



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 7, 2020 – 09:01 AM BST

PDB ID : 5HUF  
Title : The crystal structure of hemagglutinin from A/gyrfalcon/Washington/41088-6/2014 influenza virus  
Authors : Yang, H.; Carney, P.J.; Guo, Z.; Chang, J.C.; Stevens, J.  
Deposited on : 2016-01-27  
Resolution : 2.81 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 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

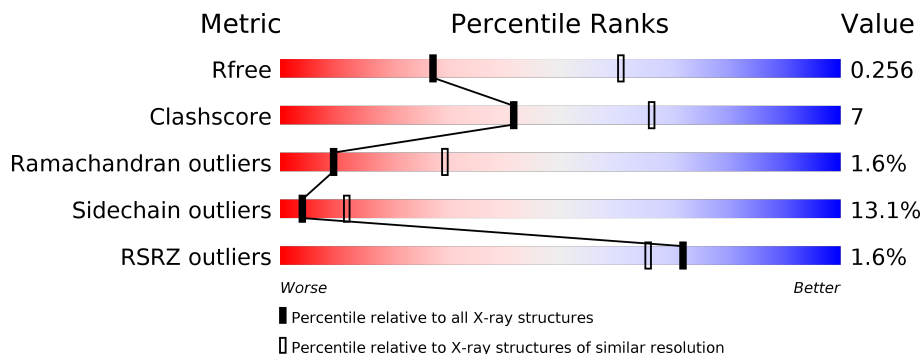
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.81 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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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	334	 % 74% 19% . . .
1	C	334	 2% 73% 20% . . .
1	E	334	 % 74% 20% . . .
2	B	181	 73% 19% . . .
2	D	181	 3% 72% 19% . . .
2	F	181	 3% 71% 19% 5% . .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	3	 100%
3	I	3	 100%
3	K	3	 100%
4	H	2	 100%
4	J	2	 50% 50%
4	L	2	 100%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	MAN	G	3	X	-	-	-
3	MAN	I	3	X	-	-	-
3	MAN	K	3	X	-	-	X
4	NAG	J	2	-	-	-	X
5	NAG	E	505	-	-	-	X

## 2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 12219 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 HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	324	2559	1618	445	481	15	0	0	0
1	C	324	2559	1618	445	481	15	0	0	0
1	E	324	2559	1618	445	481	15	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	ALA	-	expression tag	UNP A0A0C4X0C0
A	-3	ASP	-	expression tag	UNP A0A0C4X0C0
A	-2	LEU	-	expression tag	UNP A0A0C4X0C0
A	-1	GLY	-	expression tag	UNP A0A0C4X0C0
C	-4	ALA	-	expression tag	UNP A0A0C4X0C0
C	-3	ASP	-	expression tag	UNP A0A0C4X0C0
C	-2	LEU	-	expression tag	UNP A0A0C4X0C0
C	-1	GLY	-	expression tag	UNP A0A0C4X0C0
E	-4	ALA	-	expression tag	UNP A0A0C4X0C0
E	-3	ASP	-	expression tag	UNP A0A0C4X0C0
E	-2	LEU	-	expression tag	UNP A0A0C4X0C0
E	-1	GLY	-	expression tag	UNP A0A0C4X0C0

- Molecule 2 is a protein called hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	176	1419	884	244	283	8	0	0	0
2	D	176	1419	884	244	283	8	0	0	0
2	F	176	1419	884	244	283	8	0	0	0

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	175	SER	-	expression tag	UNP A0A0C4X0C0
B	176	GLY	-	expression tag	UNP A0A0C4X0C0
B	177	ARG	-	expression tag	UNP A0A0C4X0C0
B	178	LEU	-	expression tag	UNP A0A0C4X0C0
B	179	VAL	-	expression tag	UNP A0A0C4X0C0
B	180	PRO	-	expression tag	UNP A0A0C4X0C0
B	181	ARG	-	expression tag	UNP A0A0C4X0C0
D	175	SER	-	expression tag	UNP A0A0C4X0C0
D	176	GLY	-	expression tag	UNP A0A0C4X0C0
D	177	ARG	-	expression tag	UNP A0A0C4X0C0
D	178	LEU	-	expression tag	UNP A0A0C4X0C0
D	179	VAL	-	expression tag	UNP A0A0C4X0C0
D	180	PRO	-	expression tag	UNP A0A0C4X0C0
D	181	ARG	-	expression tag	UNP A0A0C4X0C0
F	175	SER	-	expression tag	UNP A0A0C4X0C0
F	176	GLY	-	expression tag	UNP A0A0C4X0C0
F	177	ARG	-	expression tag	UNP A0A0C4X0C0
F	178	LEU	-	expression tag	UNP A0A0C4X0C0
F	179	VAL	-	expression tag	UNP A0A0C4X0C0
F	180	PRO	-	expression tag	UNP A0A0C4X0C0
F	181	ARG	-	expression tag	UNP A0A0C4X0C0

- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



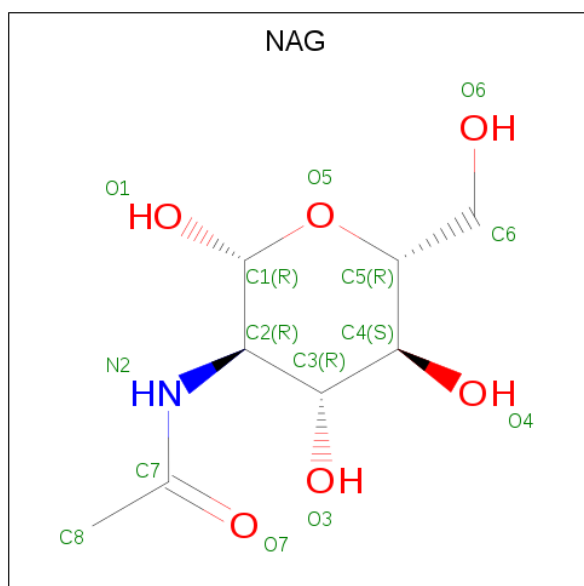
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	G	3	39	22	2	15	0	0	0
3	I	3	39	22	2	15	0	0	0
3	K	3	39	22	2	15	0	0	0

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



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

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

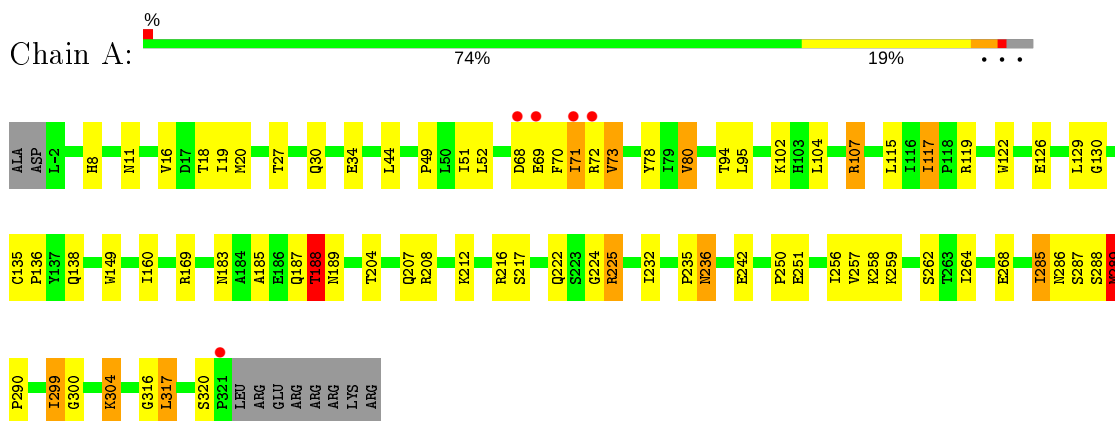


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	C	1	Total	C	N	O	0	0
			14	8	1	5		
5	C	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		

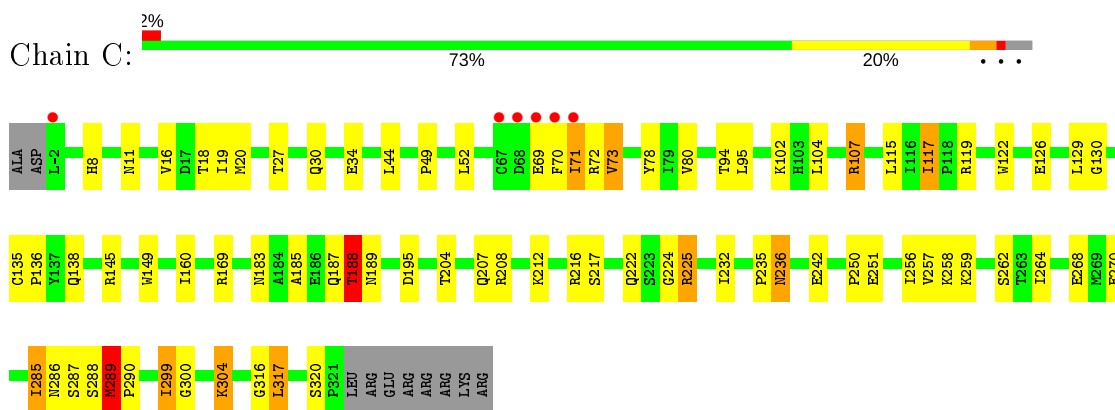
### 3 Residue-property plots [\(i\)](#)

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.

