



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 6, 2020 – 01:32 PM BST

PDB ID : 4JTV  
Title : Crystal structure of 2009 pandemic influenza virus hemagglutinin complexed with human receptor analogue LSTc  
Authors : Zhang, W.; Shi, Y.; Qi, J.; Gao, F.; Li, Q.; Fan, Z.; Yan, J.; Gao, G.F.  
Deposited on : 2013-03-24  
Resolution : 3.00 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  
buster-report : 1.1.7 (2018)  
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

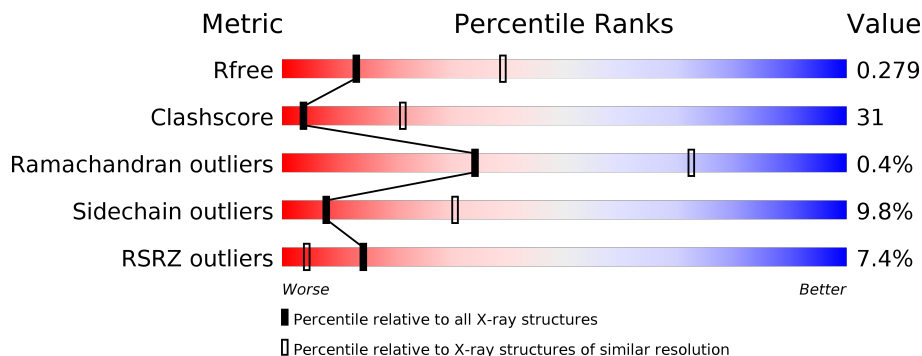
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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.



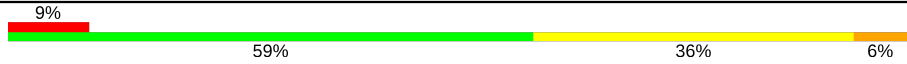




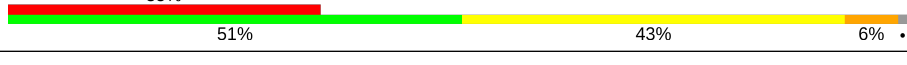

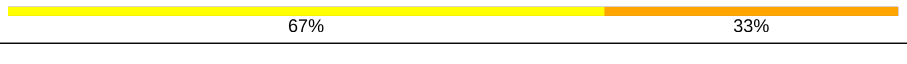
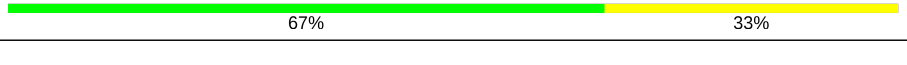
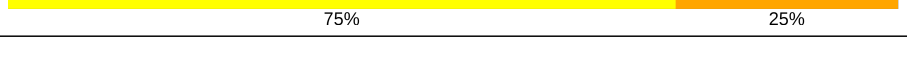
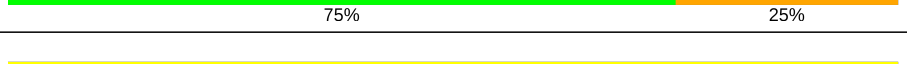
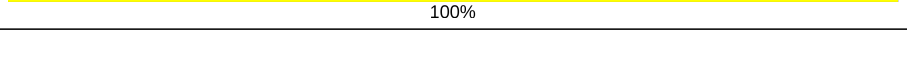
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	321	
1	C	321	
1	E	321	
1	G	321	
1	I	321	
1	K	321	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	162	
2	D	162	
2	F	162	
2	H	162	
2	J	162	
2	L	162	
3	M	3	
3	Q	3	
3	R	3	
4	N	4	
4	P	4	
5	O	2	

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	NAG	G	601	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 23528 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.

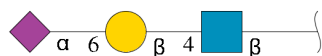
- Molecule 1 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	321	2505	1584	433	477	11	0	0	0
1	C	321	2509	1586	433	479	11	0	0	0
1	E	321	2509	1586	433	479	11	0	0	0
1	G	321	2505	1584	433	477	11	0	0	0
1	I	321	2510	1587	433	479	11	0	0	0
1	K	321	2509	1586	433	479	11	0	0	0

- Molecule 2 is a protein called Hemagglutinin.

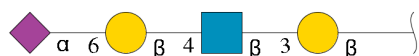
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	162	1305	822	220	257	6	0	0	0
2	D	162	1300	818	219	257	6	0	0	0
2	F	161	1302	821	219	256	6	0	0	0
2	H	162	1305	822	220	257	6	0	0	0
2	J	162	1300	818	219	257	6	0	0	0
2	L	161	1302	821	219	256	6	0	0	0

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	M	3	Total	C	N	O	0	0	0
			45	25	2	18			
3	Q	3	Total	C	N	O	0	0	0
			45	25	2	18			
3	R	3	Total	C	N	O	0	0	0
			45	25	2	18			

- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	N	4	Total	C	N	O	0	0	0
			56	31	2	23			
4	P	4	Total	C	N	O	0	0	0
			56	31	2	23			

