	2:D:146:ASN:OD1	2.28	0.66
2:B:143:LYS:NZ	11:B:302:HOH:O	2.27	0.66

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	320/326 (98%)	312 (98%)	8 (2%)	0	100	100
1	C	320/326 (98%)	311 (97%)	8 (2%)	1 (0%)	37	43
1	E	320/326 (98%)	313 (98%)	7 (2%)	0	100	100
1	G	320/326 (98%)	312 (98%)	7 (2%)	1 (0%)	37	43
2	B	169/176 (96%)	166 (98%)	3 (2%)	0	100	100
2	D	169/176 (96%)	163 (96%)	6 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	169/176 (96%)	166 (98%)	3 (2%)	0	100	100
2	H	169/176 (96%)	163 (96%)	6 (4%)	0	100	100
3	N	11/13 (85%)	9 (82%)	2 (18%)	0	100	100
3	O	11/13 (85%)	9 (82%)	2 (18%)	0	100	100
3	P	11/13 (85%)	9 (82%)	1 (9%)	1 (9%)	0	0
3	Q	11/13 (85%)	10 (91%)	1 (9%)	0	100	100
All	All	2000/2060 (97%)	1943 (97%)	54 (3%)	3 (0%)	44	52

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	P	12	CYS
1	G	93	SER
1	C	125(C)	SER

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	285/289 (99%)	285 (100%)	0	100	100
1	C	282/289 (98%)	282 (100%)	0	100	100
1	E	284/289 (98%)	284 (100%)	0	100	100
1	G	284/289 (98%)	284 (100%)	0	100	100
2	B	146/151 (97%)	146 (100%)	0	100	100
2	D	147/151 (97%)	147 (100%)	0	100	100
2	F	147/151 (97%)	147 (100%)	0	100	100
2	H	146/151 (97%)	146 (100%)	0	100	100
3	N	11/11 (100%)	11 (100%)	0	100	100
3	O	11/11 (100%)	11 (100%)	0	100	100
3	P	11/11 (100%)	10 (91%)	1 (9%)	7	7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	Q	11/11 (100%)	10 (91%)	1 (9%)	7 7
All	All	1765/1804 (98%)	1763 (100%)	2 (0%)	92 96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	P	12	CYS
3	Q	12	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	81	ASN
1	G	197	ASN
1	E	129	ASN
2	H	25	HIS
2	F	38	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](#)

13 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
5	NAG	I	1	5,1	14,14,15	0.36	0	17,19,21	0.68	1 (5%)
5	FUC	I	2	5	10,10,11	0.79	0	14,14,16	0.93	0
5	NAG	I	3	5	14,14,15	0.30	0	17,19,21	0.34	0
4	NAG	J	1	1,4	14,14,15	0.75	1 (7%)	17,19,21	0.57	0
4	NAG	J	2	4	14,14,15	0.18	0	17,19,21	0.57	0
4	FUC	J	3	4	10,10,11	0.78	0	14,14,16	0.85	0
6	NAG	K	1	6,1	14,14,15	0.53	0	17,19,21	0.57	0
6	NAG	K	2	6	14,14,15	0.25	0	17,19,21	0.55	0
5	NAG	L	1	5,1	14,14,15	0.33	0	17,19,21	0.69	1 (5%)
5	FUC	L	2	5	10,10,11	0.56	0	14,14,16	0.73	0
5	NAG	L	3	5	14,14,15	0.40	0	17,19,21	0.43	0
6	NAG	M	1	6,2	14,14,15	0.28	0	17,19,21	0.61	0
6	NAG	M	2	6	14,14,15	0.36	0	17,19,21	0.41	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
5	NAG	I	1	5,1	-	2/6/23/26	0/1/1/1
5	FUC	I	2	5	-	-	0/1/1/1
5	NAG	I	3	5	-	0/6/23/26	0/1/1/1
4	NAG	J	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	J	2	4	-	0/6/23/26	0/1/1/1
4	FUC	J	3	4	-	-	0/1/1/1
6	NAG	K	1	6,1	-	2/6/23/26	0/1/1/1
6	NAG	K	2	6	-	0/6/23/26	0/1/1/1
5	NAG	L	1	5,1	-	1/6/23/26	0/1/1/1
5	FUC	L	2	5	-	-	0/1/1/1
5	NAG	L	3	5	-	0/6/23/26	0/1/1/1
6	NAG	M	1	6,2	-	2/6/23/26	0/1/1/1
6	NAG	M	2	6	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	J	1	NAG	C1-C2	2.57	1.55	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	I	1	NAG	C1-O5-C5	2.26	115.21	112.19
5	L	1	NAG	C1-O5-C5	2.25	115.20	112.19

There are no chirality outliers.