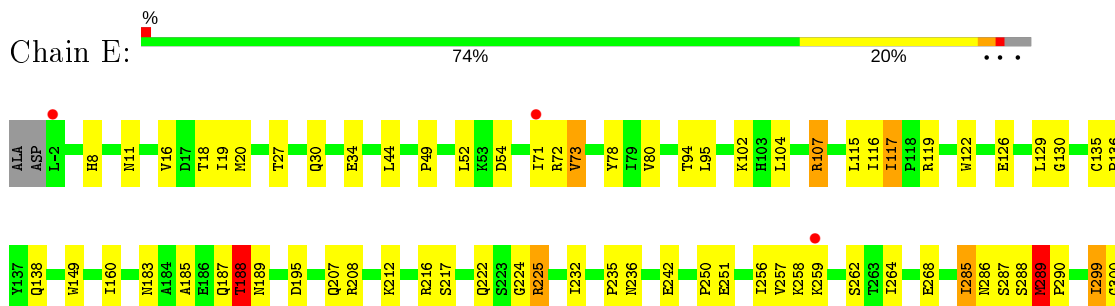
- Molecule 1: hemagglutinin HA1



- Molecule 1: hemagglutinin HA1



- Molecule 1: hemagglutinin HA1





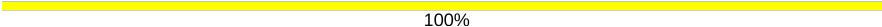


- Molecule 3: alpha-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K:  100%

MAG1  
MAG2  
MAG3

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

Chain H:  100%

MAG1  
MAG2

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

Chain J:  50%

MAG1  
MAG2

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

Chain L:  100%

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	125.33Å 252.38Å 70.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.66 – 2.81 49.66 – 2.81	Depositor EDS
% Data completeness (in resolution range)	96.0 (49.66-2.81) 96.1 (49.66-2.81)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.90 (at 2.81Å)	Xtrriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.228 , 0.256 0.228 , 0.256	Depositor DCC
$R_{free}$ test set	2723 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	57.2	Xtrriage
Anisotropy	0.236	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 20.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	12219	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.83% 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: NAG, MAN

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	0.57	0/2622	0.81	1/3565 (0.0%)
1	C	0.61	0/2622	0.83	2/3565 (0.1%)
1	E	0.62	0/2622	0.83	1/3565 (0.0%)
2	B	0.70	1/1446 (0.1%)	0.83	1/1943 (0.1%)
2	D	0.63	1/1446 (0.1%)	0.82	2/1943 (0.1%)
2	F	0.63	0/1446	0.81	2/1943 (0.1%)
All	All	0.62	2/12204 (0.0%)	0.82	9/16524 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	4
2	D	0	4
2	F	0	4
All	All	0	12

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	64	GLU	CG-CD	5.52	1.60	1.51
2	D	64	GLU	CG-CD	5.22	1.59	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	289	MET	CG-SD-CE	6.28	110.24	100.20
1	C	289	MET	CG-SD-CE	5.57	109.10	100.20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	289	MET	CG-SD-CE	5.53	109.05	100.20
2	F	65	ALA	N-CA-C	5.36	125.48	111.00
2	B	65	ALA	N-CA-C	5.30	125.30	111.00

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	5	ALA	Peptide
2	B	61	THR	Peptide
2	B	64	GLU	Peptide
2	B	67	GLY	Peptide
2	D	5	ALA	Peptide

## 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	2559	0	2512	37	0
1	C	2559	0	2512	33	0
1	E	2559	0	2512	37	0
2	B	1419	0	1329	21	0
2	D	1419	0	1329	24	0
2	F	1419	0	1329	27	0
3	G	39	0	34	0	0
3	I	39	0	34	0	0
3	K	39	0	34	0	0
4	H	28	0	25	0	0
4	J	28	0	25	1	0
4	L	28	0	25	0	0
5	A	28	0	26	0	0
5	C	28	0	26	0	0
5	E	28	0	26	0	0
All	All	12219	0	11778	167	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 7.

The worst 5 of 167 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:A:117:ILE:O	1:A:117:ILE:HD13	1.74	0.87
1:C:117:ILE:O	1:C:117:ILE:HD13	1.74	0.87
1:E:117:ILE:HD13	1:E:117:ILE:O	1.75	0.87
2:B:59:MET:O	2:B:61:THR:N	2.10	0.84
2:D:59:MET:O	2:D:61:THR:N	2.10	0.83

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	322/334 (96%)	287 (89%)	31 (10%)	4 (1%)	13	37
1	C	322/334 (96%)	289 (90%)	29 (9%)	4 (1%)	13	37
1	E	322/334 (96%)	287 (89%)	32 (10%)	3 (1%)	17	44
2	B	174/181 (96%)	161 (92%)	8 (5%)	5 (3%)	4	15
2	D	174/181 (96%)	160 (92%)	10 (6%)	4 (2%)	6	20
2	F	174/181 (96%)	159 (91%)	11 (6%)	4 (2%)	6	20
All	All	1488/1545 (96%)	1343 (90%)	121 (8%)	24 (2%)	9	29

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	71	ILE
2	B	60	ASN
1	C	71	ILE
2	D	60	ASN
1	E	71	ILE

### 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	288/297 (97%)	252 (88%)	36 (12%)	4	14
1	C	288/297 (97%)	250 (87%)	38 (13%)	4	11
1	E	288/297 (97%)	253 (88%)	35 (12%)	5	14
2	B	150/155 (97%)	130 (87%)	20 (13%)	4	11
2	D	150/155 (97%)	130 (87%)	20 (13%)	4	11
2	F	150/155 (97%)	127 (85%)	23 (15%)	2	8
All	All	1314/1356 (97%)	1142 (87%)	172 (13%)	4	12

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	217	SER
2	D	29	GLU
2	F	66	VAL
1	C	236	ASN
1	C	287	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	222	GLN
1	C	236	ASN
1	E	183	ASN
1	C	183	ASN
1	E	222	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](#)