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



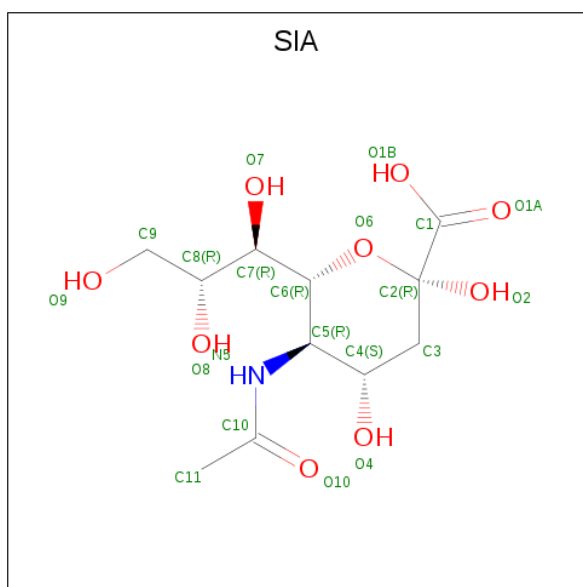
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	O	2	Total	C	N	O	0	0	0
			28	16	2	10			

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



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

- Molecule 7 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula: C<sub>11</sub>H<sub>19</sub>NO<sub>9</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
7	G	1	20	11	1	8	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	40	Total	O	0	0
			40	40		
8	B	13	Total	O	0	0
			13	13		
8	C	29	Total	O	0	0
			29	29		
8	D	13	Total	O	0	0
			13	13		
8	E	31	Total	O	0	0
			31	31		
8	F	7	Total	O	0	0
			7	7		
8	G	19	Total	O	0	0
			19	19		
8	H	13	Total	O	0	0
			13	13		
8	I	35	Total	O	0	0
			35	35		
8	J	18	Total	O	0	0
			18	18		
8	K	25	Total	O	0	0
			25	25		

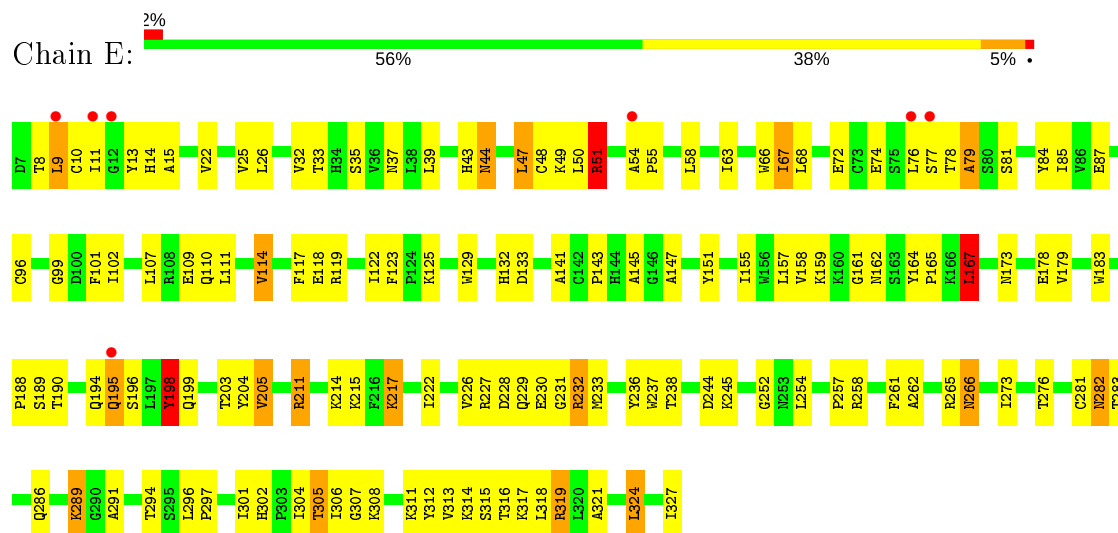
*Continued on next page...*

*Continued from previous page...*

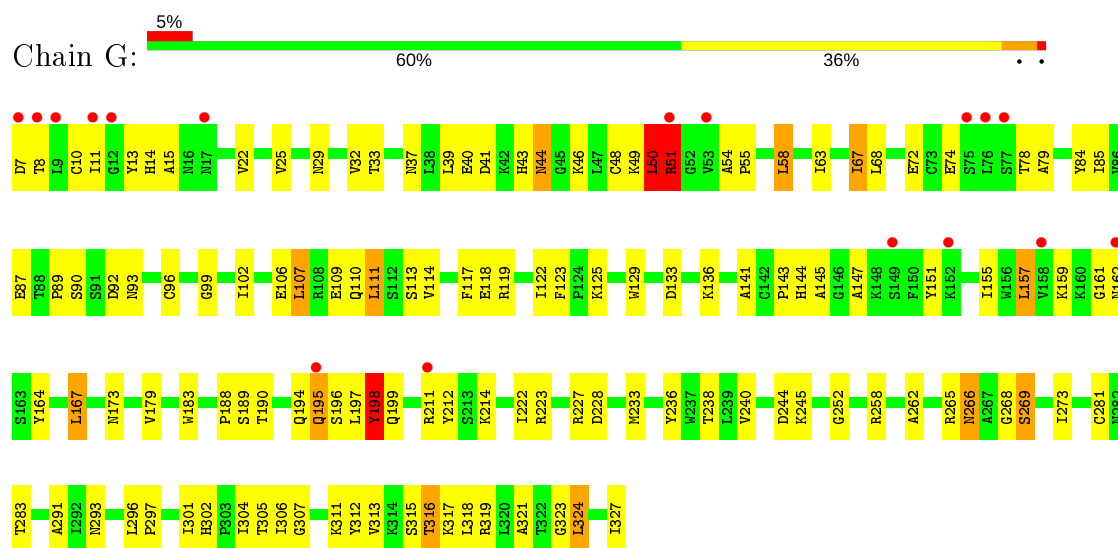
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
8	L	31	Total	O	0	0
			31	31		



