



# wwPDB X-ray Structure Validation Summary Report

Dec 12, 2023 – 06:51 pm GMT

PDB ID : 4CYZ  
Title : Structure of the A\_mallard\_Sweden\_51\_2002 H10 Avian Haemmagglutinin in complex with avian receptor analog LSTA  
Authors : Vachieri, S.G.; Xiong, X.; Collins, P.J.; Walker, P.A.; Martin, S.R.; Haire, L.F.; McCauley, J.W.; Gamblin, S.J.; Skehel, J.J.  
Deposited on : 2014-04-16  
Resolution : 2.40 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 : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

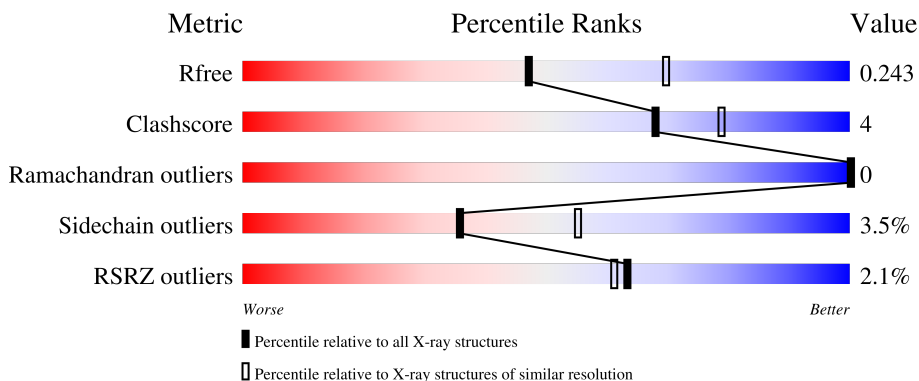
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	318	 89% 10% 2%
1	C	318	 92% 8% 2%
1	E	318	 91% 8% 3%
2	B	172	 84% 16% 2%
2	D	172	 90% 10% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	172	
3	G	3	
4	H	5	
4	J	5	
5	I	2	
6	K	2	
7	L	4	

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
8	NAG	A	420	X	-	-	-
8	NAG	B	208	-	-	-	X
8	NAG	C	420	X	-	-	-
9	EDO	E	1328	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 12227 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.

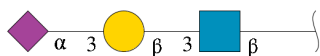
- Molecule 1 is a protein called HEMAGGLUTININ.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	318	Total 2423	C 1498	N 436	O 473	S 16	0	0	0
1	C	318	Total 2419	C 1496	N 435	O 472	S 16	0	0	0
1	E	318	Total 2419	C 1495	N 435	O 473	S 16	0	0	0

- Molecule 2 is a protein called HEMAGGLUTININ.

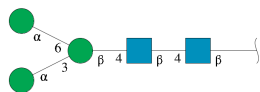
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	172	Total 1381	C 853	N 238	O 282	S 8	0	0	0
2	D	172	Total 1376	C 851	N 238	O 279	S 8	0	0	0
2	F	172	Total 1383	C 854	N 239	O 282	S 8	0	0	0

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	G	3	Total 46	C 25	N 2	O 19	0	0	0

- Molecule 4 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



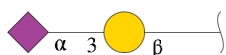
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	H	5	61	34	2	25	0	0	0
4	J	5	61	34	2	25	0	0	0

- Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	I	2	28	16	2	10	0	0	0

- Molecule 6 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose.



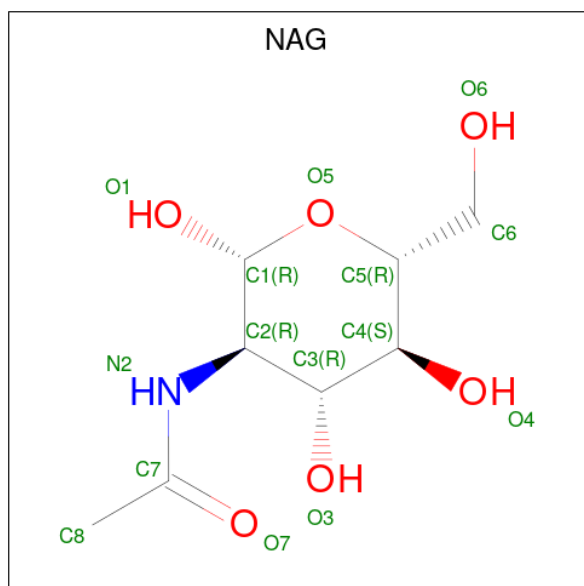
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
6	K	2	31	17	1	13	0	0	0

- Molecule 7 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
7	L	4	50	28	2	20	0	0	0

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 9 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	1	Total C O 4 2 2	0	0
9	A	1	Total C O 4 2 2	0	0
9	A	1	Total C O 4 2 2	0	0
9	B	1	Total C O 4 2 2	0	0
9	C	1	Total C O 4 2 2	0	0
9	C	1	Total C O 4 2 2	0	0
9	E	1	Total C O 4 2 2	0	0
9	E	1	Total C O 4 2 2	0	0
9	E	1	Total C O 4 2 2	0	0
9	F	1	Total C O 4 2 2	0	0
9	F	1	Total C O 4 2 2	0	0

- Molecule 10 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	132	Total O 132 132	0	0

*Continued on next page...*

*Continued from previous page...*

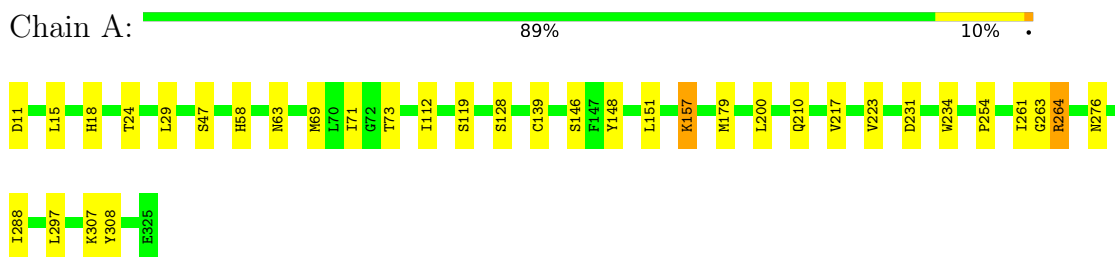
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
10	B	61	Total 61	O 61	0	0
10	C	123	Total 123	O 123	0	0
10	D	26	Total 26	O 26	0	0
10	E	57	Total 57	O 57	0	0
10	F	22	Total 22	O 22	0	0



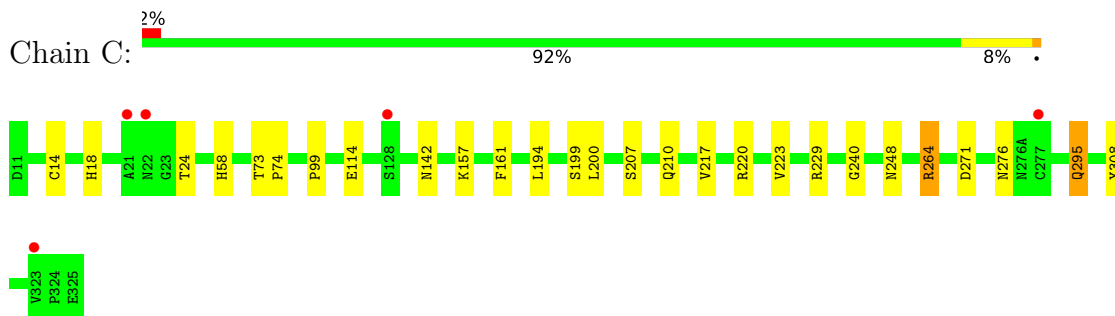
### 3 Residue-property plots [i](#)

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.