21 non-standard protein/DNA/RNA residues 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	2	3	14,14,15	0.76	0	17,19,21	1.32	2 (11%)
5	NAG	A	501	1	14,14,15	1.02	1 (7%)	17,19,21	1.95	5 (29%)
5	NAG	A	505	1	14,14,15	0.64	0	17,19,21	1.61	3 (17%)
4	NAG	J	2	4	14,14,15	0.59	0	17,19,21	2.26	3 (17%)
3	NAG	G	1	1,3	14,14,15	0.67	0	17,19,21	1.50	2 (11%)
4	NAG	H	2	4	14,14,15	0.73	0	17,19,21	1.90	3 (17%)
3	NAG	I	1	1,3	14,14,15	0.95	1 (7%)	17,19,21	1.93	3 (17%)
4	NAG	L	1	2,4	14,14,15	0.53	0	17,19,21	1.51	3 (17%)
3	MAN	G	3	3	11,11,12	0.75	0	15,15,17	1.73	3 (20%)
4	NAG	J	1	2,4	14,14,15	0.68	0	17,19,21	1.20	3 (17%)
4	NAG	H	1	2,4	14,14,15	1.02	1 (7%)	17,19,21	2.01	5 (29%)
3	NAG	K	1	1,3	14,14,15	0.64	0	17,19,21	2.59	7 (41%)
5	NAG	E	505	1	14,14,15	0.80	0	17,19,21	1.32	2 (11%)
4	NAG	L	2	4	14,14,15	0.70	0	17,19,21	1.61	5 (29%)
5	NAG	C	501	1	14,14,15	0.60	0	17,19,21	2.47	4 (23%)
3	MAN	K	3	3	11,11,12	0.90	0	15,15,17	1.69	3 (20%)
5	NAG	E	501	1	14,14,15	0.90	1 (7%)	17,19,21	1.54	2 (11%)
3	MAN	I	3	3	11,11,12	0.68	0	15,15,17	2.99	5 (33%)
5	NAG	C	505	1	14,14,15	0.72	0	17,19,21	1.85	5 (29%)
3	NAG	I	2	3	14,14,15	0.80	0	17,19,21	1.39	1 (5%)
3	NAG	K	2	3	14,14,15	0.74	0	17,19,21	1.52	3 (17%)

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	2	3	-	2/6/23/26	0/1/1/1
5	NAG	A	501	1	-	2/6/23/26	0/1/1/1
5	NAG	A	505	1	-	0/6/23/26	0/1/1/1
4	NAG	J	2	4	-	2/6/23/26	0/1/1/1
3	NAG	G	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	I	2	3	-	1/6/23/26	0/1/1/1
4	NAG	H	2	4	-	0/6/23/26	0/1/1/1
3	NAG	I	1	1,3	-	1/6/23/26	0/1/1/1
4	NAG	L	1	2,4	-	0/6/23/26	0/1/1/1
4	NAG	J	1	2,4	-	0/6/23/26	0/1/1/1
4	NAG	H	1	2,4	-	1/6/23/26	0/1/1/1
3	NAG	K	1	1,3	-	2/6/23/26	0/1/1/1
5	NAG	E	505	1	-	0/6/23/26	0/1/1/1
4	NAG	L	2	4	-	0/6/23/26	0/1/1/1
5	NAG	C	501	1	-	1/6/23/26	0/1/1/1
3	MAN	K	3	3	1/1/4/5	2/2/19/22	0/1/1/1
5	NAG	E	501	1	-	2/6/23/26	0/1/1/1
3	MAN	I	3	3	1/1/4/5	2/2/19/22	0/1/1/1
5	NAG	C	505	1	-	2/6/23/26	0/1/1/1
3	MAN	G	3	3	1/1/4/5	2/2/19/22	1/1/1/1
3	NAG	K	2	3	-	2/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	501	NAG	C1-C2	2.89	1.56	1.52
4	H	1	NAG	C1-C2	2.57	1.56	1.52
5	E	501	NAG	C1-C2	2.37	1.55	1.52
3	I	1	NAG	C1-C2	2.18	1.55	1.52

The worst 5 of 72 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	3	MAN	C1-O5-C5	8.90	124.25	112.19
5	C	501	NAG	C1-O5-C5	8.69	123.97	112.19
3	K	1	NAG	C1-O5-C5	7.49	122.34	112.19
4	J	2	NAG	C1-O5-C5	7.38	122.19	112.19
3	I	1	NAG	C1-O5-C5	6.02	120.35	112.19

All (3) chirality outliers are listed below:



Mol	Chain	Res	Type	Atom
3	G	3	MAN	C1
3	K	3	MAN	C1
3	I	3	MAN	C1

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	501	NAG	O5-C5-C6-O6
3	K	3	MAN	O5-C5-C6-O6
3	K	3	MAN	C4-C5-C6-O6
3	I	3	MAN	O5-C5-C6-O6
5	E	501	NAG	O5-C5-C6-O6

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	3	MAN	C1-C2-C3-C4-C5-O5

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	J	2	NAG	1	0

## 5.5 Carbohydrates [i](#)

15 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	0.67	0	17,19,21	1.50	2 (11%)
3	NAG	G	2	3	14,14,15	0.76	0	17,19,21	1.32	2 (11%)
3	MAN	G	3	3	11,11,12	0.75	0	15,15,17	1.73	3 (20%)
4	NAG	H	1	2,4	14,14,15	1.02	1 (7%)	17,19,21	2.01	5 (29%)
4	NAG	H	2	4	14,14,15	0.73	0	17,19,21	1.90	3 (17%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	I	1	1,3	14,14,15	0.95	1 (7%)	17,19,21	1.93	3 (17%)
3	NAG	I	2	3	14,14,15	0.80	0	17,19,21	1.39	1 (5%)
3	MAN	I	3	3	11,11,12	0.68	0	15,15,17	2.99	5 (33%)
4	NAG	J	1	2,4	14,14,15	0.68	0	17,19,21	1.20	3 (17%)
4	NAG	J	2	4	14,14,15	0.59	0	17,19,21	2.26	3 (17%)
3	NAG	K	1	1,3	14,14,15	0.64	0	17,19,21	2.59	7 (41%)
3	NAG	K	2	3	14,14,15	0.74	0	17,19,21	1.52	3 (17%)
3	MAN	K	3	3	11,11,12	0.90	0	15,15,17	1.69	3 (20%)
4	NAG	L	1	2,4	14,14,15	0.53	0	17,19,21	1.51	3 (17%)
4	NAG	L	2	4	14,14,15	0.70	0	17,19,21	1.61	5 (29%)

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	-	2/6/23/26	0/1/1/1
3	MAN	G	3	3	1/1/4/5	2/2/19/22	1/1/1/1
4	NAG	H	1	2,4	-	1/6/23/26	0/1/1/1
4	NAG	H	2	4	-	0/6/23/26	0/1/1/1
3	NAG	I	1	1,3	-	1/6/23/26	0/1/1/1
3	NAG	I	2	3	-	1/6/23/26	0/1/1/1
3	MAN	I	3	3	1/1/4/5	2/2/19/22	0/1/1/1
4	NAG	J	1	2,4	-	0/6/23/26	0/1/1/1
4	NAG	J	2	4	-	2/6/23/26	0/1/1/1
3	NAG	K	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	K	2	3	-	2/6/23/26	0/1/1/1
3	MAN	K	3	3	1/1/4/5	2/2/19/22	0/1/1/1
4	NAG	L	1	2,4	-	0/6/23/26	0/1/1/1
4	NAG	L	2	4	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	1	NAG	C1-C2	2.57	1.56	1.52
3	I	1	NAG	C1-C2	2.18	1.55	1.52

The worst 5 of 51 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	3	MAN	C1-O5-C5	8.90	124.25	112.19
3	K	1	NAG	C1-O5-C5	7.49	122.34	112.19
4	J	2	NAG	C1-O5-C5	7.38	122.19	112.19
3	I	1	NAG	C1-O5-C5	6.02	120.35	112.19
4	H	1	NAG	C1-O5-C5	5.37	119.47	112.19

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	K	3	MAN	C1
3	I	3	MAN	C1
3	G	3	MAN	C1

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	K	3	MAN	O5-C5-C6-O6
3	K	3	MAN	C4-C5-C6-O6
3	I	3	MAN	O5-C5-C6-O6
3	K	2	NAG	O5-C5-C6-O6
4	J	2	NAG	O5-C5-C6-O6