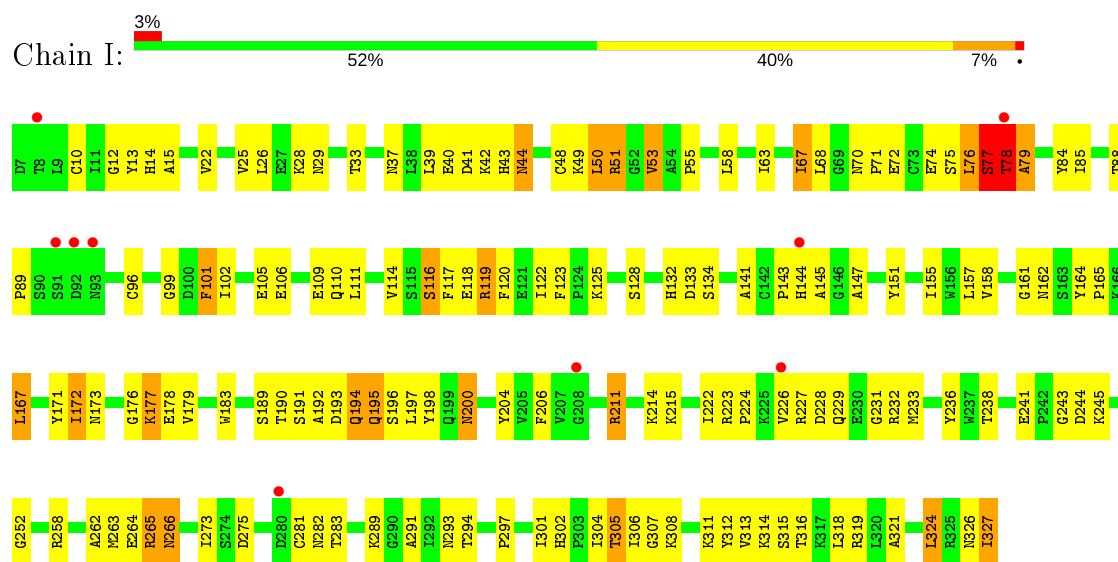




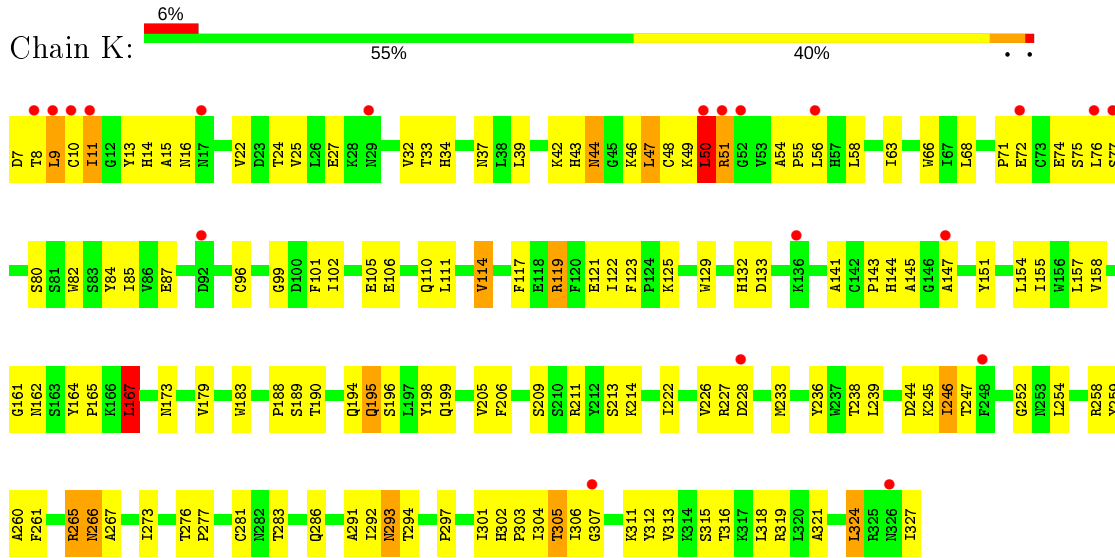
- Molecule 1: Hemagglutinin



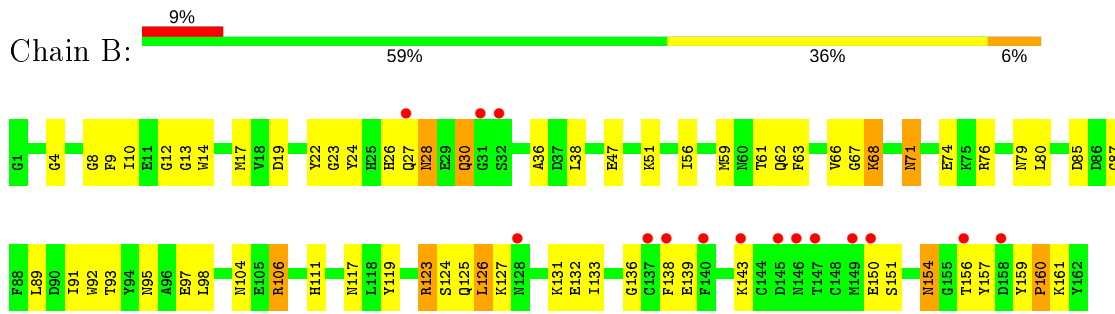
- Molecule 1: Hemagglutinin



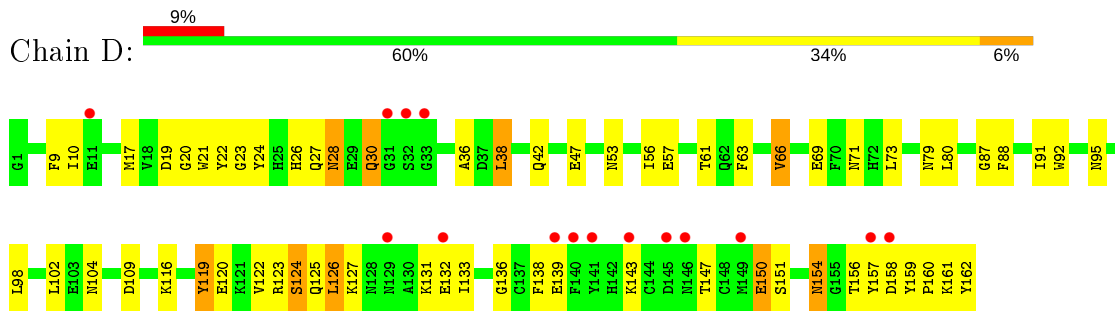
• Molecule 1: Hemagglutinin



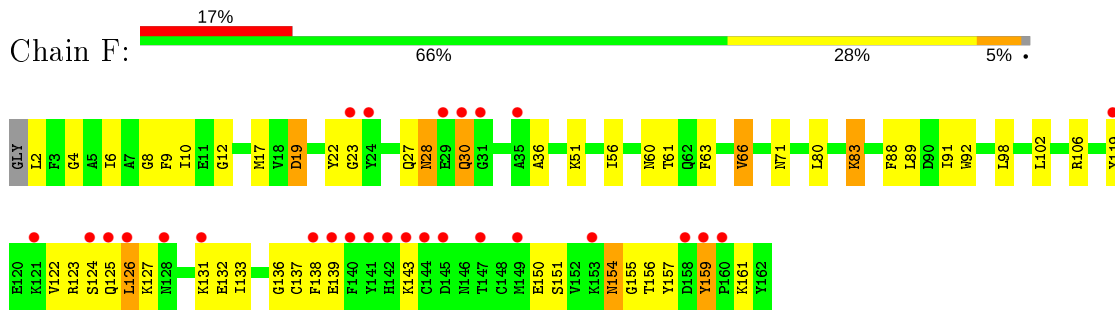
• Molecule 2: Hemagglutinin



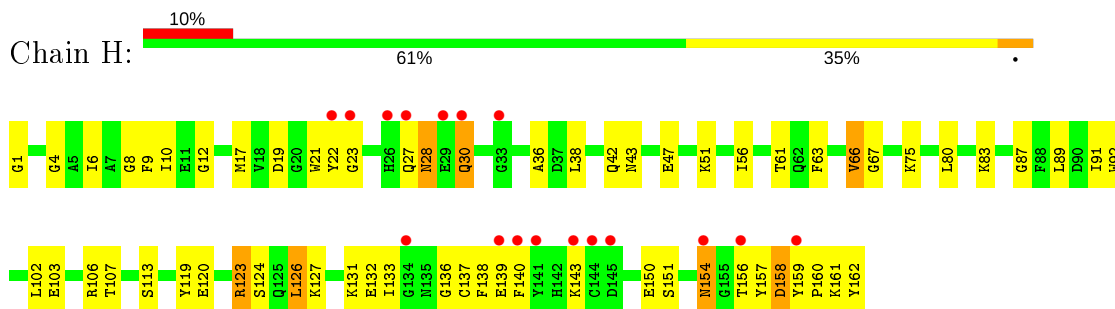
• Molecule 2: Hemagglutinin



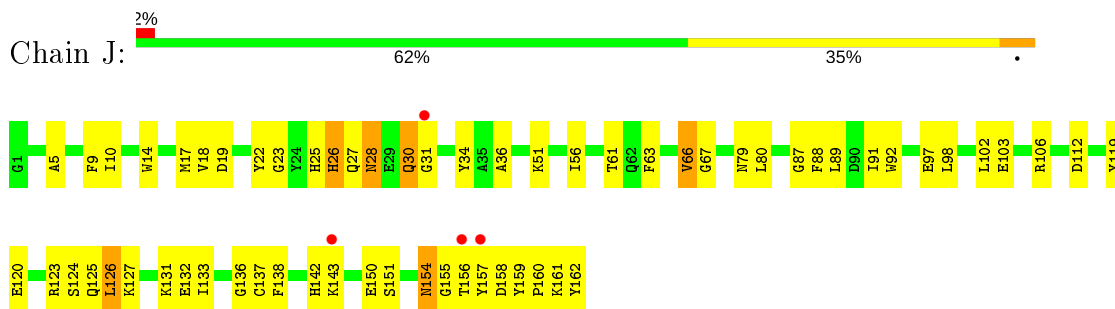
• Molecule 2: Hemagglutinin



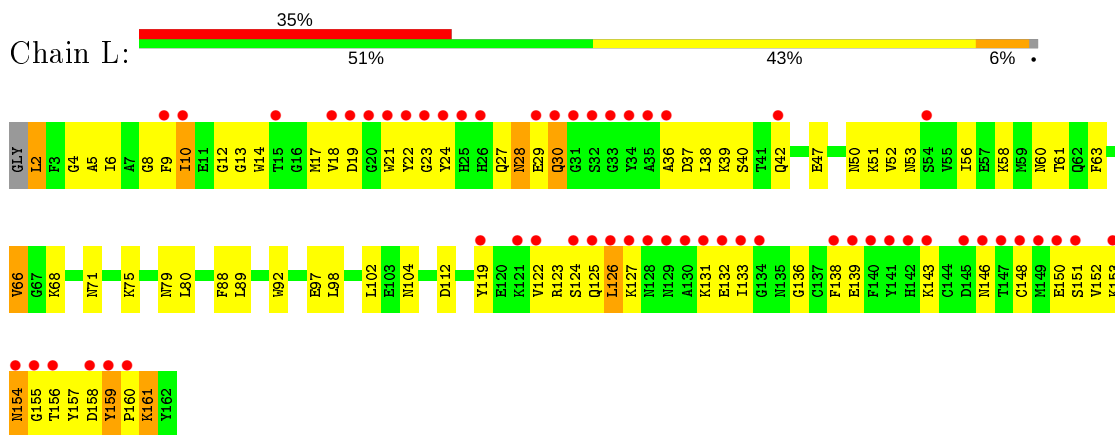
- Molecule 2: Hemagglutinin



- Molecule 2: Hemagglutinin



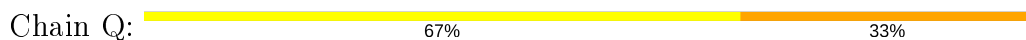
- Molecule 2: Hemagglutinin



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose




NAG1  
GAL2  
SIA3

- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain R:  67% 33%

NAG1  
GAL2  
SIA3