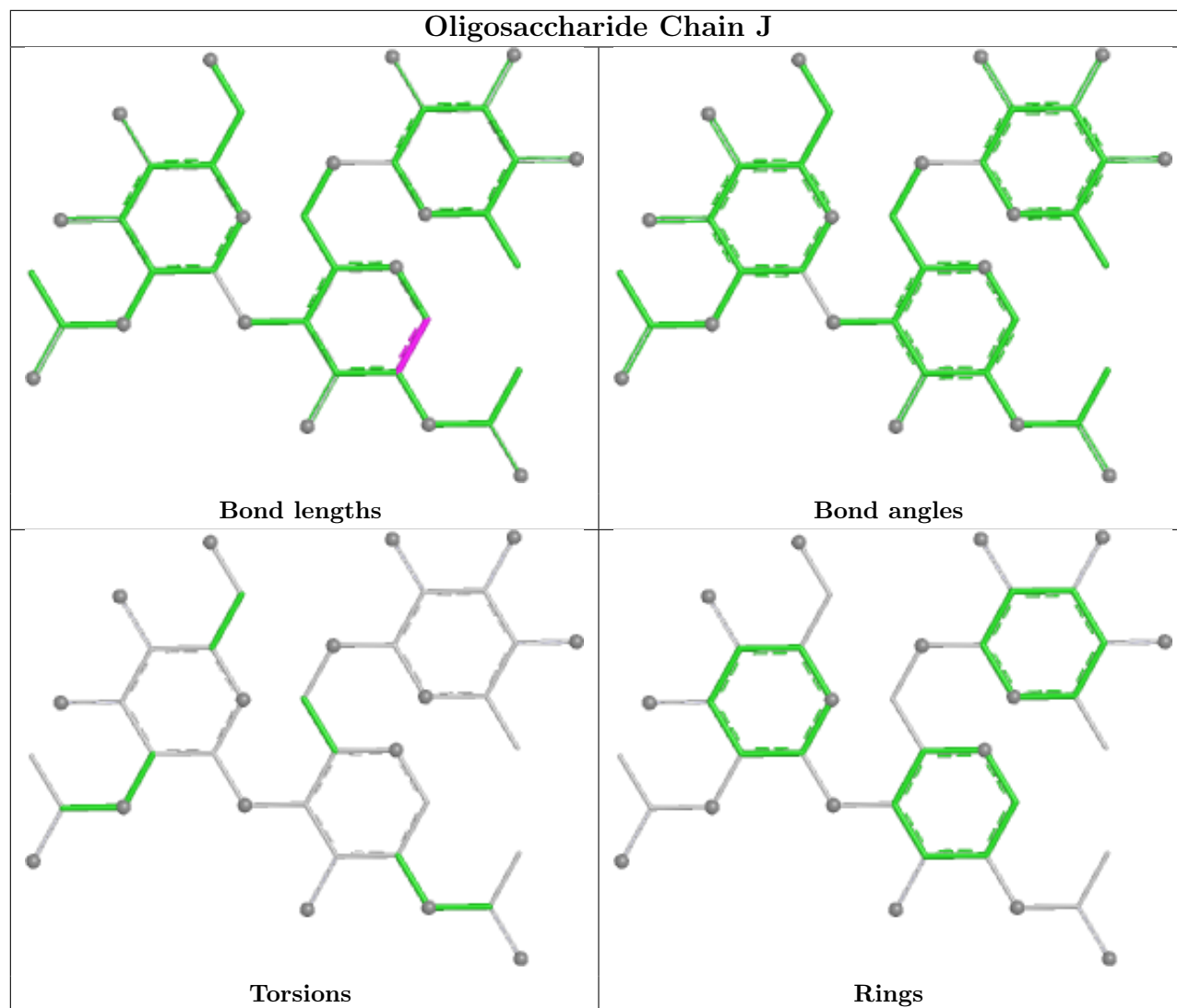
5 of 7 torsion outliers are listed below:

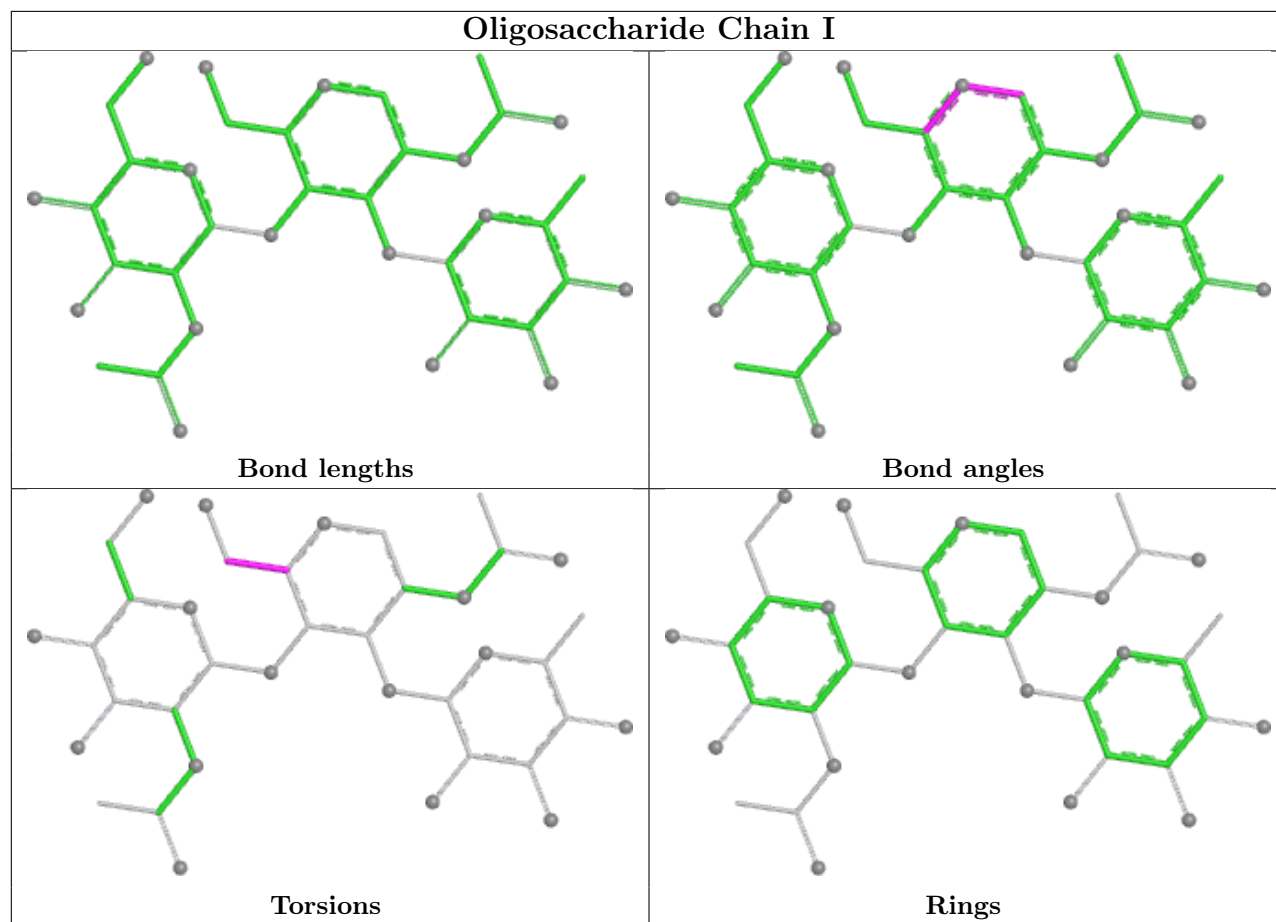
Mol	Chain	Res	Type	Atoms
5	I	1	NAG	C4-C5-C6-O6
5	I	1	NAG	O5-C5-C6-O6
6	M	1	NAG	C4-C5-C6-O6
6	K	1	NAG	C4-C5-C6-O6
6	M	1	NAG	O5-C5-C6-O6