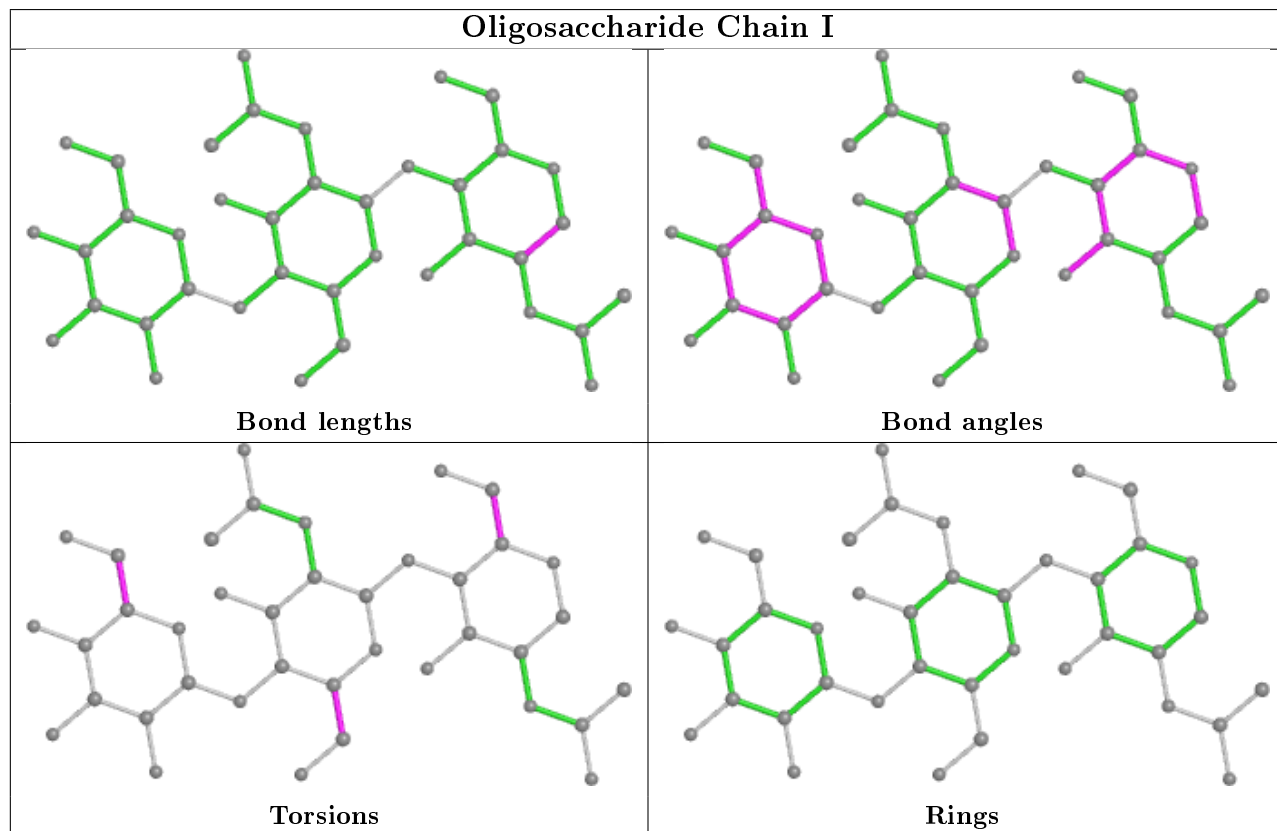
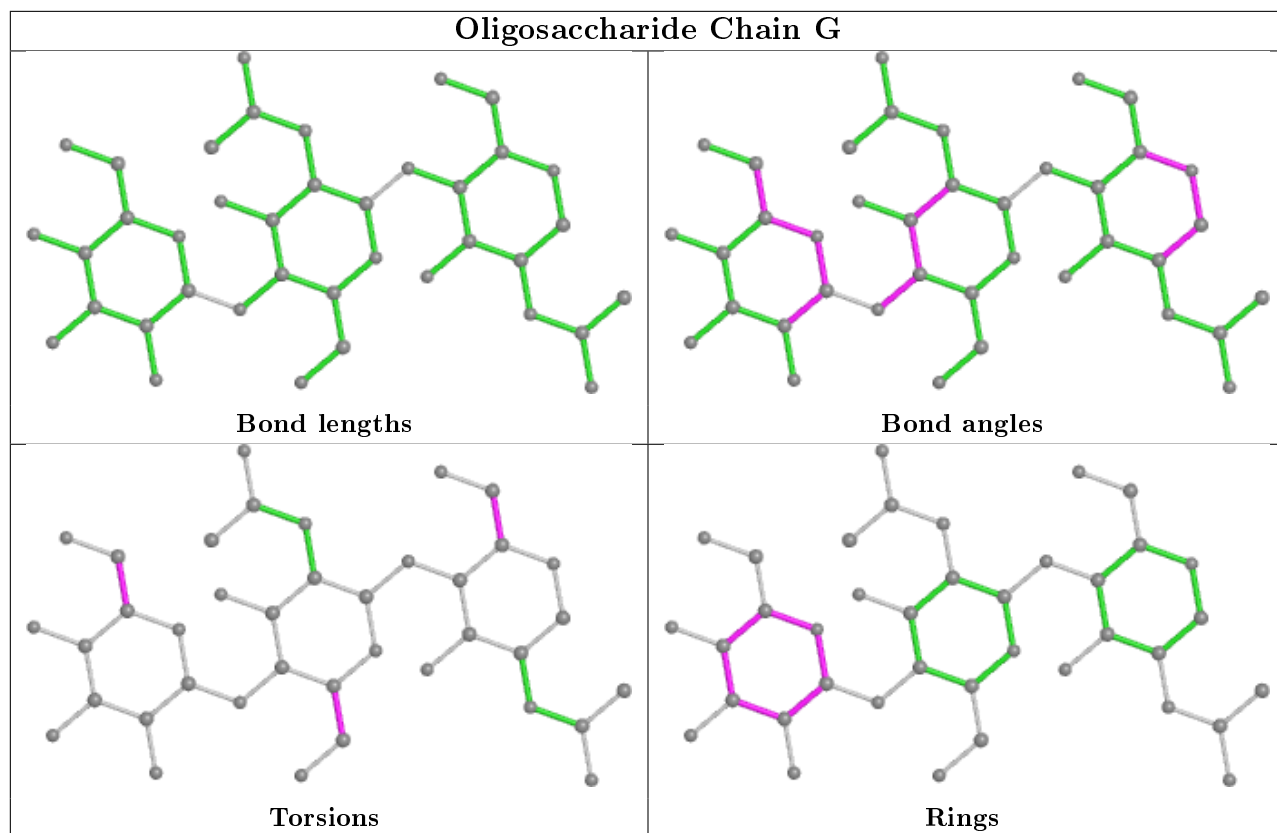
All (1) ring outliers are listed below:

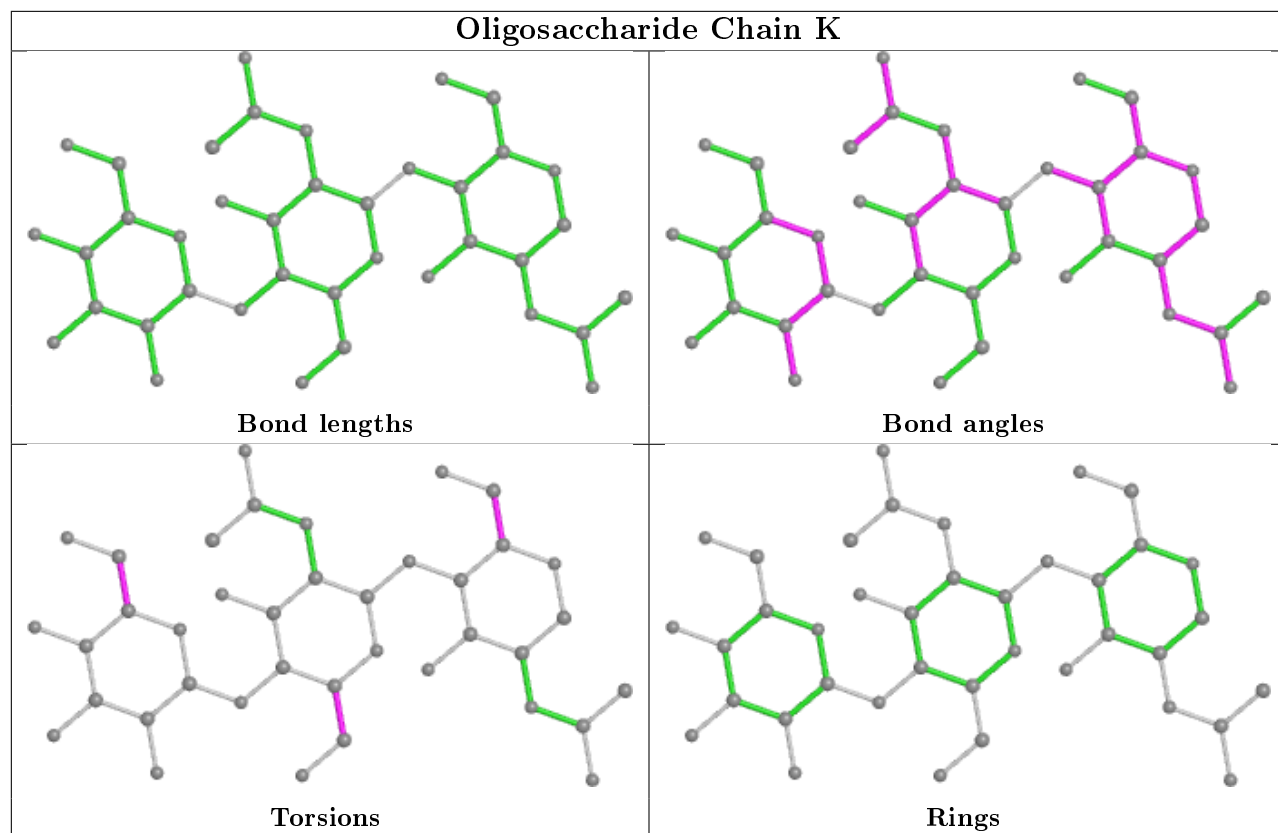
Mol	Chain	Res	Type	Atoms
3	G	3	MAN	C1-C2-C3-C4-C5-O5