- Molecule 4: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose

Chain N:  75% 25%

GAL1  
NAG2  
GAL3  
SIA4

- Molecule 4: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose

Chain P:  75% 25%

GAL3  
NAG2  
GAL3  
SIA4

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

Chain O:  100%

NAG1  
NAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.72Å 117.32Å 117.39Å 61.78° 81.82° 77.42°	Depositor
Resolution (Å)	38.39 – 3.00 49.68 – 3.00	Depositor EDS
% Data completeness (in resolution range)	81.5 (38.39-3.00) 87.5 (49.68-3.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.37 (at 3.01Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, $R_{free}$	0.224 , 0.275 0.230 , 0.279	Depositor DCC
$R_{free}$ test set	2711 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	55.4	Xtrriage
Anisotropy	0.701	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 58.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	23528	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	84.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.66% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: SIA, GAL, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.57	0/2568	0.75	5/3488 (0.1%)
1	C	0.55	0/2572	0.69	8/3493 (0.2%)
1	E	0.39	0/2572	0.58	4/3493 (0.1%)
1	G	0.33	0/2568	0.77	9/3488 (0.3%)
1	I	0.47	0/2573	0.63	3/3495 (0.1%)
1	K	0.33	0/2572	0.63	5/3493 (0.1%)
2	B	0.48	0/1333	0.54	1/1797 (0.1%)
2	D	0.38	0/1328	0.45	0/1791
2	F	0.26	0/1330	0.46	0/1794
2	H	0.26	0/1333	0.49	1/1797 (0.1%)
2	J	0.43	0/1328	0.48	0/1791
2	L	0.26	0/1330	0.53	1/1794 (0.1%)
All	All	0.42	0/23407	0.62	37/31714 (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
1	I	0	1

There are no bond length outliers.