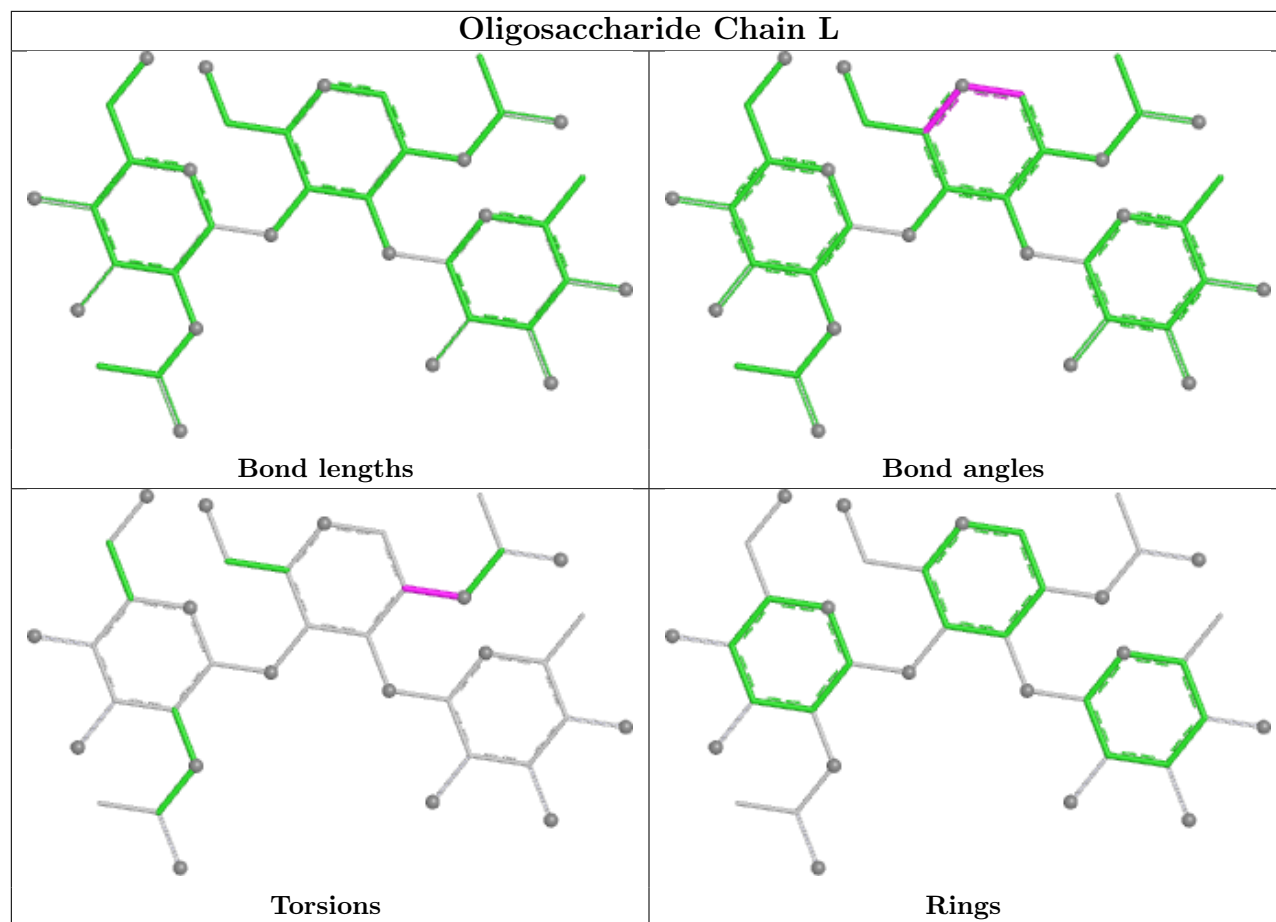
There are no ring outliers.