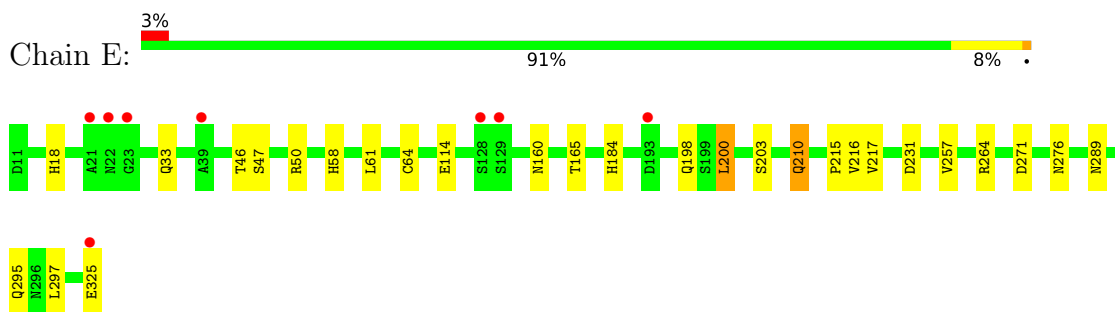
- Molecule 1: HEMAGGLUTININ



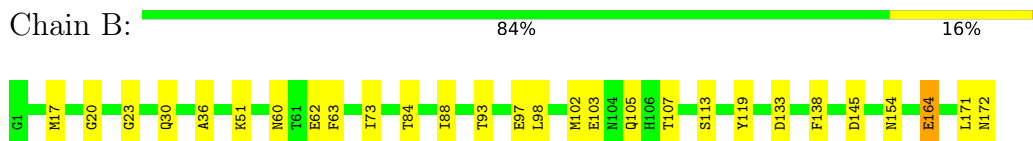
- Molecule 1: HEMAGGLUTININ



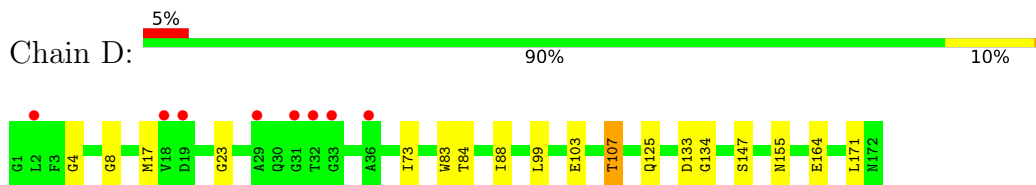
- Molecule 1: HEMAGGLUTININ



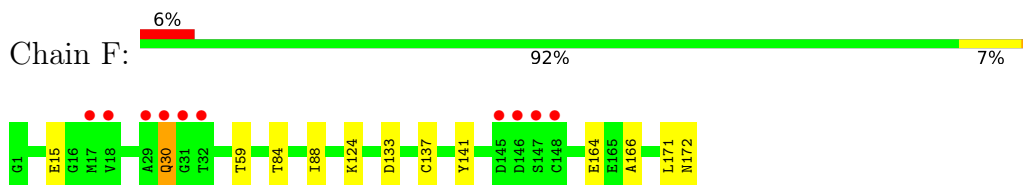
- Molecule 2: HEMAGGLUTININ



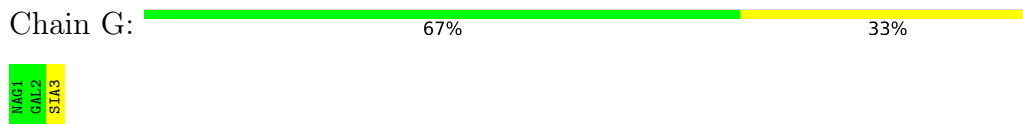
- Molecule 2: HEMAGGLUTININ



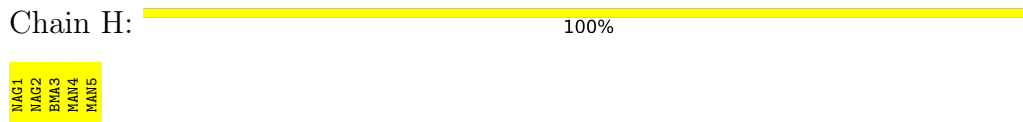
- Molecule 2: HEMAGGLUTININ



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 4: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 4: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



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



- Molecule 6: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



GAL1  
SIA2

- Molecule 7: alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain L: NAG1  
NAG2  
BMA3  
MAN4

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.07Å 230.58Å 68.46Å 90.00° 110.41° 90.00°	Depositor
Resolution (Å)	115.29 – 2.40 48.66 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.0 (115.29-2.40) 99.0 (48.66-2.40)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.70 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.205 , 0.244 0.205 , 0.243	Depositor DCC
$R_{free}$ test set	3773 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.4	Xtrriage
Anisotropy	0.313	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 37.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.029 for l,-k,h	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12227	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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: EDO, GAL, NAG, BMA, SIA, 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	0.37	0/2472	0.55	0/3350
1	C	0.39	1/2468 (0.0%)	0.56	0/3345
1	E	0.35	0/2468	0.53	0/3346
2	B	0.37	0/1406	0.52	0/1899
2	D	0.35	0/1401	0.50	0/1892
2	F	0.33	0/1408	0.50	0/1901
All	All	0.36	1/11623 (0.0%)	0.54	0/15733

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	114	GLU	CD-OE2	6.76	1.33	1.25

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	2423	0	2356	25	0
1	C	2419	0	2350	16	0
1	E	2419	0	2346	16	0
2	B	1381	0	1268	21	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	1376	0	1267	13	0
2	F	1383	0	1275	7	0
3	G	46	0	40	0	0
4	H	61	0	51	0	0
4	J	61	0	52	1	0
5	I	28	0	25	1	0
6	K	31	0	26	0	0
7	L	50	0	43	1	0
8	A	28	0	26	0	0
8	B	28	0	26	1	0
8	C	14	0	13	0	0
8	E	14	0	13	0	0
9	A	12	0	18	4	0
9	B	4	0	6	0	0
9	C	8	0	12	0	0
9	E	12	0	18	1	0
9	F	8	0	12	1	0
10	A	132	0	0	1	0
10	B	61	0	0	2	0
10	C	123	0	0	2	0
10	D	26	0	0	0	0
10	E	57	0	0	0	0
10	F	22	0	0	0	0
All	All	12227	0	11243	87	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 87 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:114:GLU:HG2	1:E:264:ARG:HH12	1.22	1.03
1:E:114:GLU:HG2	1:E:264:ARG:NH1	1.82	0.94
1:E:257:VAL:HA	9:E:1327:EDO:H11	1.56	0.88
1:C:264:ARG:HG3	1:C:264:ARG:HH11	1.43	0.83
1:A:264:ARG:HH11	1:A:264:ARG:HG3	1.45	0.80

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	316/318 (99%)	311 (98%)	5 (2%)	0	100	100
1	C	316/318 (99%)	310 (98%)	6 (2%)	0	100	100
1	E	316/318 (99%)	310 (98%)	6 (2%)	0	100	100
2	B	170/172 (99%)	166 (98%)	4 (2%)	0	100	100
2	D	170/172 (99%)	165 (97%)	5 (3%)	0	100	100
2	F	170/172 (99%)	165 (97%)	5 (3%)	0	100	100
All	All	1458/1470 (99%)	1427 (98%)	31 (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	269/270 (100%)	261 (97%)	8 (3%)	41	61
1	C	268/270 (99%)	256 (96%)	12 (4%)	27	44
1	E	268/270 (99%)	257 (96%)	11 (4%)	30	48
2	B	144/146 (99%)	139 (96%)	5 (4%)	36	55
2	D	143/146 (98%)	139 (97%)	4 (3%)	43	63
2	F	145/146 (99%)	142 (98%)	3 (2%)	53	72
All	All	1237/1248 (99%)	1194 (96%)	43 (4%)	36	55

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	147	SER
1	E	210	GLN
1	E	18	HIS
1	E	165	THR
1	E	271	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	30	GLN
2	D	125	GLN
2	F	30	GLN
1	E	160	ASN
1	E	212	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](#)