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	119	ARG	NE-CZ-NH1	-16.56	112.02	120.30
1	A	119	ARG	NE-CZ-NH2	16.50	128.55	120.30
1	A	119	ARG	NE-CZ-NH1	-16.32	112.14	120.30
1	G	119	ARG	NE-CZ-NH2	15.97	128.29	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	212	TYR	N-CA-CB	11.91	132.04	110.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	77	SER	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	2505	0	2444	227	0
1	C	2509	0	2451	186	0
1	E	2509	0	2451	159	4
1	G	2505	0	2445	137	0
1	I	2510	0	2456	183	0
1	K	2509	0	2452	188	0
2	B	1305	0	1228	97	0
2	D	1300	0	1216	81	0
2	F	1302	0	1226	65	0
2	H	1305	0	1228	73	0
2	J	1300	0	1216	89	0
2	L	1302	0	1226	126	4
3	M	45	0	38	4	0
3	Q	45	0	38	3	0
3	R	45	0	38	0	0
4	N	56	0	47	1	0
4	P	56	0	47	2	0
5	O	28	0	25	3	0
6	A	56	0	52	6	0
6	C	14	0	13	2	0
6	G	28	0	26	5	0
7	G	20	0	17	1	0
8	A	40	0	0	59	0
8	B	13	0	0	11	0
8	C	29	0	0	34	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	D	13	0	0	6	0
8	E	31	0	0	27	0
8	F	7	0	0	6	0
8	G	19	0	0	22	0
8	H	13	0	0	22	0
8	I	35	0	0	32	0
8	J	18	0	0	22	0
8	K	25	0	0	47	0
8	L	31	0	0	48	0
All	All	23528	0	22380	1414	4

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 31.

The worst 5 of 1414 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:11:ILE:HD13	2:D:24:TYR:CD2	1.53	1.42
1:I:190:THR:HG21	1:I:193:ASP:OD1	1.24	1.35
1:K:106:GLU:HB3	8:K:915:HOH:O	1.20	1.34
2:L:19:ASP:HA	8:L:228:HOH:O	1.33	1.28
1:A:8:THR:HG22	2:B:138:PHE:O	1.18	1.27

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:119:ARG:NH2	2:L:38:LEU:CD2[1_565]	1.96	0.24
1:E:79:ALA:N	2:L:150:GLU:OE2[1_565]	2.01	0.19
1:E:258:ARG:NH2	2:L:42:GLN:OE1[1_565]	2.04	0.16
1:E:81:SER:OG	2:L:146:ASN:OD1[1_565]	2.12	0.08

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	319/321 (99%)	295 (92%)	23 (7%)	1 (0%)	41	76
1	C	319/321 (99%)	296 (93%)	19 (6%)	4 (1%)	12	45
1	E	319/321 (99%)	295 (92%)	23 (7%)	1 (0%)	41	76
1	G	319/321 (99%)	292 (92%)	27 (8%)	0	100	100
1	I	319/321 (99%)	291 (91%)	26 (8%)	2 (1%)	25	64
1	K	319/321 (99%)	290 (91%)	29 (9%)	0	100	100
2	B	160/162 (99%)	146 (91%)	13 (8%)	1 (1%)	25	64
2	D	160/162 (99%)	143 (89%)	16 (10%)	1 (1%)	25	64
2	F	159/162 (98%)	145 (91%)	14 (9%)	0	100	100
2	H	160/162 (99%)	146 (91%)	14 (9%)	0	100	100
2	J	160/162 (99%)	144 (90%)	16 (10%)	0	100	100
2	L	159/162 (98%)	145 (91%)	13 (8%)	1 (1%)	25	64
All	All	2872/2898 (99%)	2628 (92%)	233 (8%)	11 (0%)	34	72

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	160	PRO
1	A	146	GLY
1	C	77	SER
1	C	78	THR
1	C	161	GLY

### 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	281/282 (100%)	248 (88%)	33 (12%)	5	22
1	C	282/282 (100%)	248 (88%)	34 (12%)	5	21

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	282/282 (100%)	254 (90%)	28 (10%)	8	30
1	G	281/282 (100%)	258 (92%)	23 (8%)	11	39
1	I	282/282 (100%)	247 (88%)	35 (12%)	4	20
1	K	282/282 (100%)	255 (90%)	27 (10%)	8	32
2	B	139/139 (100%)	128 (92%)	11 (8%)	12	41
2	D	138/139 (99%)	125 (91%)	13 (9%)	8	32
2	F	139/139 (100%)	126 (91%)	13 (9%)	8	32
2	H	139/139 (100%)	130 (94%)	9 (6%)	17	50
2	J	138/139 (99%)	128 (93%)	10 (7%)	14	45
2	L	139/139 (100%)	128 (92%)	11 (8%)	12	41
All	All	2522/2526 (100%)	2275 (90%)	247 (10%)	8	30

5 of 247 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	232	ARG
1	G	107	LEU
1	K	228	ASP
1	E	305	THR
2	F	80	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 91 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	282	ASN
1	G	162	ASN
1	K	293	ASN
2	F	27	GLN
2	F	146	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](#)