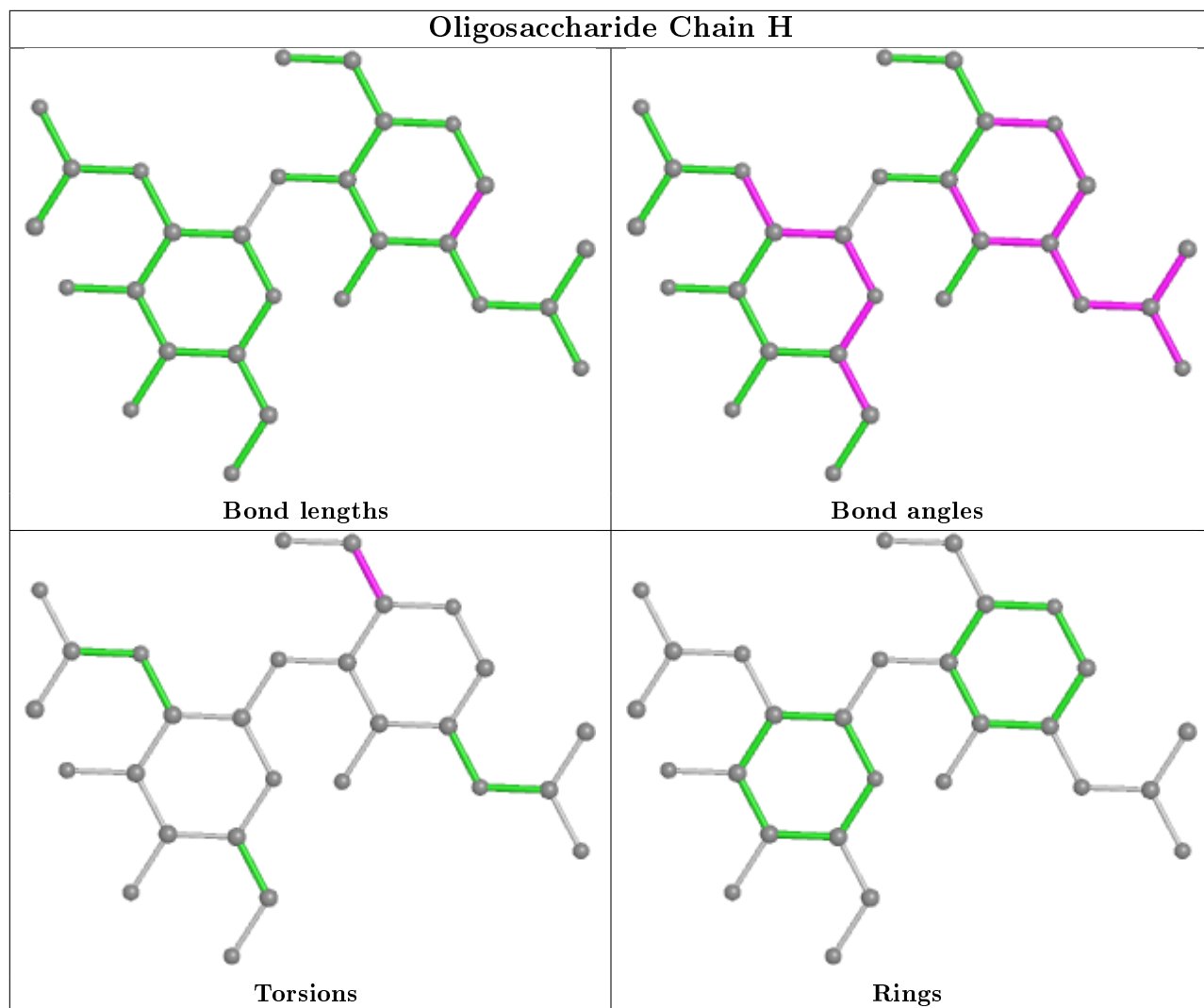
1 monomer is involved in 1 short contact:

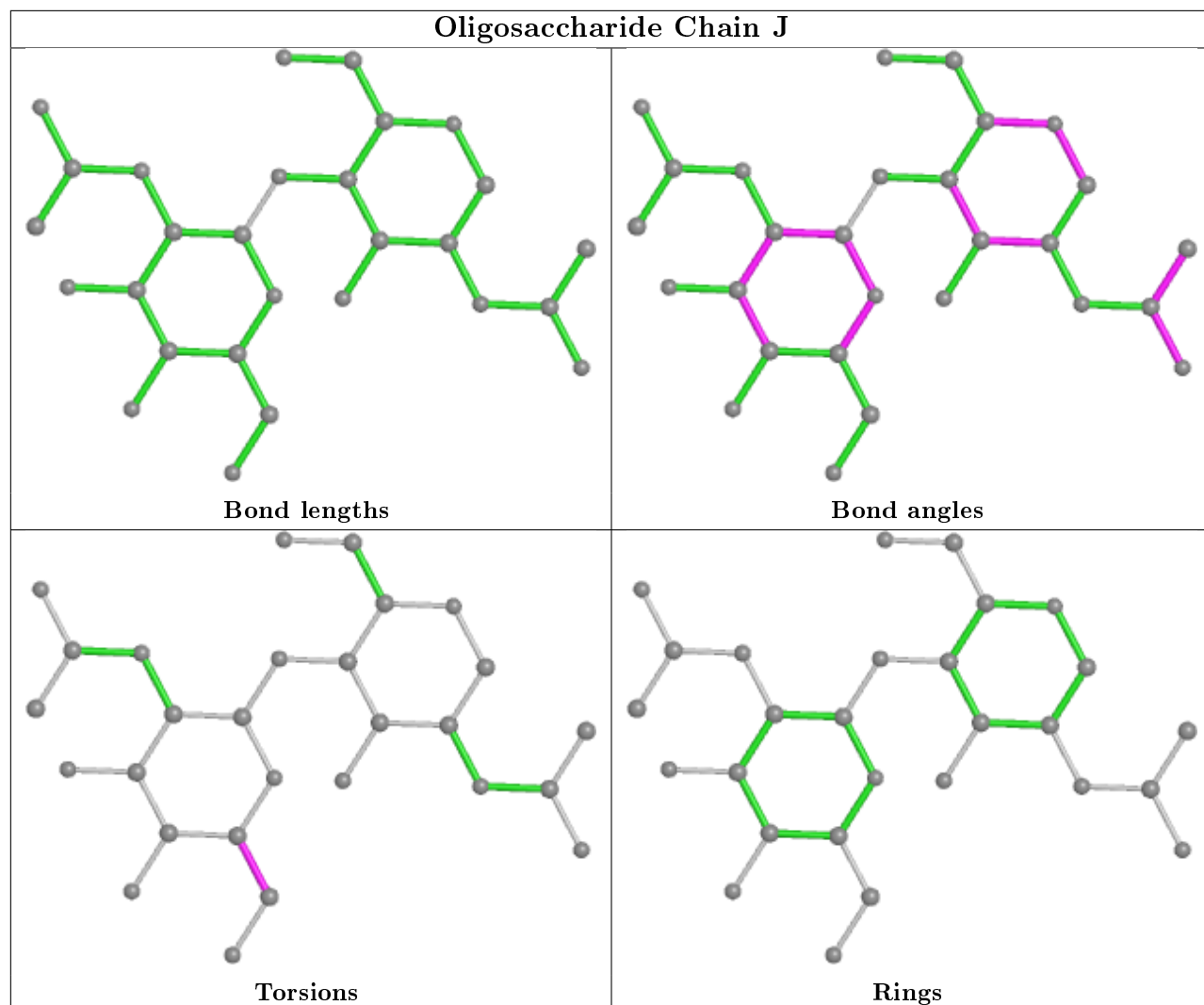
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	J	2	NAG	1	0