21 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	3	15,15,15	0.42	0	21,21,21	0.82	0
3	GAL	G	2	3	11,11,12	0.64	0	15,15,17	0.92	0
3	SIA	G	3	3	20,20,21	0.62	0	24,28,31	1.34	3 (12%)
4	NAG	H	1	2,4	14,14,15	0.66	0	17,19,21	2.17	5 (29%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	H	2	4	14,14,15	0.67	0	17,19,21	0.91	1 (5%)
4	BMA	H	3	4	11,11,12	0.43	0	15,15,17	1.07	2 (13%)
4	MAN	H	4	4	11,11,12	0.78	0	15,15,17	1.79	2 (13%)
4	MAN	H	5	4	11,11,12	0.56	0	15,15,17	1.05	1 (6%)
5	NAG	I	1	1,5	14,14,15	0.50	0	17,19,21	2.27	3 (17%)
5	NAG	I	2	5	14,14,15	0.63	0	17,19,21	1.06	1 (5%)
4	NAG	J	1	2,4	14,14,15	0.58	0	17,19,21	1.17	1 (5%)
4	NAG	J	2	4	14,14,15	0.63	0	17,19,21	0.74	0
4	BMA	J	3	4	11,11,12	0.41	0	15,15,17	1.44	3 (20%)
4	MAN	J	4	4	11,11,12	0.61	0	15,15,17	0.90	1 (6%)
4	MAN	J	5	4	11,11,12	0.59	0	15,15,17	1.97	5 (33%)
6	GAL	K	1	6	11,11,12	0.66	0	15,15,17	1.08	2 (13%)
6	SIA	K	2	6	20,20,21	0.54	0	24,28,31	1.20	3 (12%)
7	NAG	L	1	7,2	14,14,15	0.62	0	17,19,21	0.86	0
7	NAG	L	2	7	14,14,15	0.61	0	17,19,21	1.21	1 (5%)
7	BMA	L	3	7	11,11,12	0.71	0	15,15,17	2.35	4 (26%)
7	MAN	L	4	7	11,11,12	0.57	0	15,15,17	1.10	1 (6%)

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. <sup>1,2</sup> means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	G	1	3	-	2/6/26/26	0/1/1/1
3	GAL	G	2	3	-	0/2/19/22	0/1/1/1
3	SIA	G	3	3	-	3/18/34/38	0/1/1/1
4	NAG	H	1	2,4	-	2/6/23/26	0/1/1/1
4	NAG	H	2	4	-	2/6/23/26	0/1/1/1
4	BMA	H	3	4	-	0/2/19/22	0/1/1/1
4	MAN	H	4	4	-	2/2/19/22	0/1/1/1
4	MAN	H	5	4	-	1/2/19/22	0/1/1/1
5	NAG	I	1	1,5	-	3/6/23/26	0/1/1/1
5	NAG	I	2	5	-	0/6/23/26	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
4	BMA	J	3	4	-	1/2/19/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	MAN	J	4	4	-	2/2/19/22	0/1/1/1
4	MAN	J	5	4	-	2/2/19/22	0/1/1/1
6	GAL	K	1	6	-	0/2/19/22	0/1/1/1
6	SIA	K	2	6	-	4/18/34/38	0/1/1/1
7	NAG	L	1	7,2	-	0/6/23/26	0/1/1/1
7	NAG	L	2	7	-	0/6/23/26	0/1/1/1
7	BMA	L	3	7	-	0/2/19/22	0/1/1/1
7	MAN	L	4	7	-	2/2/19/22	1/1/1/1

There are no bond length outliers.

The worst 5 of 39 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	I	1	NAG	C2-N2-C7	7.38	133.41	122.90
7	L	3	BMA	C1-O5-C5	6.75	121.33	112.19
4	H	4	MAN	C1-C2-C3	5.70	116.68	109.67
4	H	1	NAG	C8-C7-N2	5.24	124.97	116.10
4	J	5	MAN	C1-C2-C3	4.15	114.77	109.67

There are no chirality outliers.

5 of 28 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	I	1	NAG	C3-C2-N2-C7
4	J	4	MAN	O5-C5-C6-O6
4	J	5	MAN	O5-C5-C6-O6
4	H	4	MAN	C4-C5-C6-O6
4	J	4	MAN	C4-C5-C6-O6

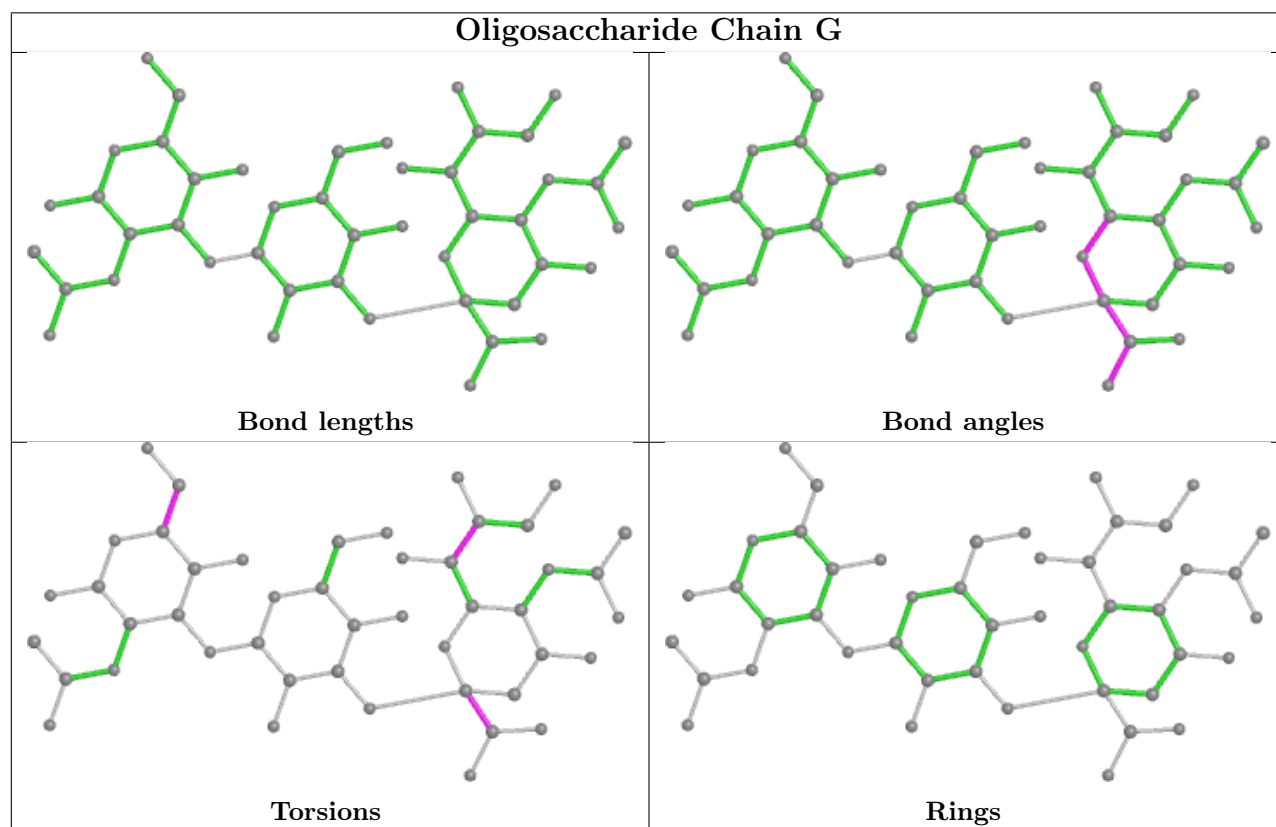
All (1) ring outliers are listed below:

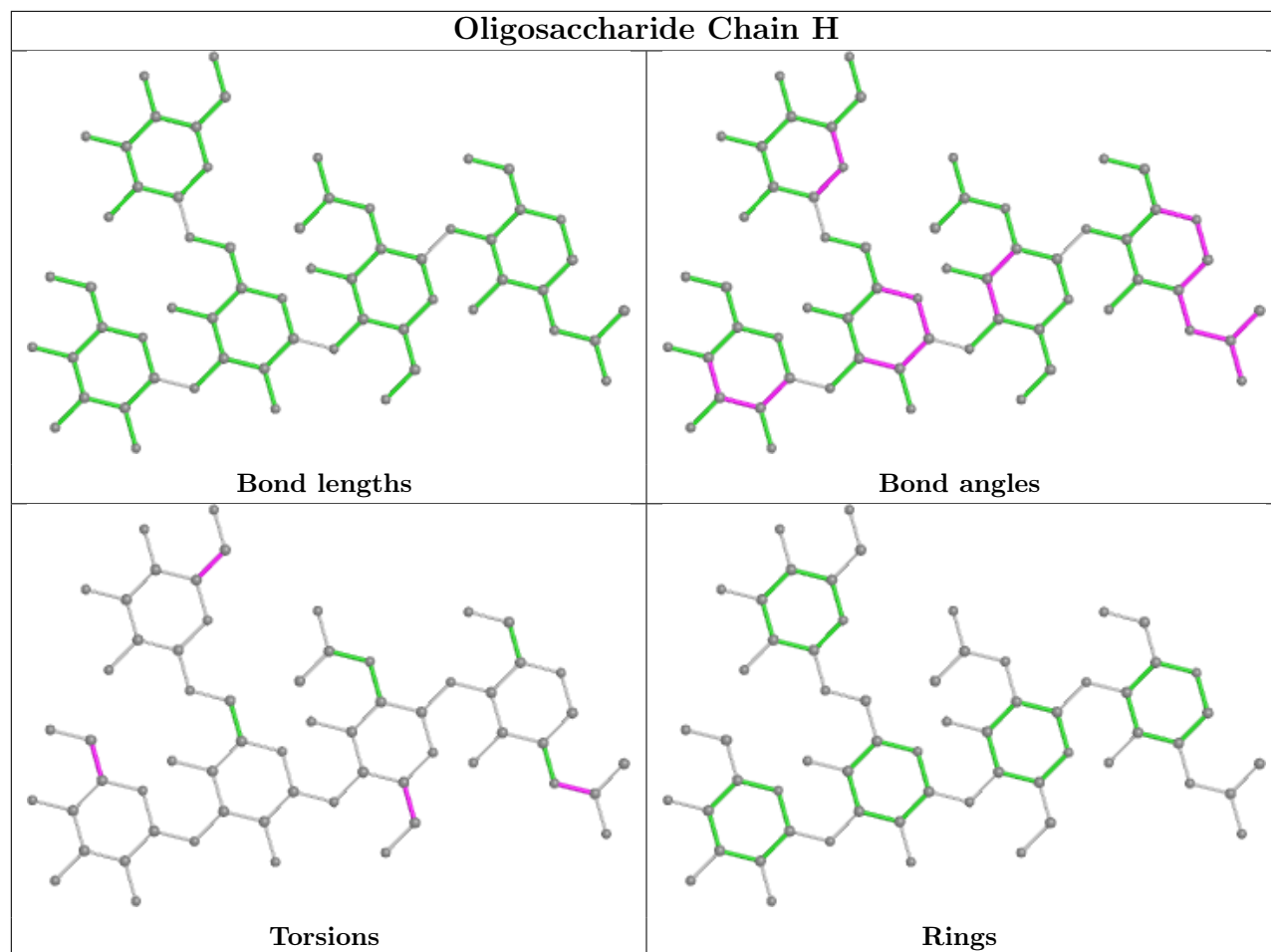
Mol	Chain	Res	Type	Atoms
7	L	4	MAN	C1-C2-C3-C4-C5-O5

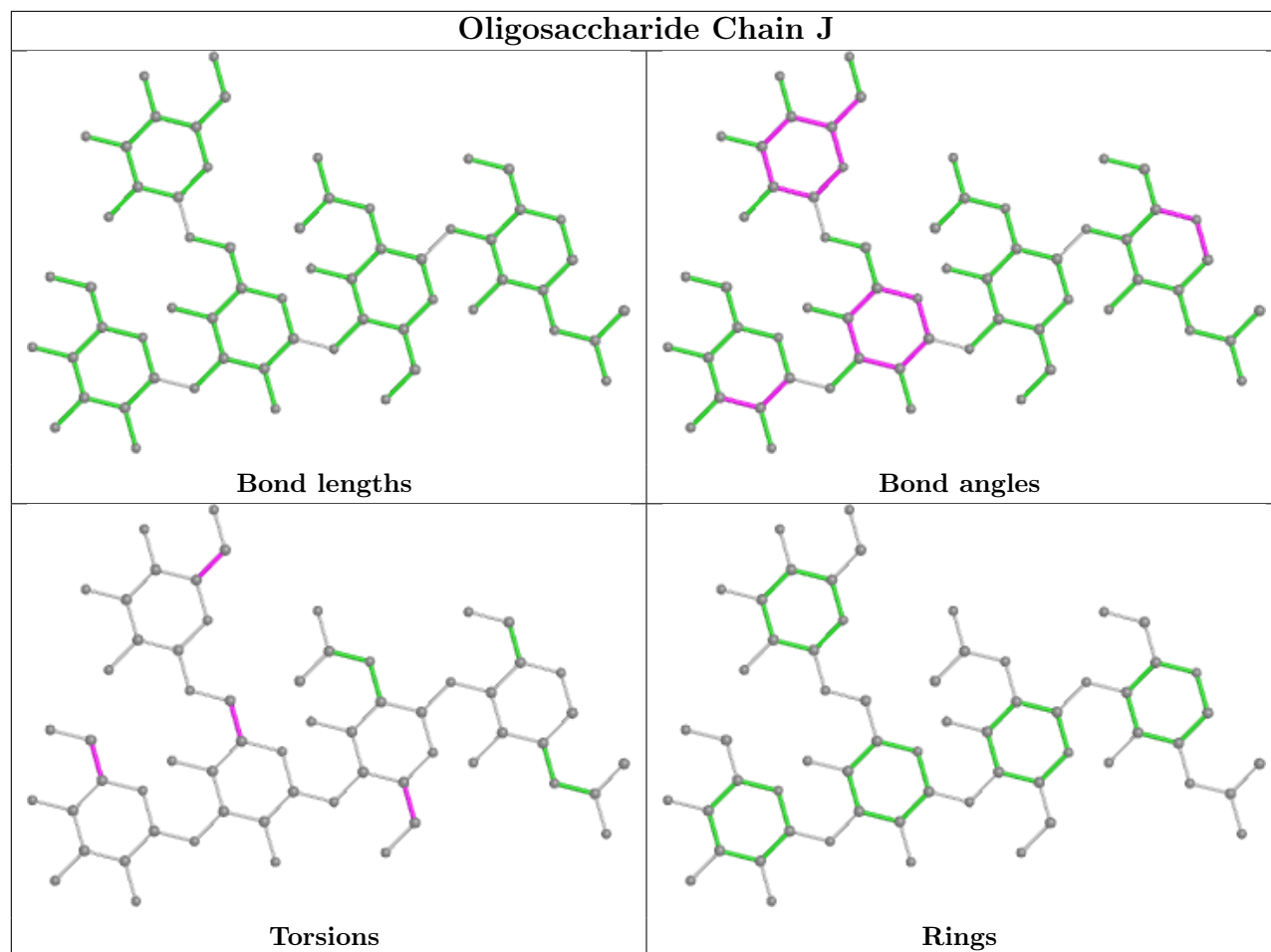
3 monomers are involved in 3 short contacts:

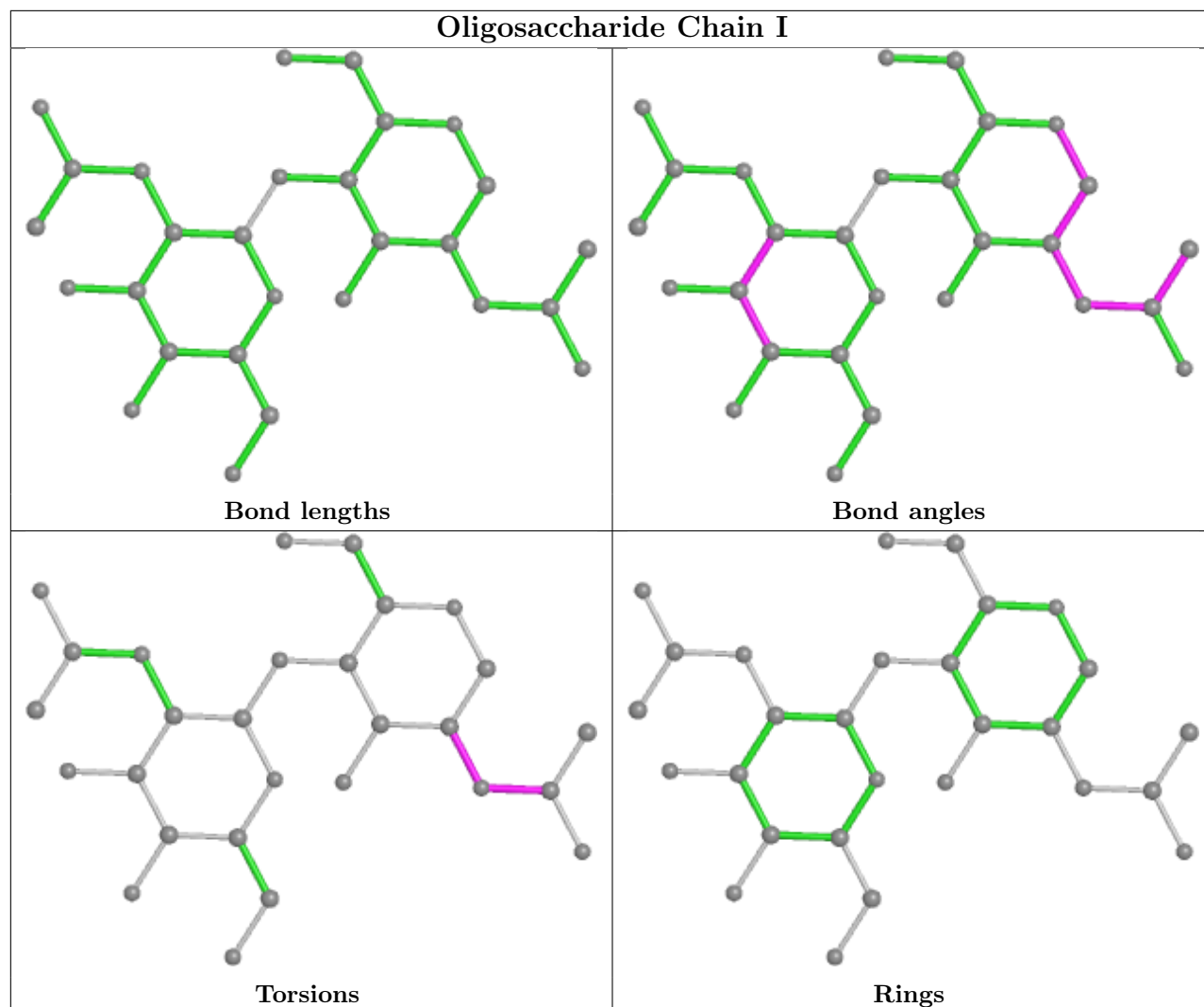
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	J	1	NAG	1	0
5	I	1	NAG	1	0
7	L	4	MAN	1	0

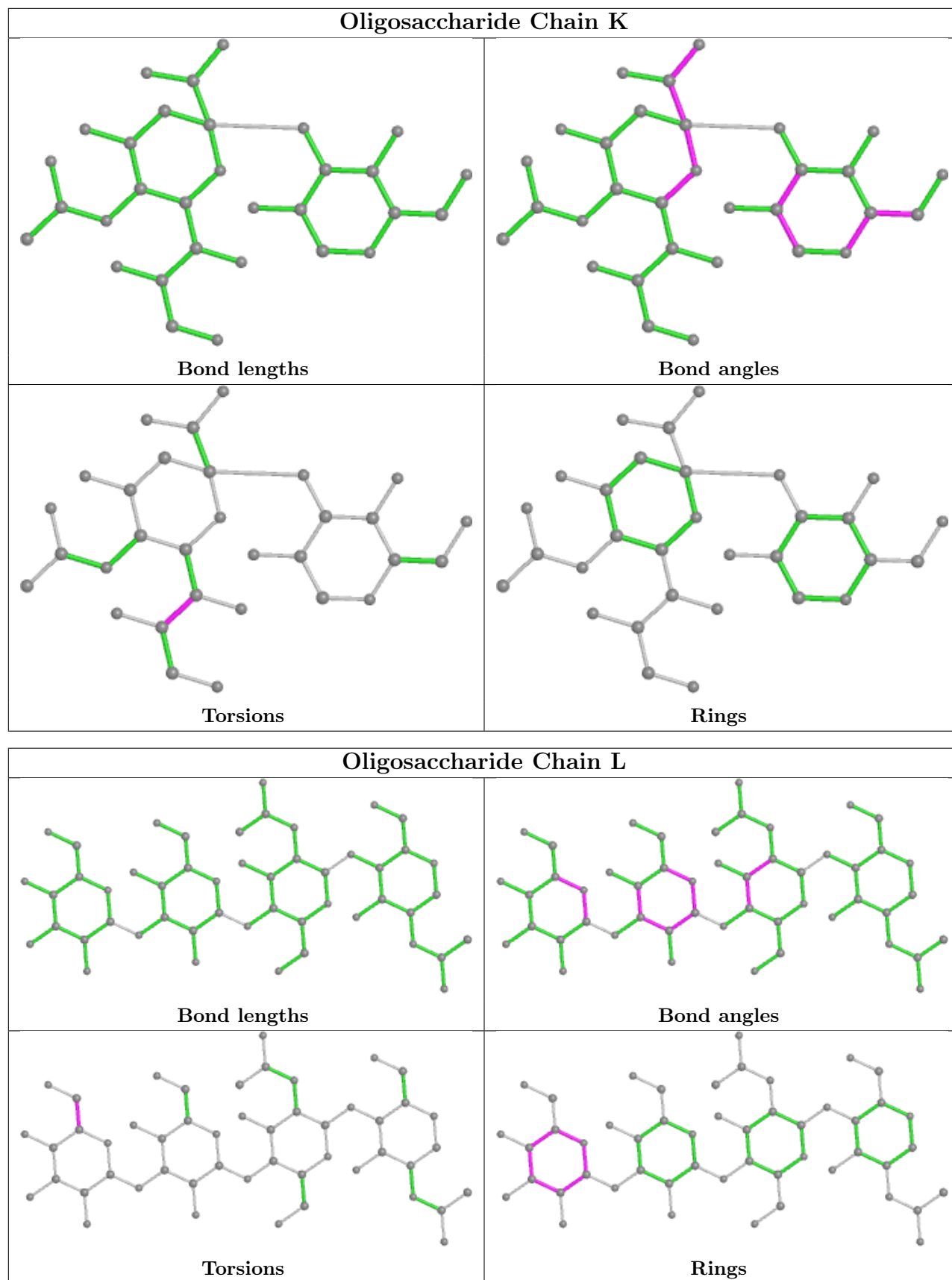
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