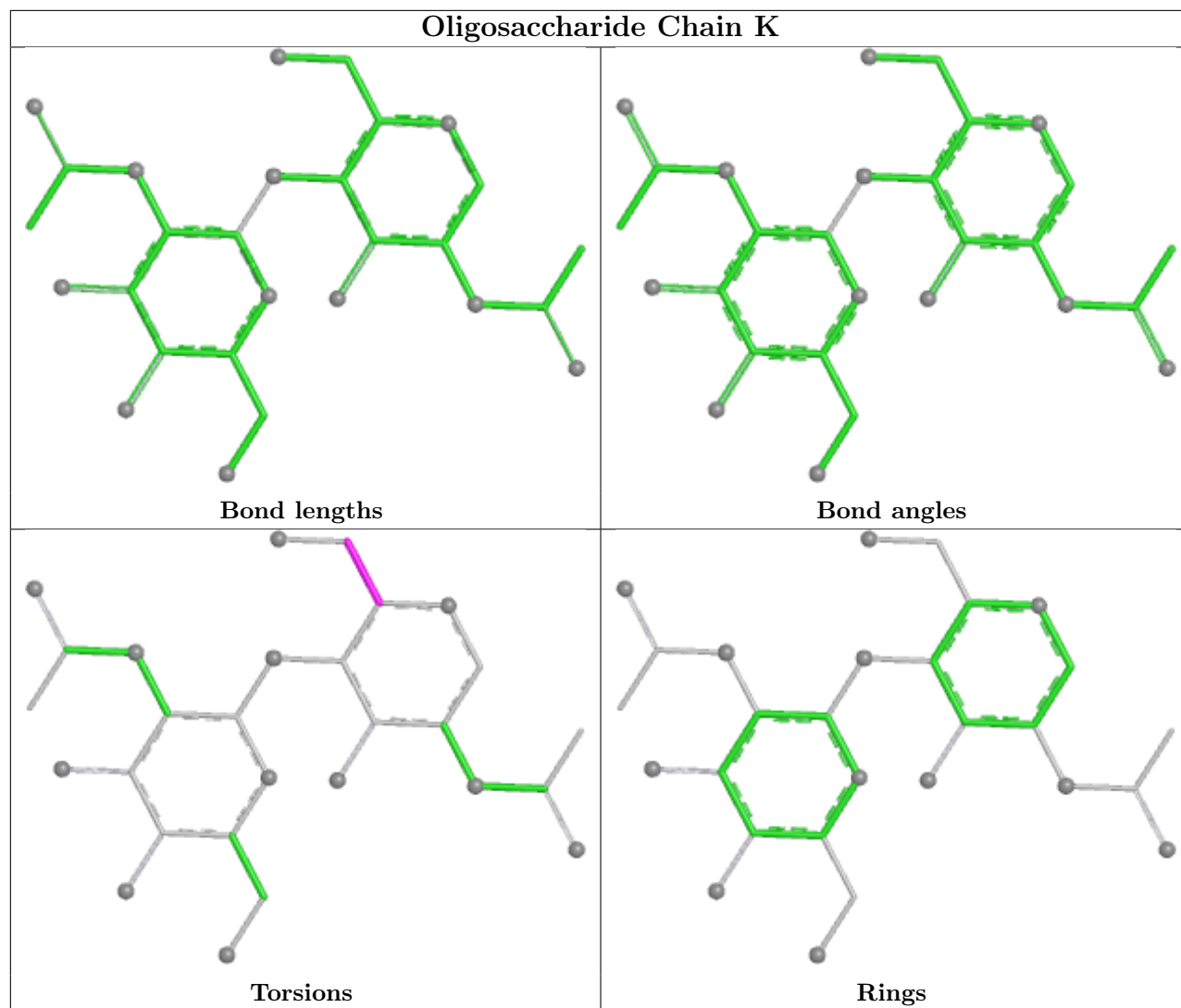
No monomer is involved in short contacts.

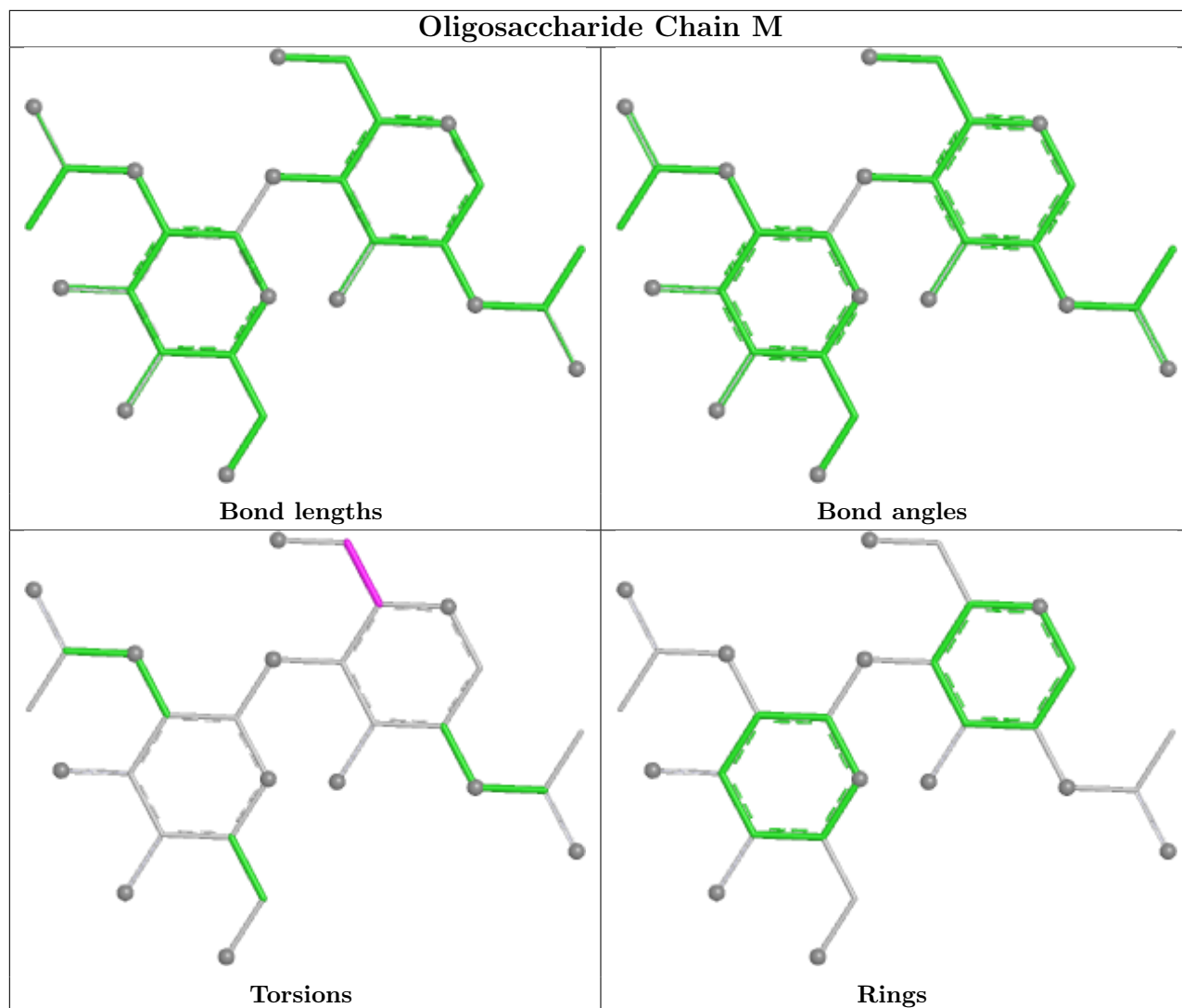
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