19 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	M	1	3	14,14,15	0.61	0	17,19,21	0.97	0
3	GAL	M	2	3	11,11,12	0.64	0	15,15,17	0.96	1 (6%)
3	SIA	M	3	3	17,20,21	0.55	0	21,28,31	3.26	4 (19%)
4	GAL	N	1	4	11,11,12	0.72	0	15,15,17	1.02	1 (6%)
4	NAG	N	2	4	14,14,15	0.58	0	17,19,21	1.27	2 (11%)
4	GAL	N	3	4	11,11,12	0.59	0	15,15,17	1.11	1 (6%)
4	SIA	N	4	4	17,20,21	0.39	0	21,28,31	3.51	3 (14%)
5	NAG	O	1	1,5	14,14,15	0.58	0	17,19,21	0.75	0
5	NAG	O	2	5	14,14,15	0.46	0	17,19,21	0.84	0
4	GAL	P	1	4	11,11,12	0.64	0	15,15,17	0.67	0
4	NAG	P	2	4	14,14,15	0.63	0	17,19,21	0.76	0
4	GAL	P	3	4	11,11,12	0.70	0	15,15,17	0.85	0
4	SIA	P	4	4	17,20,21	0.52	0	21,28,31	3.26	4 (19%)
3	NAG	Q	1	3	14,14,15	0.50	0	17,19,21	0.76	0
3	GAL	Q	2	3	11,11,12	0.73	0	15,15,17	1.27	3 (20%)
3	SIA	Q	3	3	17,20,21	0.44	0	21,28,31	3.26	4 (19%)
3	NAG	R	1	3	14,14,15	0.53	0	17,19,21	0.69	0
3	GAL	R	2	3	11,11,12	0.62	0	15,15,17	0.90	0
3	SIA	R	3	3	17,20,21	0.43	0	21,28,31	3.45	5 (23%)

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	M	1	3	-	0/6/23/26	0/1/1/1
3	GAL	M	2	3	-	0/2/19/22	0/1/1/1
3	SIA	M	3	3	-	2/14/34/38	0/1/1/1
4	GAL	N	1	4	-	1/2/19/22	0/1/1/1
4	NAG	N	2	4	-	0/6/23/26	0/1/1/1
4	GAL	N	3	4	-	0/2/19/22	0/1/1/1
4	SIA	N	4	4	-	2/14/34/38	0/1/1/1
5	NAG	O	1	1,5	-	0/6/23/26	0/1/1/1
5	NAG	O	2	5	-	0/6/23/26	0/1/1/1
4	GAL	P	1	4	-	0/2/19/22	0/1/1/1
4	NAG	P	2	4	-	0/6/23/26	0/1/1/1
4	GAL	P	3	4	-	0/2/19/22	0/1/1/1
4	SIA	P	4	4	-	2/14/34/38	0/1/1/1
3	NAG	Q	1	3	-	0/6/23/26	0/1/1/1
3	GAL	Q	2	3	-	0/2/19/22	0/1/1/1
3	SIA	Q	3	3	-	2/14/34/38	0/1/1/1
3	NAG	R	1	3	-	0/6/23/26	0/1/1/1
3	GAL	R	2	3	-	2/2/19/22	0/1/1/1
3	SIA	R	3	3	-	4/14/34/38	0/1/1/1

There are no bond length outliers.

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	N	4	SIA	O6-C2-C3	-11.28	90.09	109.87
3	R	3	SIA	C3-C2-C1	-10.94	88.04	111.93
4	N	4	SIA	C3-C2-C1	-10.81	88.34	111.93
4	P	4	SIA	C3-C2-C1	-10.52	88.97	111.93
3	M	3	SIA	O6-C2-C3	-10.23	91.93	109.87

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

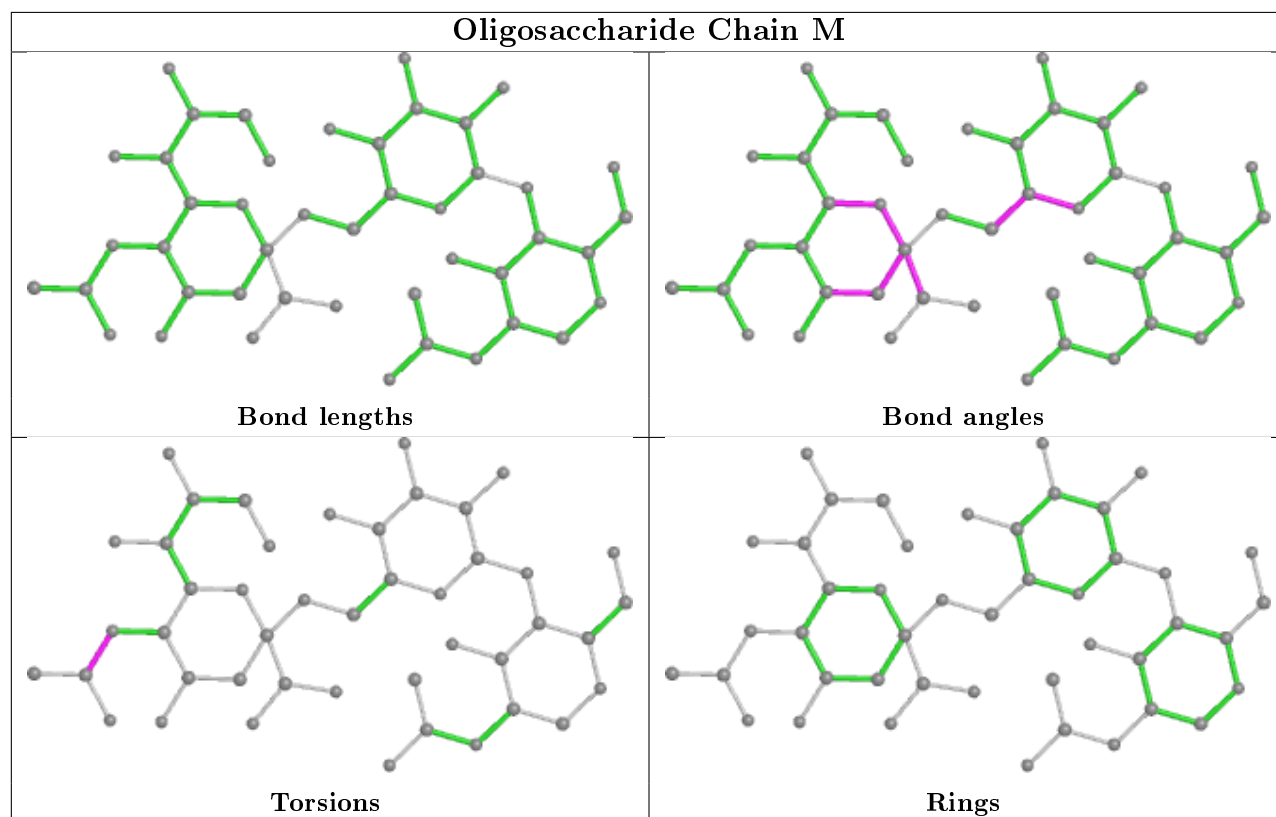
Mol	Chain	Res	Type	Atoms
3	R	2	GAL	C4-C5-C6-O6
3	M	3	SIA	C11-C10-N5-C5
3	M	3	SIA	O10-C10-N5-C5
3	R	3	SIA	C11-C10-N5-C5
3	R	3	SIA	O10-C10-N5-C5