## 5.6 Ligand geometry

17 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
8	NAG	B	211	-	14,14,15	0.52	0	17,19,21	0.83	0
8	NAG	B	208	-	14,14,15	0.58	0	17,19,21	1.59	2 (11%)
8	NAG	C	420	1	14,14,15	0.59	0	17,19,21	0.99	1 (5%)
9	EDO	C	1327	-	3,3,3	0.47	0	2,2,2	0.21	0
9	EDO	C	1326	-	3,3,3	0.45	0	2,2,2	0.38	0
9	EDO	B	1173	-	3,3,3	0.40	0	2,2,2	0.59	0
9	EDO	F	1173	-	3,3,3	0.48	0	2,2,2	0.29	0
8	NAG	A	401	1	14,14,15	0.48	0	17,19,21	1.12	1 (5%)
8	NAG	E	401	1	14,14,15	0.58	0	17,19,21	1.04	1 (5%)
9	EDO	A	1326	-	3,3,3	0.47	0	2,2,2	0.34	0
9	EDO	A	1327	-	3,3,3	0.47	0	2,2,2	0.37	0
9	EDO	E	1326	-	3,3,3	0.48	0	2,2,2	0.15	0
9	EDO	E	1328	-	3,3,3	0.56	0	2,2,2	0.27	0
9	EDO	F	1174	-	3,3,3	0.48	0	2,2,2	0.31	0
9	EDO	A	1328	-	3,3,3	0.41	0	2,2,2	0.33	0
8	NAG	A	420	1	14,14,15	0.53	0	17,19,21	1.54	2 (11%)
9	EDO	E	1327	-	3,3,3	0.43	0	2,2,2	0.40	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
8	NAG	B	211	-	-	2/6/23/26	0/1/1/1
8	NAG	C	420	1	1/1/5/7	0/6/23/26	0/1/1/1
8	NAG	B	208	-	-	4/6/23/26	0/1/1/1
9	EDO	C	1327	-	-	0/1/1/1	-
9	EDO	C	1326	-	-	1/1/1/1	-
9	EDO	B	1173	-	-	0/1/1/1	-

*Continued on next page...*



Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	EDO	F	1173	-	-	1/1/1/1	-
8	NAG	A	401	1	-	0/6/23/26	0/1/1/1
8	NAG	E	401	1	-	0/6/23/26	0/1/1/1
9	EDO	A	1326	-	-	0/1/1/1	-
9	EDO	A	1327	-	-	0/1/1/1	-
9	EDO	E	1326	-	-	1/1/1/1	-
9	EDO	E	1328	-	-	1/1/1/1	-
9	EDO	F	1174	-	-	1/1/1/1	-
9	EDO	A	1328	-	-	1/1/1/1	-
8	NAG	A	420	1	1/1/5/7	0/6/23/26	0/1/1/1
9	EDO	E	1327	-	-	1/1/1/1	-

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	A	420	NAG	O5-C1-C2	-5.12	103.20	111.29
8	B	208	NAG	C2-N2-C7	4.73	129.64	122.90
8	A	401	NAG	C1-O5-C5	2.41	115.46	112.19
8	E	401	NAG	C4-C3-C2	2.19	114.23	111.02
8	C	420	NAG	C4-C3-C2	2.05	114.02	111.02

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
8	A	420	NAG	C1
8	C	420	NAG	C1

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	B	208	NAG	C3-C2-N2-C7
8	B	211	NAG	O5-C5-C6-O6
8	B	211	NAG	C4-C5-C6-O6
9	A	1328	EDO	O1-C1-C2-O2
8	B	208	NAG	C4-C5-C6-O6

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	211	NAG	1	0
9	F	1173	EDO	1	0
9	A	1327	EDO	3	0
9	A	1328	EDO	1	0
9	E	1327	EDO	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	318/318 (100%)	-0.24	0 <b>100</b> <b>100</b>	27, 37, 52, 69	0
1	C	318/318 (100%)	-0.24	5 (1%) 72 70	24, 37, 57, 94	0
1	E	318/318 (100%)	0.08	8 (2%) 57 55	34, 51, 73, 95	0
2	B	172/172 (100%)	0.02	0 <b>100</b> <b>100</b>	21, 45, 61, 70	0
2	D	172/172 (100%)	0.32	8 (4%) 31 30	23, 52, 76, 88	0
2	F	172/172 (100%)	0.46	10 (5%) 23 22	23, 58, 86, 94	0
All	All	1470/1470 (100%)	0.01	31 (2%) 63 61	21, 45, 74, 95	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	21	ALA	3.9
1	C	21	ALA	3.8
2	D	33	GLY	3.7
2	F	146	ASP	3.5
2	D	32	THR	3.2

### 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](#)

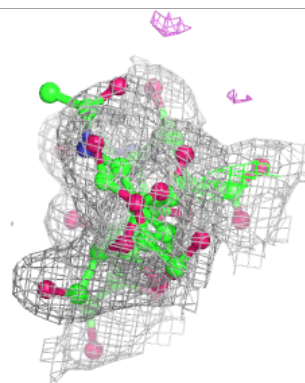
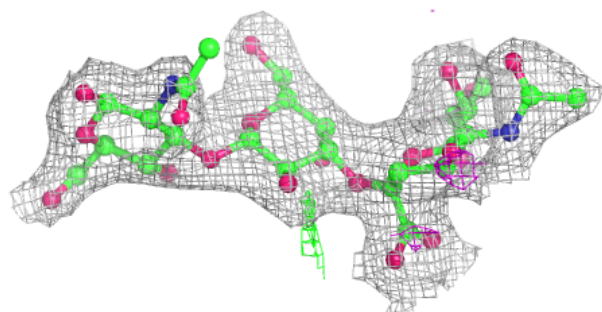
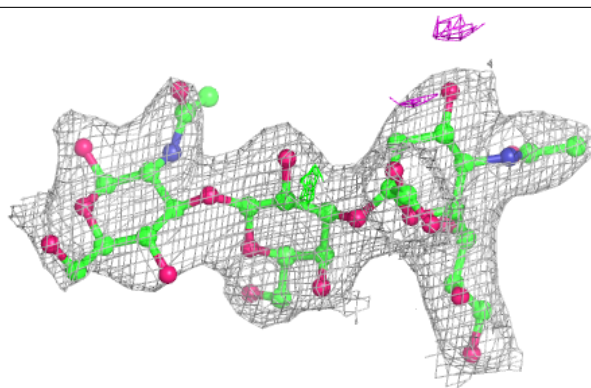
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 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
6	GAL	K	1	11/12	0.72	0.26	122,132,137,139	0
4	BMA	J	3	11/12	0.76	0.17	78,82,85,86	0
7	BMA	L	3	11/12	0.76	0.26	54,57,59,61	0
5	NAG	I	2	14/15	0.77	0.36	103,107,113,113	0
4	MAN	H	4	11/12	0.80	0.29	74,79,86,88	0
4	MAN	J	5	11/12	0.80	0.23	81,84,90,91	0
5	NAG	I	1	14/15	0.80	0.25	93,97,100,102	0
6	SIA	K	2	20/21	0.81	0.24	102,111,120,120	0
7	MAN	L	4	11/12	0.82	0.32	63,65,66,67	0
4	MAN	J	4	11/12	0.83	0.32	84,87,89,91	0
4	NAG	J	2	14/15	0.83	0.20	68,71,75,75	0
3	NAG	G	1	15/15	0.84	0.24	73,79,87,88	0
4	BMA	H	3	11/12	0.85	0.19	69,73,76,77	0
4	NAG	H	1	14/15	0.85	0.26	58,59,61,62	0
4	MAN	H	5	11/12	0.87	0.32	72,74,78,79	0
4	NAG	J	1	14/15	0.89	0.26	62,64,65,67	0
4	NAG	H	2	14/15	0.90	0.19	62,64,66,68	0
7	NAG	L	2	14/15	0.91	0.17	46,48,50,52	0
7	NAG	L	1	14/15	0.93	0.22	42,43,44,45	0
3	GAL	G	2	11/12	0.94	0.14	61,67,70,71	0
3	SIA	G	3	20/21	0.94	0.12	52,56,59,60	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