## 5.6 Ligand geometry [i](#)

22 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
7	NAG	C	401	1	14,14,15	0.50	0	17,19,21	0.67	1 (5%)
8	PEG	F	202	-	6,6,6	0.48	0	5,5,5	0.25	0
7	NAG	D	201	2	14,14,15	0.28	0	17,19,21	0.51	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	NAG	E	401	1	14,14,15	0.40	0	17,19,21	0.46	0
10	SEZ	P	101	3	9,9,9	0.38	0	12,12,12	1.09	0
8	PEG	E	403	-	6,6,6	0.48	0	5,5,5	0.22	0
8	PEG	N	103	-	6,6,6	0.49	0	5,5,5	0.20	0
8	PEG	B	201	-	6,6,6	0.47	0	5,5,5	0.28	0
9	FUC	G	402	-	10,10,11	0.88	1 (10%)	14,14,16	1.52	3 (21%)
7	NAG	G	401	1	14,14,15	0.42	0	17,19,21	0.61	0
10	SEZ	O	101	3	9,9,9	0.55	0	12,12,12	0.81	0
10	SEZ	N	101	3	9,9,9	0.45	0	12,12,12	0.90	0
7	NAG	H	201	2	14,14,15	0.27	0	17,19,21	0.51	0
8	PEG	N	102	-	6,6,6	0.50	0	5,5,5	0.23	0
8	PEG	G	403	-	6,6,6	0.49	0	5,5,5	0.21	0
7	NAG	F	201	2	14,14,15	0.27	0	17,19,21	0.58	0
8	PEG	A	403	-	6,6,6	0.49	0	5,5,5	0.24	0
7	NAG	A	401	1	14,14,15	0.27	0	17,19,21	0.54	0
8	PEG	C	402	-	6,6,6	0.50	0	5,5,5	0.23	0
8	PEG	A	402	-	6,6,6	0.49	0	5,5,5	0.29	0
10	SEZ	Q	101	3	9,9,9	0.36	0	12,12,12	1.16	1 (8%)
8	PEG	E	402	-	6,6,6	0.50	0	5,5,5	0.24	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
7	NAG	C	401	1	-	2/6/23/26	0/1/1/1
8	PEG	F	202	-	-	1/4/4/4	-
7	NAG	D	201	2	-	2/6/23/26	0/1/1/1
7	NAG	E	401	1	-	2/6/23/26	0/1/1/1
10	SEZ	P	101	3	-	-	0/1/1/1
8	PEG	E	403	-	-	2/4/4/4	-
8	PEG	N	103	-	-	2/4/4/4	-
8	PEG	B	201	-	-	2/4/4/4	-
9	FUC	G	402	-	-	-	0/1/1/1
7	NAG	G	401	1	-	2/6/23/26	0/1/1/1
10	SEZ	O	101	3	-	-	0/1/1/1
10	SEZ	N	101	3	-	-	0/1/1/1
7	NAG	H	201	2	-	2/6/23/26	0/1/1/1
8	PEG	N	102	-	-	2/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	PEG	G	403	-	-	2/4/4/4	-
7	NAG	F	201	2	-	2/6/23/26	0/1/1/1
8	PEG	A	403	-	-	1/4/4/4	-
7	NAG	A	401	1	-	0/6/23/26	0/1/1/1
8	PEG	C	402	-	-	3/4/4/4	-
8	PEG	A	402	-	-	2/4/4/4	-
10	SEZ	Q	101	3	-	-	0/1/1/1
8	PEG	E	402	-	-	2/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	G	402	FUC	C1-C2	2.30	1.57	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	G	402	FUC	C1-C2-C3	3.28	114.42	109.64
9	G	402	FUC	C1-O5-C5	2.84	119.68	112.97
9	G	402	FUC	O5-C1-C2	2.66	117.13	110.79
10	Q	101	SEZ	CH-CT2-CZ2	2.13	120.48	118.11
7	C	401	NAG	C1-O5-C5	2.11	115.01	112.19

There are no chirality outliers.

5 of 31 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	D	201	NAG	C4-C5-C6-O6
7	H	201	NAG	C4-C5-C6-O6
8	B	201	PEG	O1-C1-C2-O2
8	E	403	PEG	O2-C3-C4-O4
8	F	202	PEG	O2-C3-C4-O4

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	N	103	PEG	1	0
8	B	201	PEG	3	0
8	A	403	PEG	1	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	C	402	PEG	1	0
8	A	402	PEG	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, 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	322/326 (98%)	-1.68	0 100 100	13, 23, 39, 57	0
1	C	322/326 (98%)	-1.57	0 100 100	17, 32, 50, 68	0
1	E	322/326 (98%)	-1.66	0 100 100	14, 23, 39, 55	0
1	G	322/326 (98%)	-1.57	0 100 100	16, 32, 50, 67	0
2	B	171/176 (97%)	-1.69	0 100 100	16, 23, 37, 62	0
2	D	171/176 (97%)	-1.47	0 100 100	20, 38, 63, 77	0
2	F	171/176 (97%)	-1.70	0 100 100	15, 23, 37, 53	0
2	H	171/176 (97%)	-1.48	0 100 100	20, 38, 66, 80	0
3	N	11/13 (84%)	-1.59	0 100 100	21, 27, 47, 57	0
3	O	11/13 (84%)	-1.57	0 100 100	22, 28, 47, 58	0
3	P	11/13 (84%)	-0.96	0 100 100	56, 66, 78, 80	0
3	Q	11/13 (84%)	-0.95	0 100 100	54, 64, 78, 78	0
All	All	2016/2060 (97%)	-1.60	0 100 100	13, 28, 56, 80	0

There are no RSRZ outliers to report.