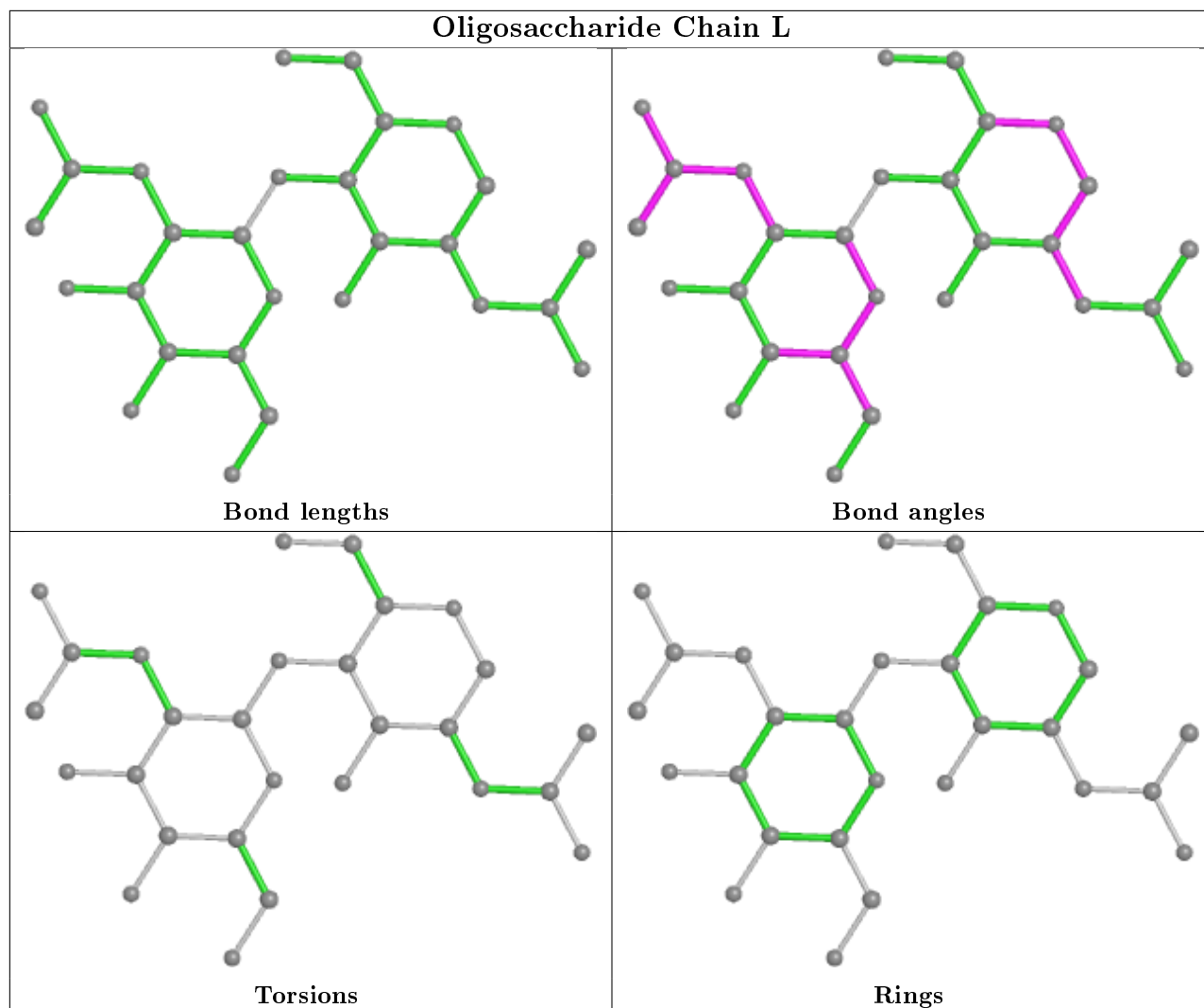
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











## 5.6 Ligand geometry [i](#)

6 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
5	NAG	A	501	1	14,14,15	1.02	1 (7%)	17,19,21	1.95	5 (29%)
5	NAG	E	505	1	14,14,15	0.80	0	17,19,21	1.32	2 (11%)
5	NAG	C	501	1	14,14,15	0.60	0	17,19,21	2.47	4 (23%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	A	505	1	14,14,15	0.64	0	17,19,21	1.61	3 (17%)
5	NAG	C	505	1	14,14,15	0.72	0	17,19,21	1.85	5 (29%)
5	NAG	E	501	1	14,14,15	0.90	1 (7%)	17,19,21	1.54	2 (11%)

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	A	501	1	-	2/6/23/26	0/1/1/1
5	NAG	E	505	1	-	0/6/23/26	0/1/1/1
5	NAG	C	501	1	-	1/6/23/26	0/1/1/1
5	NAG	A	505	1	-	0/6/23/26	0/1/1/1
5	NAG	C	505	1	-	2/6/23/26	0/1/1/1
5	NAG	E	501	1	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	501	NAG	C1-C2	2.89	1.56	1.52
5	E	501	NAG	C1-C2	2.37	1.55	1.52

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	501	NAG	C1-O5-C5	8.69	123.97	112.19
5	A	501	NAG	C1-O5-C5	5.27	119.33	112.19
5	E	501	NAG	C1-O5-C5	4.63	118.47	112.19
5	A	505	NAG	O5-C5-C6	4.38	114.07	107.20
5	C	505	NAG	C1-O5-C5	3.98	117.58	112.19

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	501	NAG	O5-C5-C6-O6
5	E	501	NAG	O5-C5-C6-O6
5	A	501	NAG	C4-C5-C6-O6
5	E	501	NAG	C4-C5-C6-O6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
5	C	505	NAG	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	324/334 (97%)	-0.32	5 (1%) 73 67	38, 57, 89, 190	0
1	C	324/334 (97%)	-0.29	6 (1%) 66 59	32, 51, 92, 162	0
1	E	324/334 (97%)	-0.34	3 (0%) 84 80	32, 49, 88, 134	0
2	B	176/181 (97%)	-0.22	0 100 100	34, 57, 83, 99	0
2	D	176/181 (97%)	0.03	5 (2%) 53 43	33, 79, 144, 197	0
2	F	176/181 (97%)	0.09	5 (2%) 53 43	34, 82, 125, 149	0
All	All	1500/1545 (97%)	-0.21	24 (1%) 72 65	32, 56, 114, 197	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	71	ILE	6.8
1	C	68	ASP	6.3
1	A	69	GLU	6.0
2	D	173	ILE	5.9
1	C	67	CYS	5.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
3	MAN	I	3	11/12	0.55	0.33	116,135,142,150	0
5	NAG	E	505	14/15	0.59	0.41	107,133,147,147	0
3	MAN	G	3	11/12	0.65	0.27	111,131,136,145	0
4	NAG	J	1	14/15	0.65	0.38	131,139,152,159	0

*Continued on next page...*

Continued from previous page...