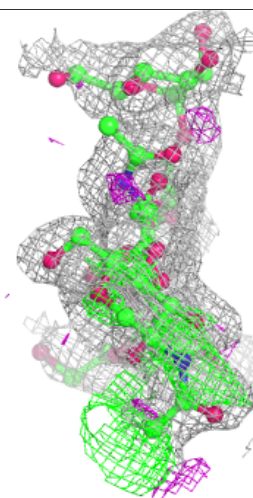
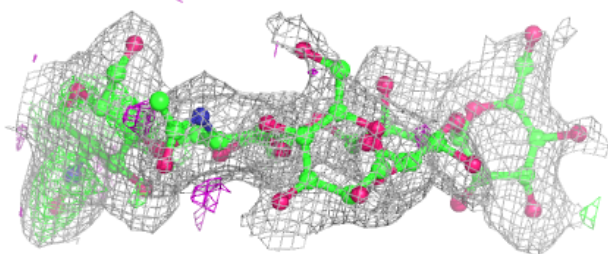
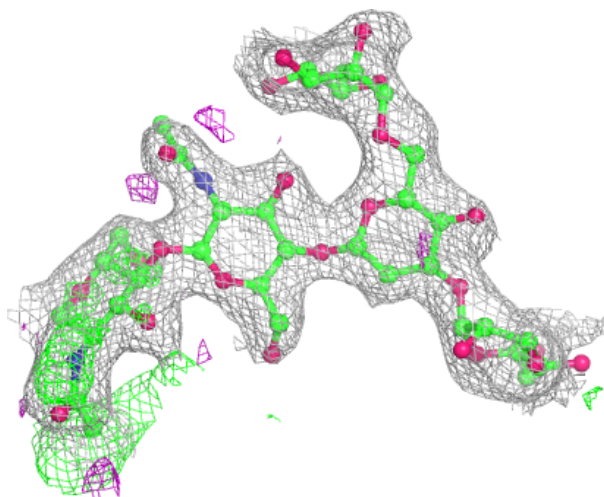
**Electron density around Chain G:**

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



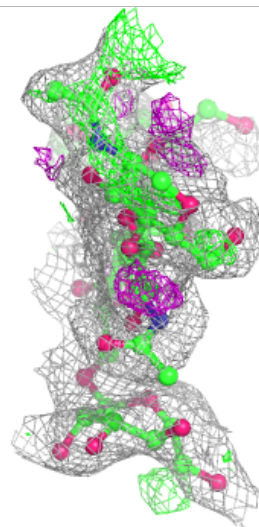
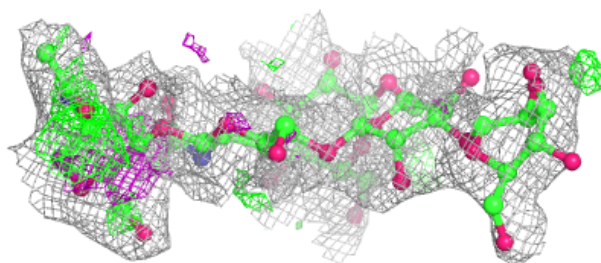
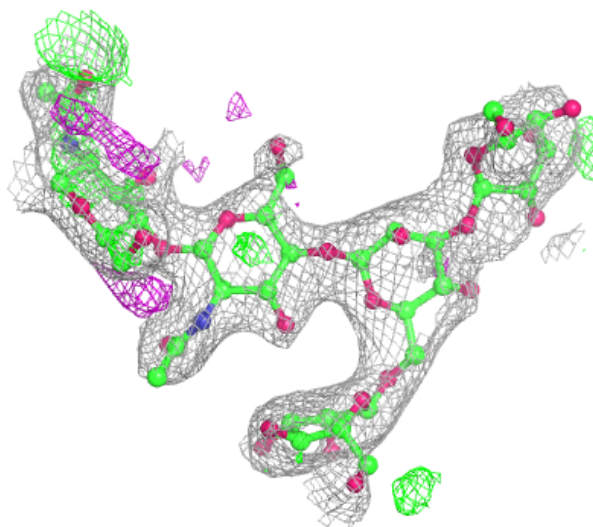
**Electron density around Chain H:**

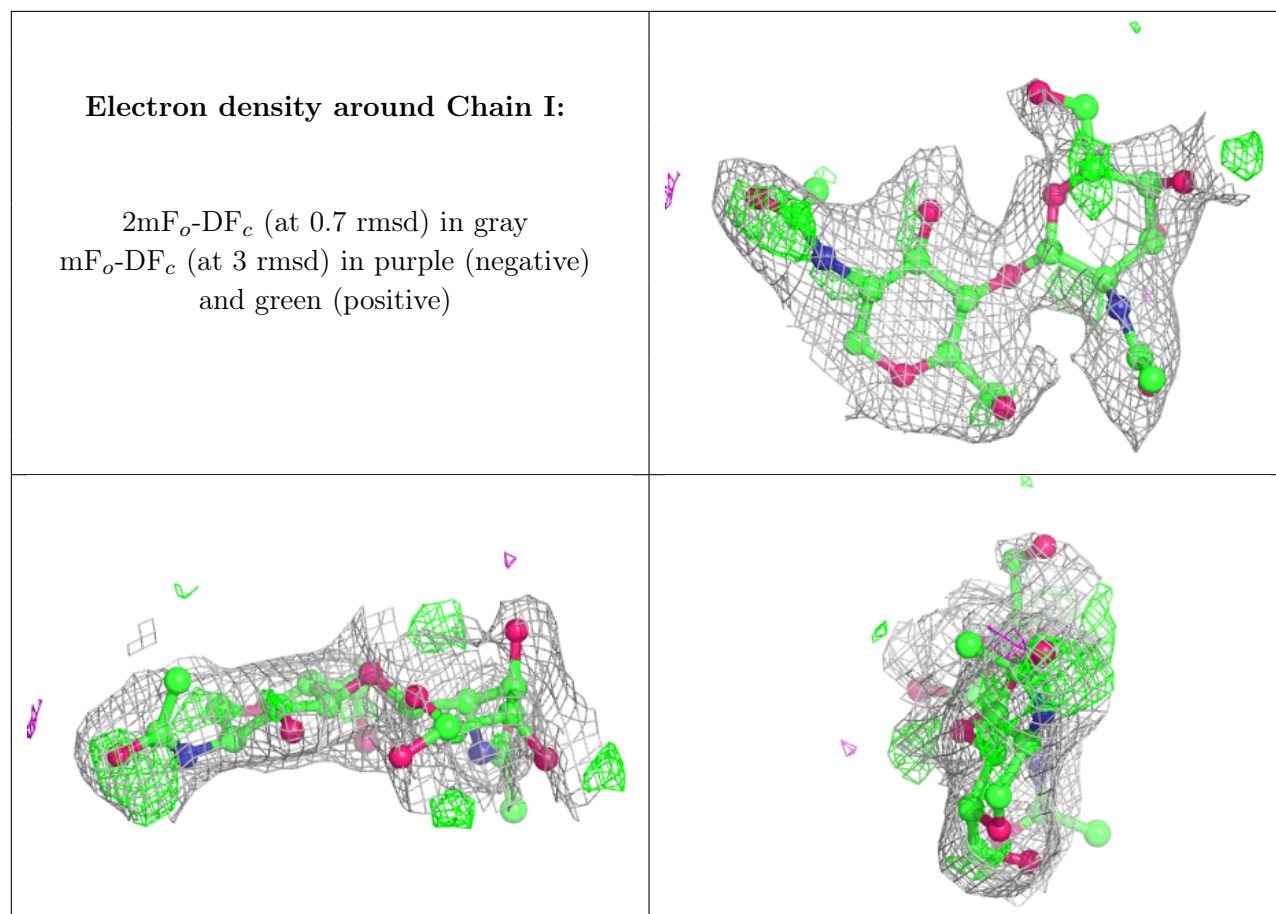
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain J:**