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

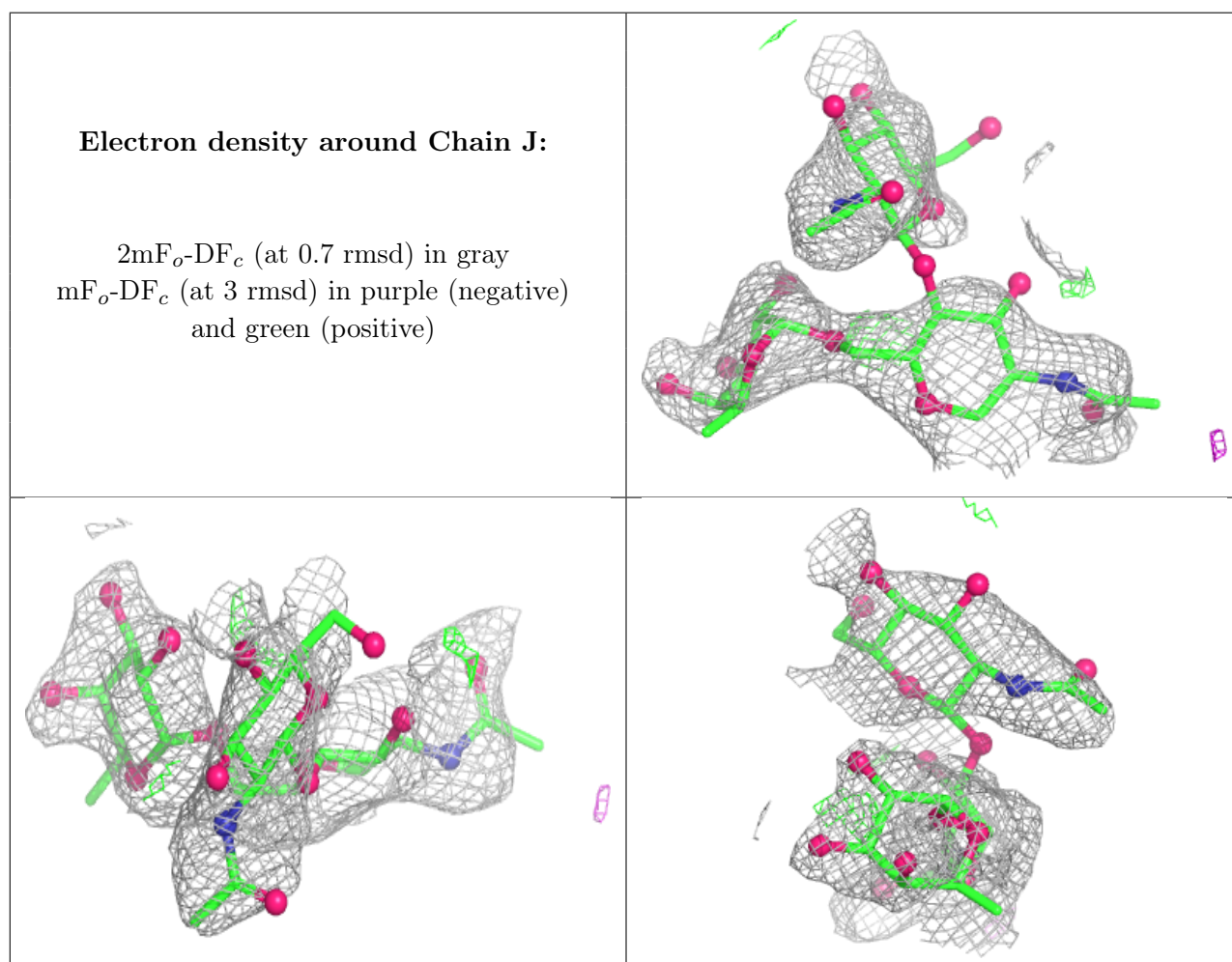
There are no non-standard protein/DNA/RNA residues in this entry.

### 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.

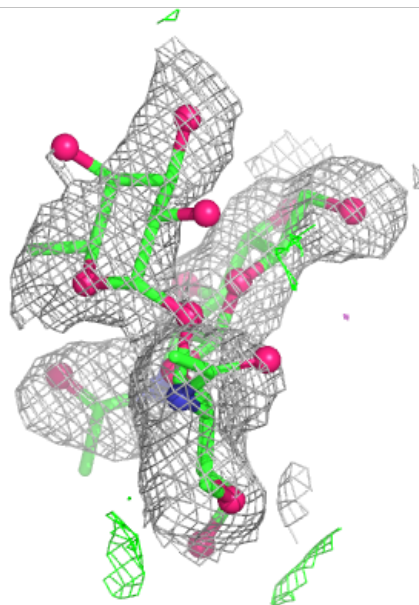
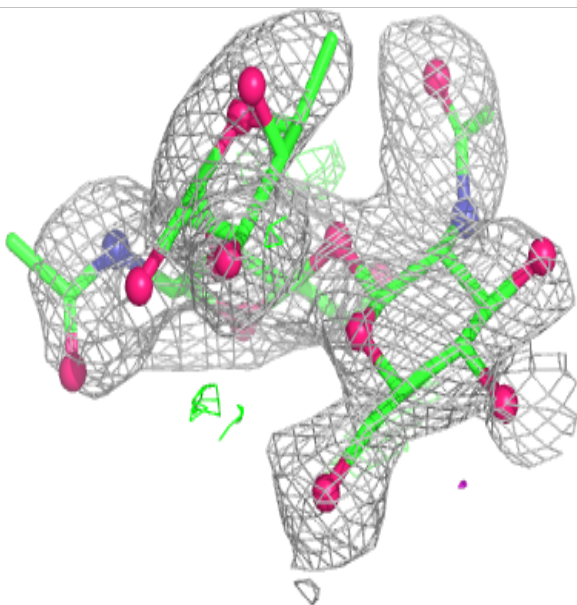
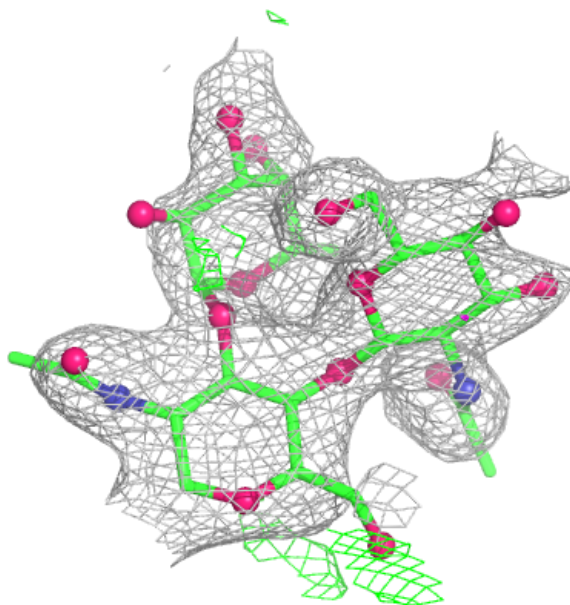
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	J	1	14/15	-	-	68,75,83,94	0
4	NAG	J	2	14/15	-	-	72,86,93,104	0
4	FUC	J	3	10/11	-	-	71,76,79,80	0
5	NAG	I	1	14/15	-	-	45,58,70,77	0
5	FUC	I	2	10/11	-	-	52,58,65,65	0
5	NAG	I	3	14/15	-	-	37,66,70,71	0
5	NAG	L	1	14/15	-	-	45,59,67,76	0
5	FUC	L	2	10/11	-	-	53,58,62,63	0
5	NAG	L	3	14/15	-	-	40,65,68,70	0
6	NAG	K	1	14/15	-	-	60,70,77,85	0
6	NAG	K	2	14/15	-	-	71,80,86,94	0
6	NAG	M	2	14/15	0.97	0.06	56,69,74,75	0
6	NAG	M	1	14/15	0.98	0.06	44,54,61,64	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.