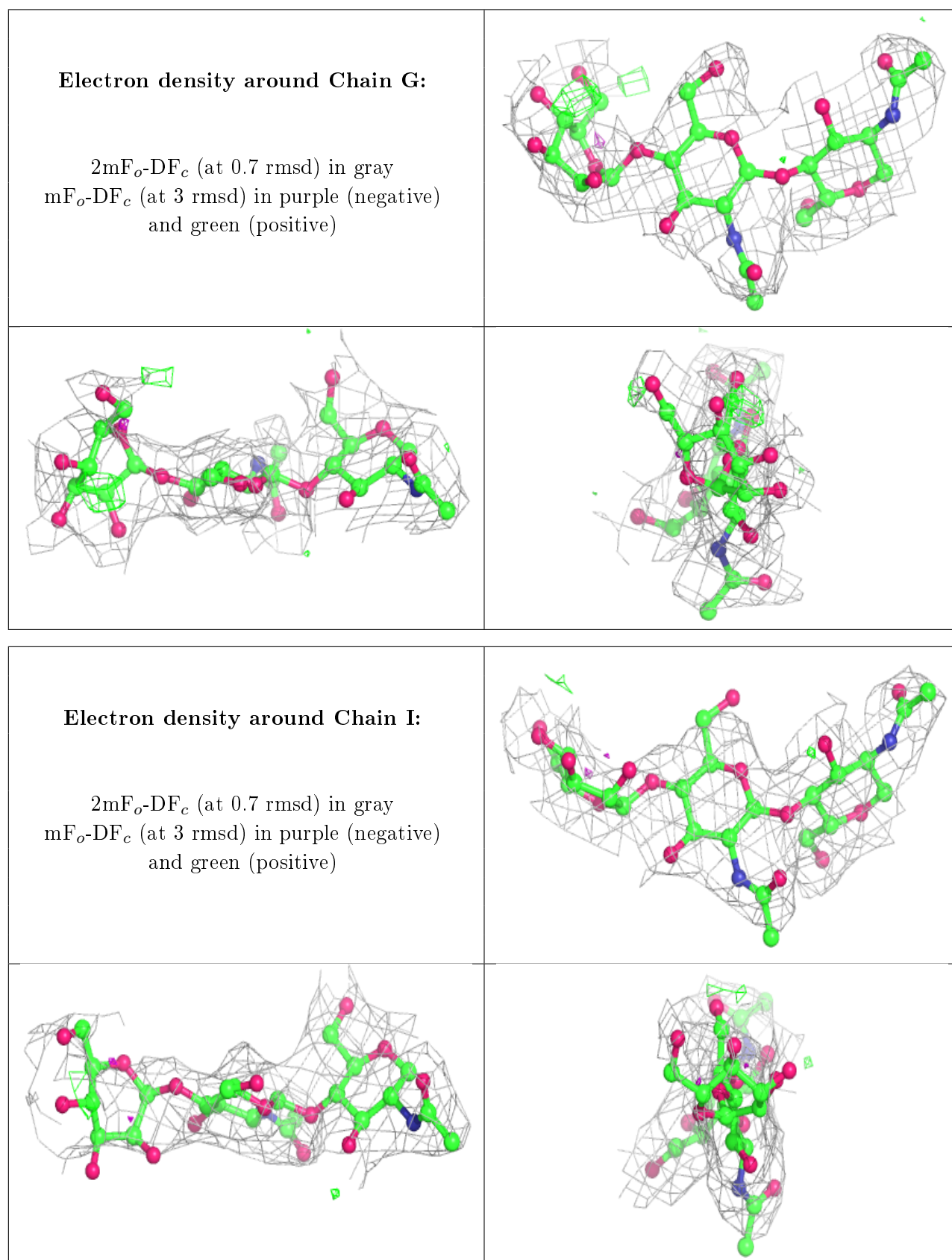
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	L	2	14/15	0.68	0.39	123,138,146,149	0
3	MAN	K	3	11/12	0.69	0.41	106,129,137,141	0
4	NAG	J	2	14/15	0.69	0.45	123,143,156,159	0
4	NAG	H	2	14/15	0.71	0.28	108,129,134,134	0
4	NAG	L	1	14/15	0.72	0.28	123,146,153,160	0
5	NAG	E	501	14/15	0.73	0.21	100,110,120,121	0
3	NAG	K	2	14/15	0.77	0.36	98,109,123,137	0
5	NAG	A	505	14/15	0.78	0.23	98,123,132,138	0
4	NAG	H	1	14/15	0.80	0.32	102,115,119,123	0
5	NAG	C	505	14/15	0.80	0.23	90,103,114,117	0
3	NAG	G	2	14/15	0.81	0.22	102,113,125,137	0
5	NAG	A	501	14/15	0.82	0.22	82,99,111,111	0
5	NAG	C	501	14/15	0.87	0.14	72,87,93,99	0
3	NAG	K	1	14/15	0.88	0.15	66,74,88,100	0
3	NAG	I	2	14/15	0.90	0.26	83,103,121,130	0
3	NAG	G	1	14/15	0.92	0.11	69,72,82,95	0
3	NAG	I	1	14/15	0.94	0.15	60,66,75,89	0

### 6.3 Carbohydrates i

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

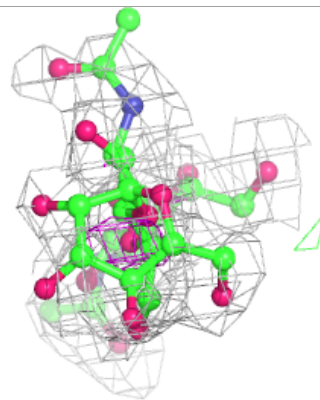
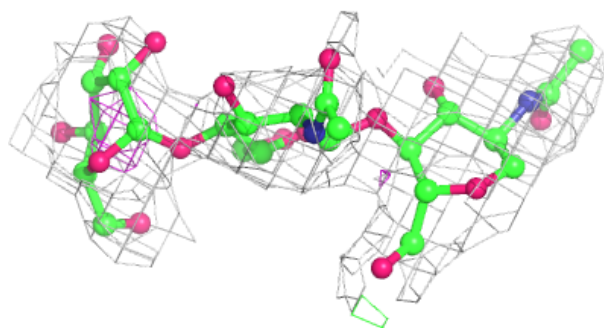
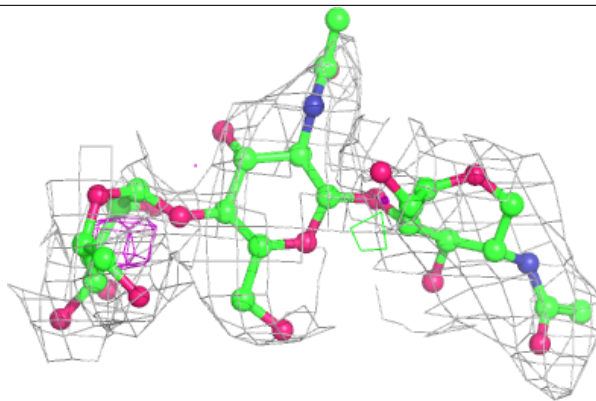
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	MAN	I	3	11/12	0.55	0.33	116,135,142,150	0
4	NAG	J	1	14/15	0.65	0.38	131,139,152,159	0
3	MAN	G	3	11/12	0.65	0.27	111,131,136,145	0
4	NAG	L	2	14/15	0.68	0.39	123,138,146,149	0
4	NAG	J	2	14/15	0.69	0.45	123,143,156,159	0
3	MAN	K	3	11/12	0.69	0.41	106,129,137,141	0
4	NAG	H	2	14/15	0.71	0.28	108,129,134,134	0
4	NAG	L	1	14/15	0.72	0.28	123,146,153,160	0
3	NAG	K	2	14/15	0.77	0.36	98,109,123,137	0
4	NAG	H	1	14/15	0.80	0.32	102,115,119,123	0
3	NAG	G	2	14/15	0.81	0.22	102,113,125,137	0
3	NAG	K	1	14/15	0.88	0.15	66,74,88,100	0
3	NAG	I	2	14/15	0.90	0.26	83,103,121,130	0
3	NAG	G	1	14/15	0.92	0.11	69,72,82,95	0
3	NAG	I	1	14/15	0.94	0.15	60,66,75,89	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.