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

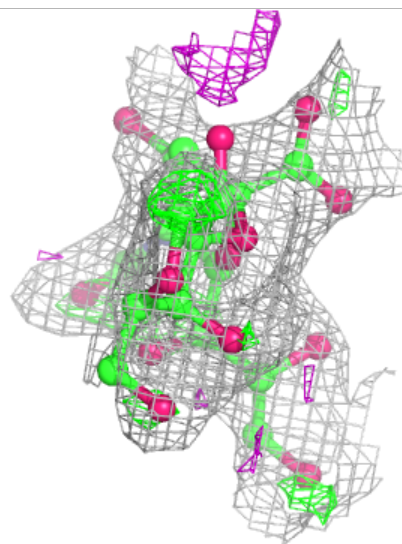
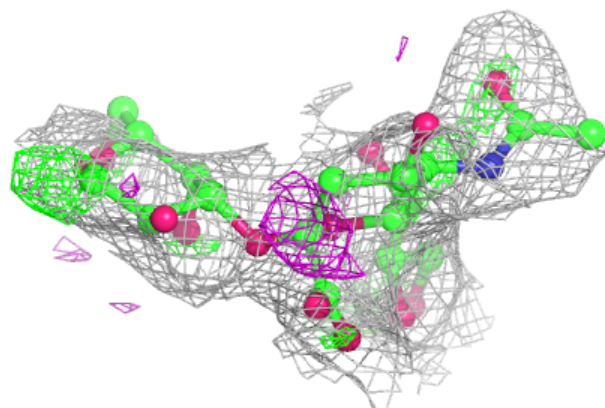
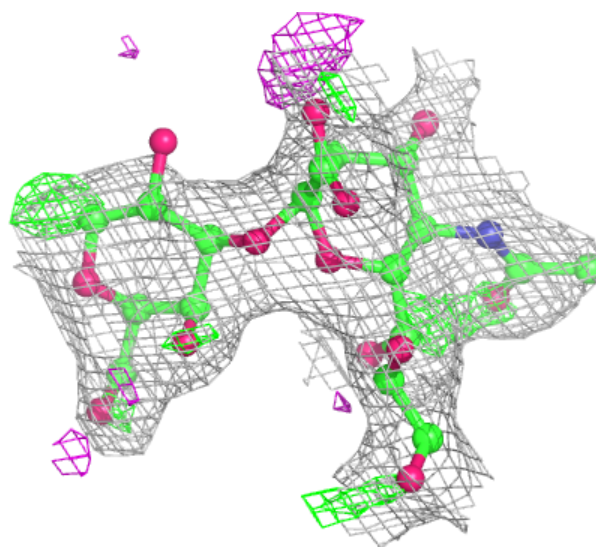


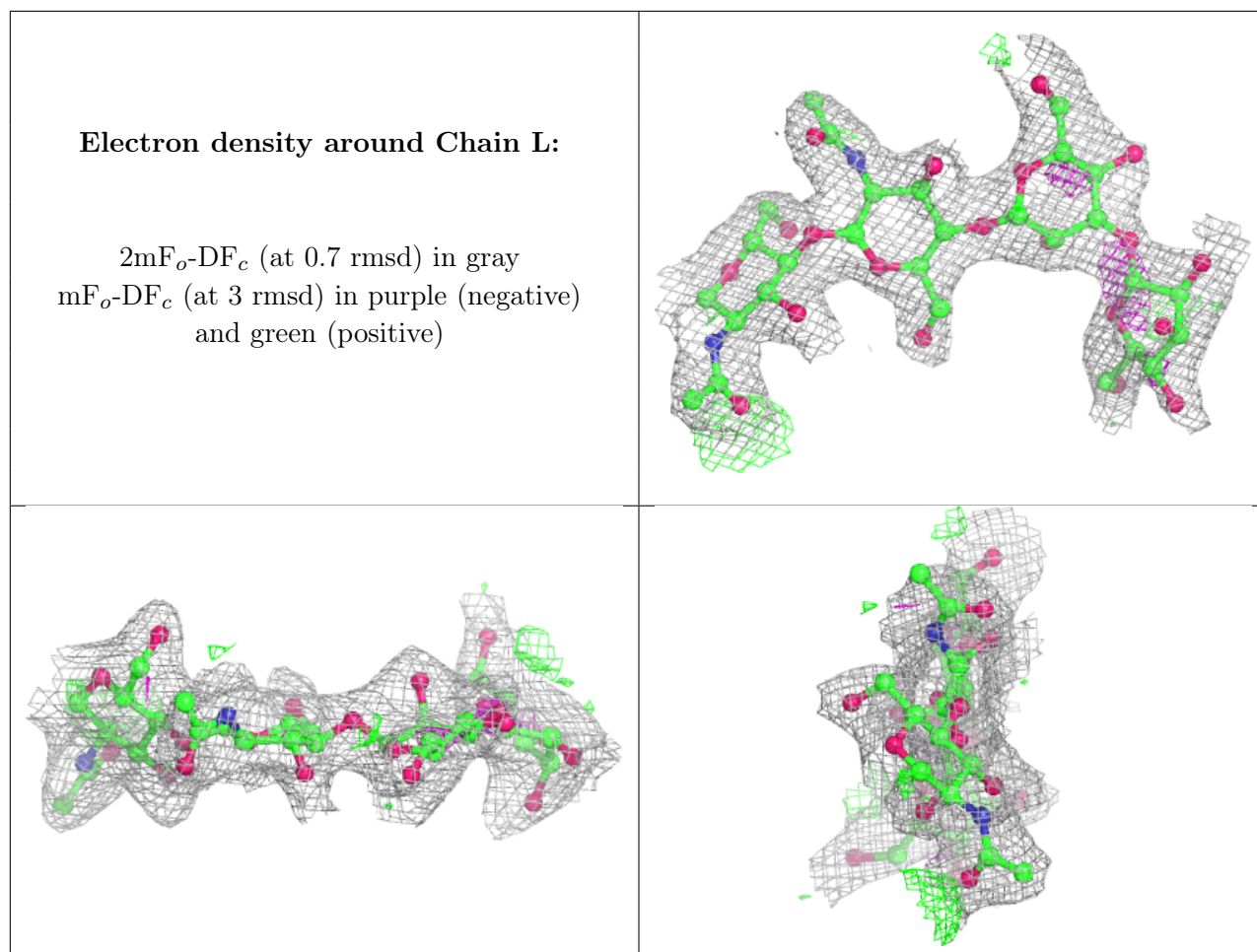




**Electron density around Chain K:**

$2mF_o-DF_c$  (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
8	NAG	B	208	14/15	0.73	0.42	81,92,102,102	0
8	NAG	E	401	14/15	0.73	0.34	156,165,171,179	0
9	EDO	E	1328	4/4	0.73	0.56	66,75,75,76	0
8	NAG	A	420	14/15	0.79	0.25	82,91,97,102	0
8	NAG	B	211	14/15	0.79	0.27	80,90,103,104	0
8	NAG	A	401	14/15	0.80	0.20	87,92,98,100	0
8	NAG	C	420	14/15	0.82	0.21	148,164,180,184	0
9	EDO	A	1327	4/4	0.90	0.38	45,46,46,47	0
9	EDO	F	1174	4/4	0.91	0.12	53,54,54,55	0
9	EDO	E	1326	4/4	0.94	0.19	50,51,52,52	0
9	EDO	B	1173	4/4	0.94	0.31	43,43,44,44	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
9	EDO	F	1173	4/4	0.94	0.23	53,55,55,58	0
9	EDO	C	1326	4/4	0.94	0.28	38,40,41,45	0
9	EDO	A	1326	4/4	0.95	0.21	36,37,37,37	0
9	EDO	A	1328	4/4	0.95	0.23	53,54,55,58	0
9	EDO	E	1327	4/4	0.96	0.51	47,47,48,48	0
9	EDO	C	1327	4/4	0.97	0.12	40,41,42,42	0

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

There are no such residues in this entry.