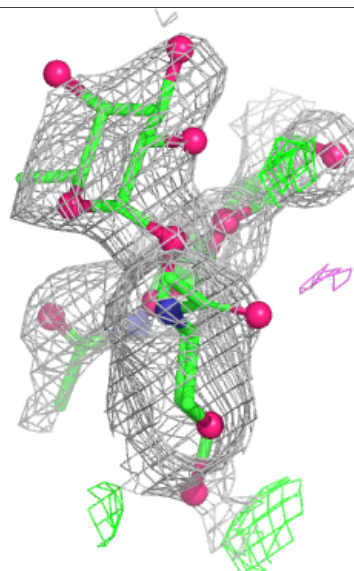
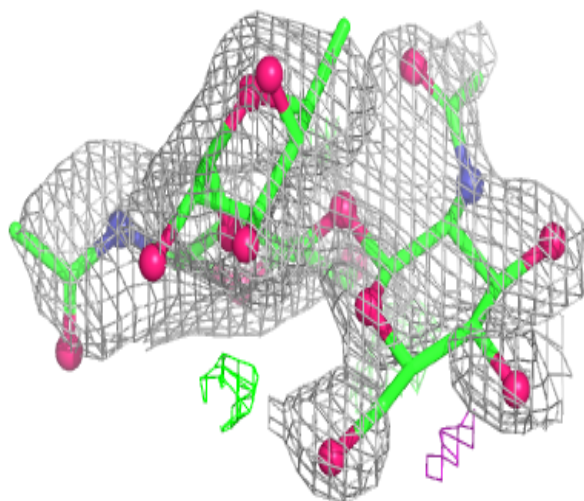
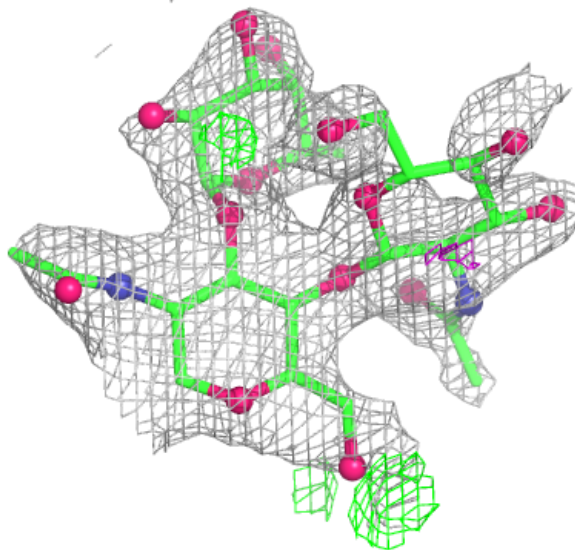
**Electron density around Chain I:**

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



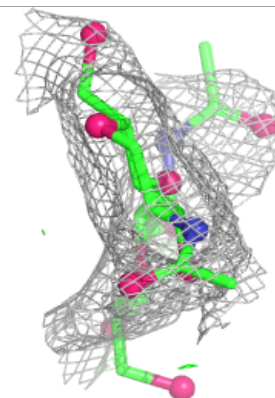
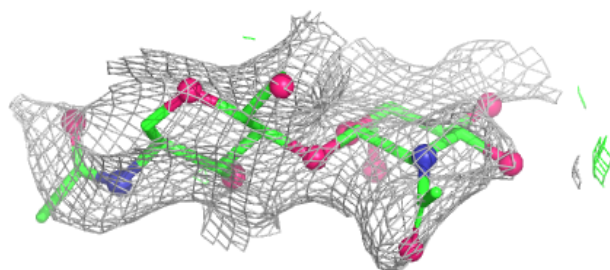
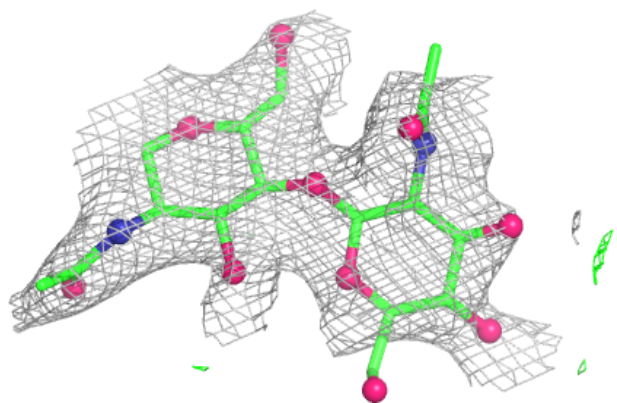
**Electron density around Chain L:**