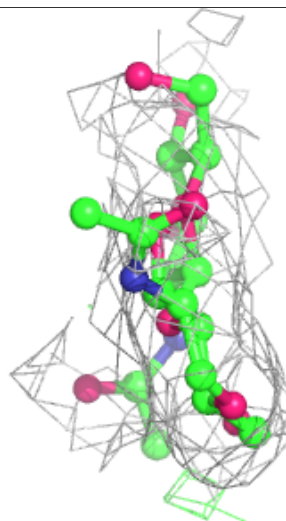
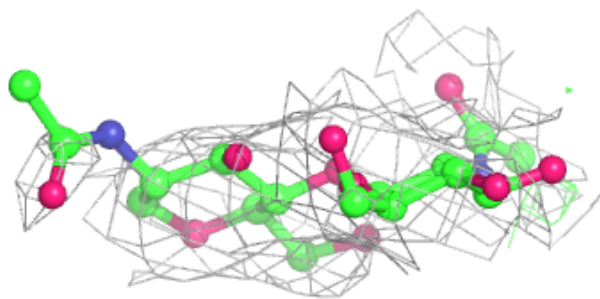
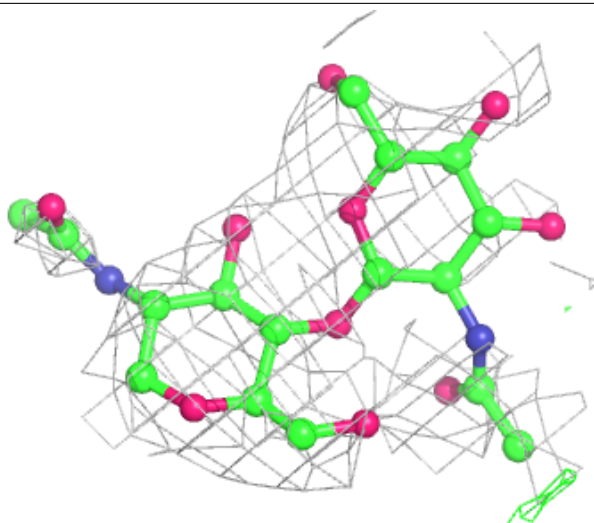
**Electron density around Chain K:**

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



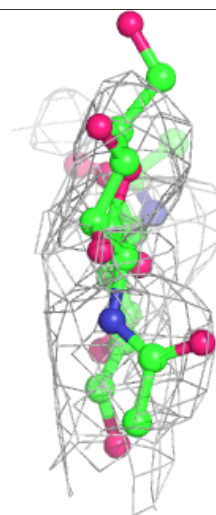
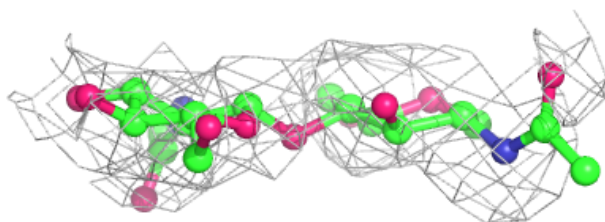
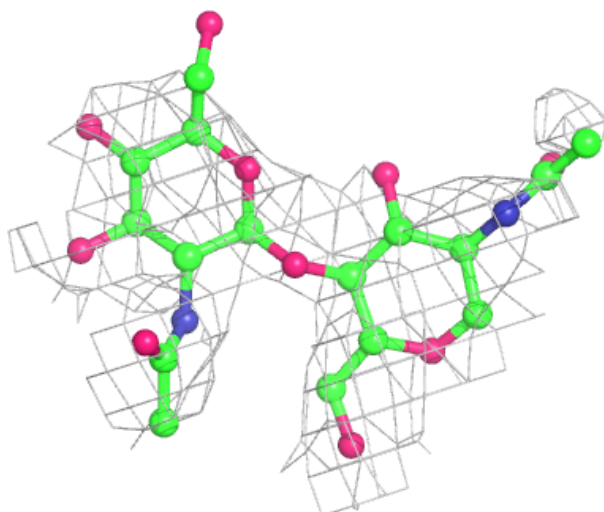
**Electron density around Chain H:**

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

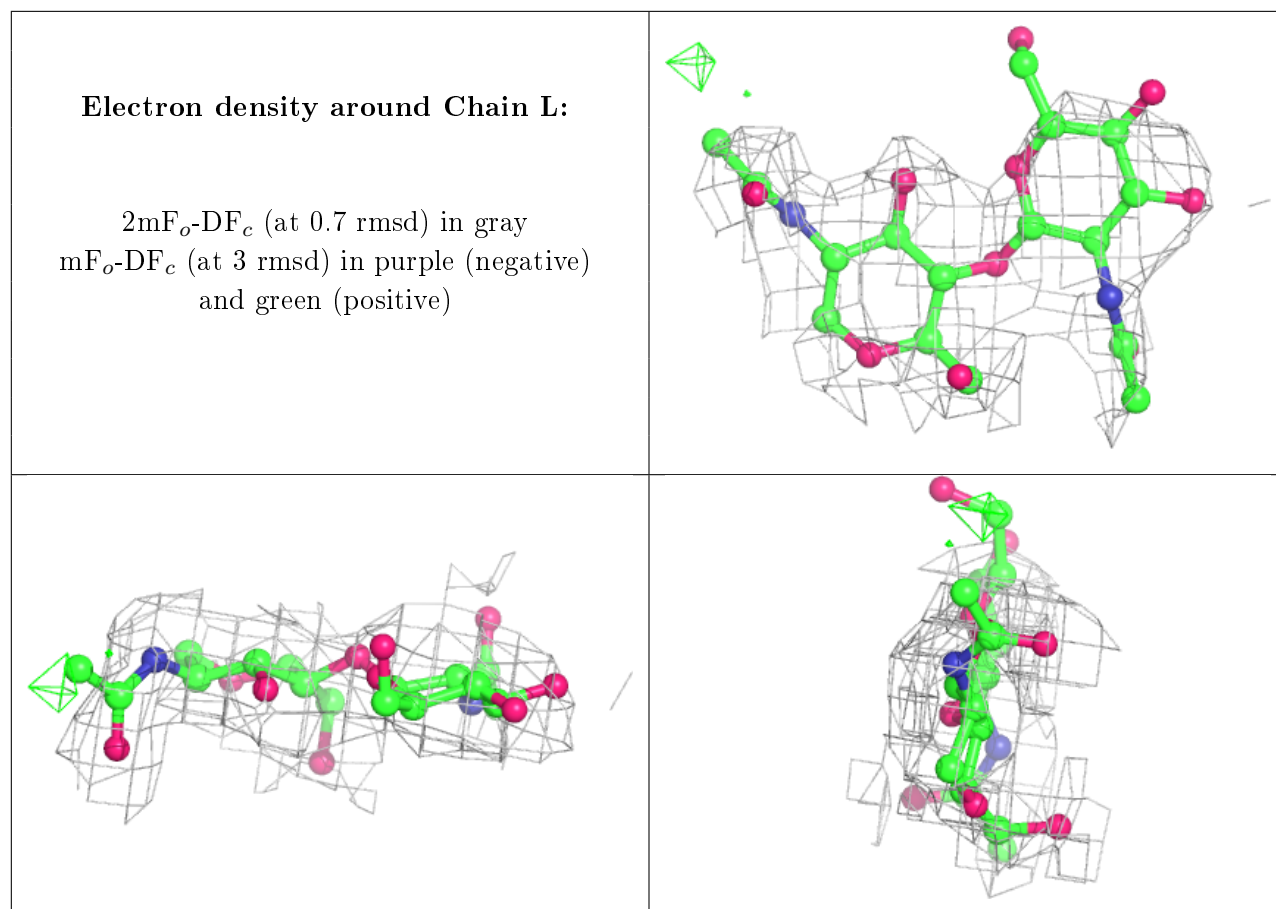


**Electron density around Chain J:**

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







## 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
5	NAG	E	505	14/15	0.59	0.41	107,133,147,147	0
5	NAG	E	501	14/15	0.73	0.21	100,110,120,121	0
5	NAG	A	505	14/15	0.78	0.23	98,123,132,138	0
5	NAG	C	505	14/15	0.80	0.23	90,103,114,117	0
5	NAG	A	501	14/15	0.82	0.22	82,99,111,111	0
5	NAG	C	501	14/15	0.87	0.14	72,87,93,99	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.