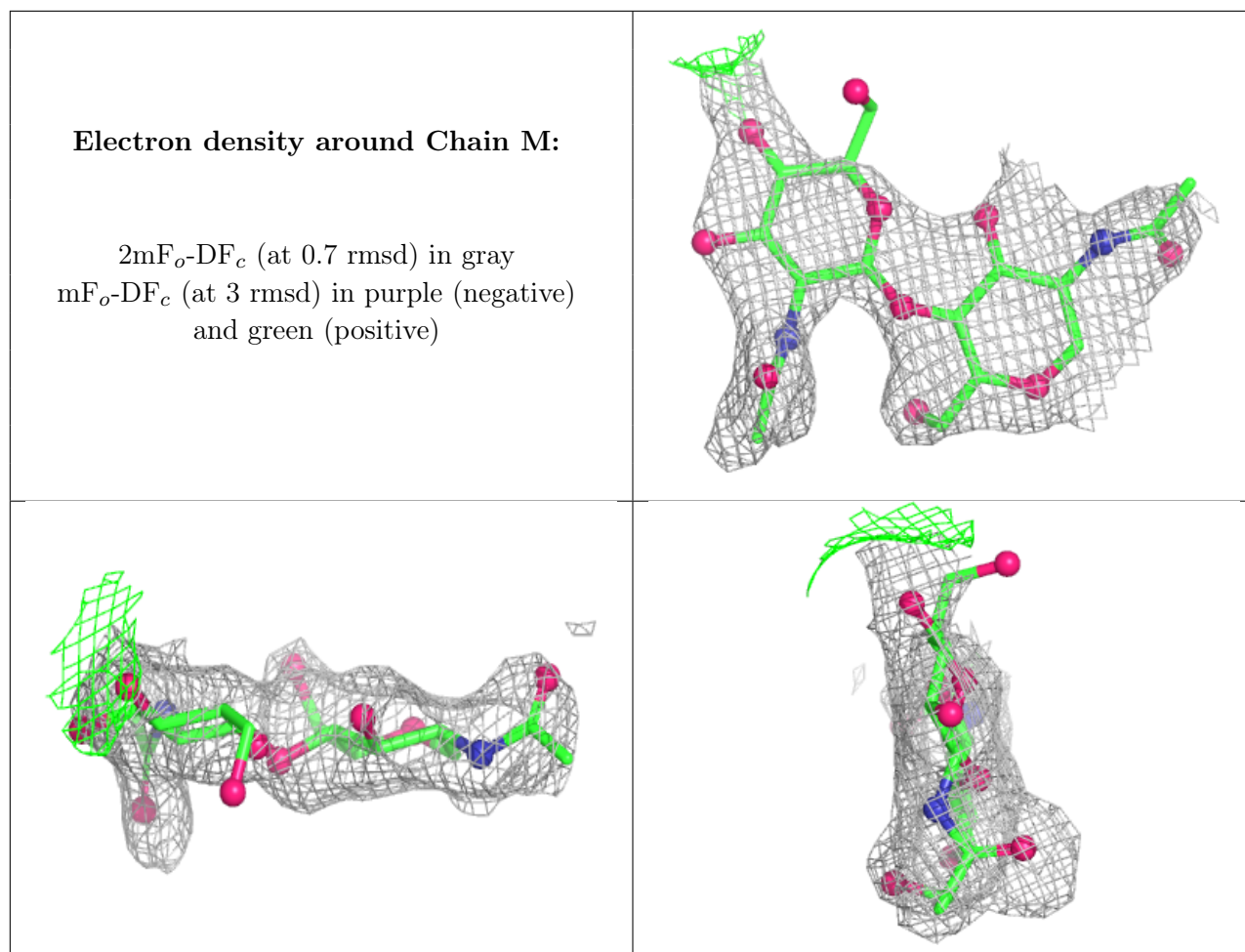
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain 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, 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
7	NAG	G	401	14/15	0.93	0.07	75,82,88,90	0
7	NAG	C	401	14/15	0.96	0.07	80,87,94,100	0
7	NAG	E	401	14/15	0.97	0.06	47,65,72,74	0
7	NAG	H	201	14/15	0.97	0.07	69,83,87,89	0
7	NAG	F	201	14/15	0.98	0.04	39,56,67,68	0
7	NAG	D	201	14/15	0.98	0.06	70,78,84,86	0
7	NAG	A	401	14/15	0.98	0.05	62,70,75,79	0
8	PEG	E	402	7/7	0.98	0.04	31,36,47,53	0
8	PEG	N	102	7/7	0.98	0.08	43,46,51,56	0
9	FUC	G	402	10/11	0.98	0.09	74,82,88,91	0
10	SEZ	P	101	9/9	0.98	0.08	75,81,85,85	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
10	SEZ	Q	101	9/9	0.98	0.09	69,82,84,87	0
8	PEG	E	403	7/7	0.99	0.04	25,34,42,44	0
8	PEG	F	202	7/7	0.99	0.04	22,30,39,47	0
8	PEG	G	403	7/7	0.99	0.05	42,45,51,55	0
8	PEG	A	403	7/7	0.99	0.05	23,37,41,42	0
8	PEG	N	103	7/7	0.99	0.06	36,39,46,59	0
8	PEG	B	201	7/7	0.99	0.04	24,24,40,47	0
10	SEZ	N	101	9/9	0.99	0.03	30,35,40,40	0
10	SEZ	O	101	9/9	0.99	0.04	26,38,42,46	0
8	PEG	C	402	7/7	0.99	0.05	36,40,48,48	0
8	PEG	A	402	7/7	0.99	0.04	28,35,42,45	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.