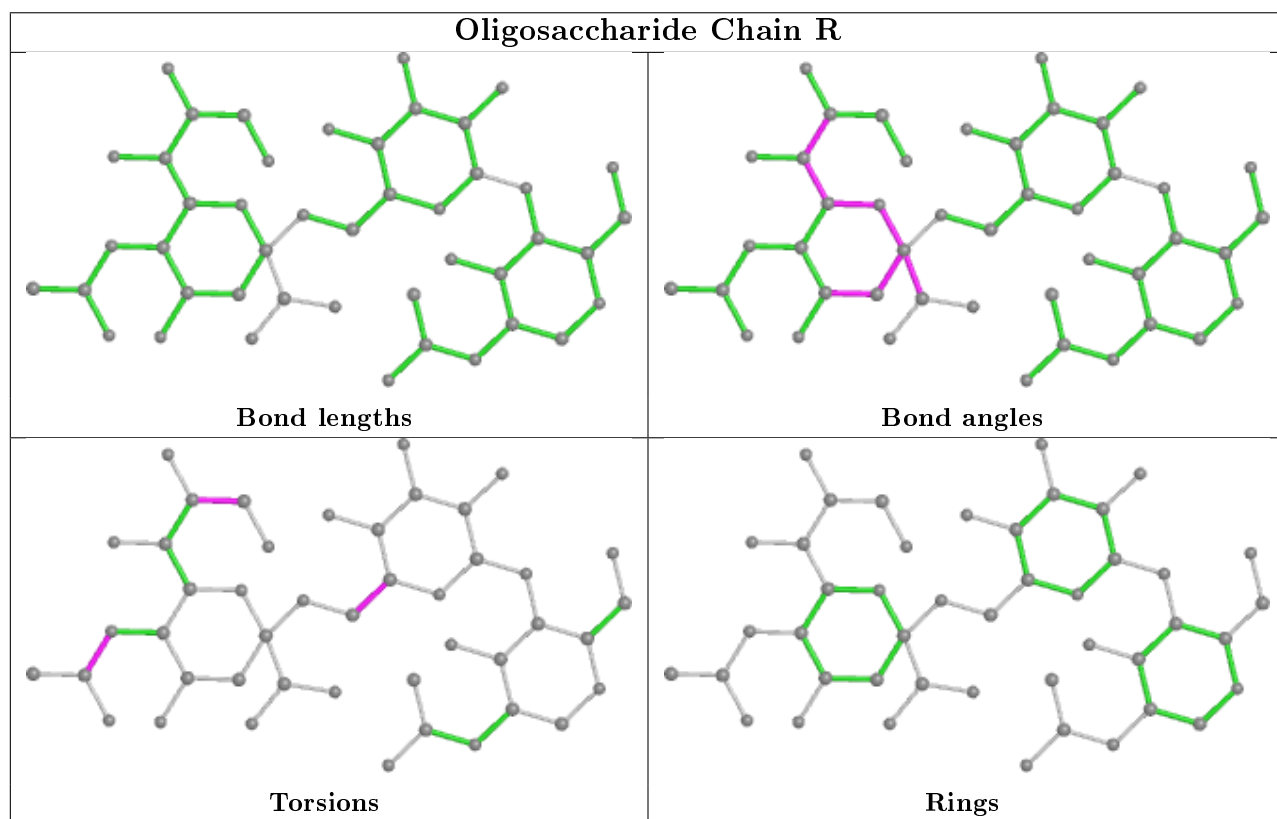
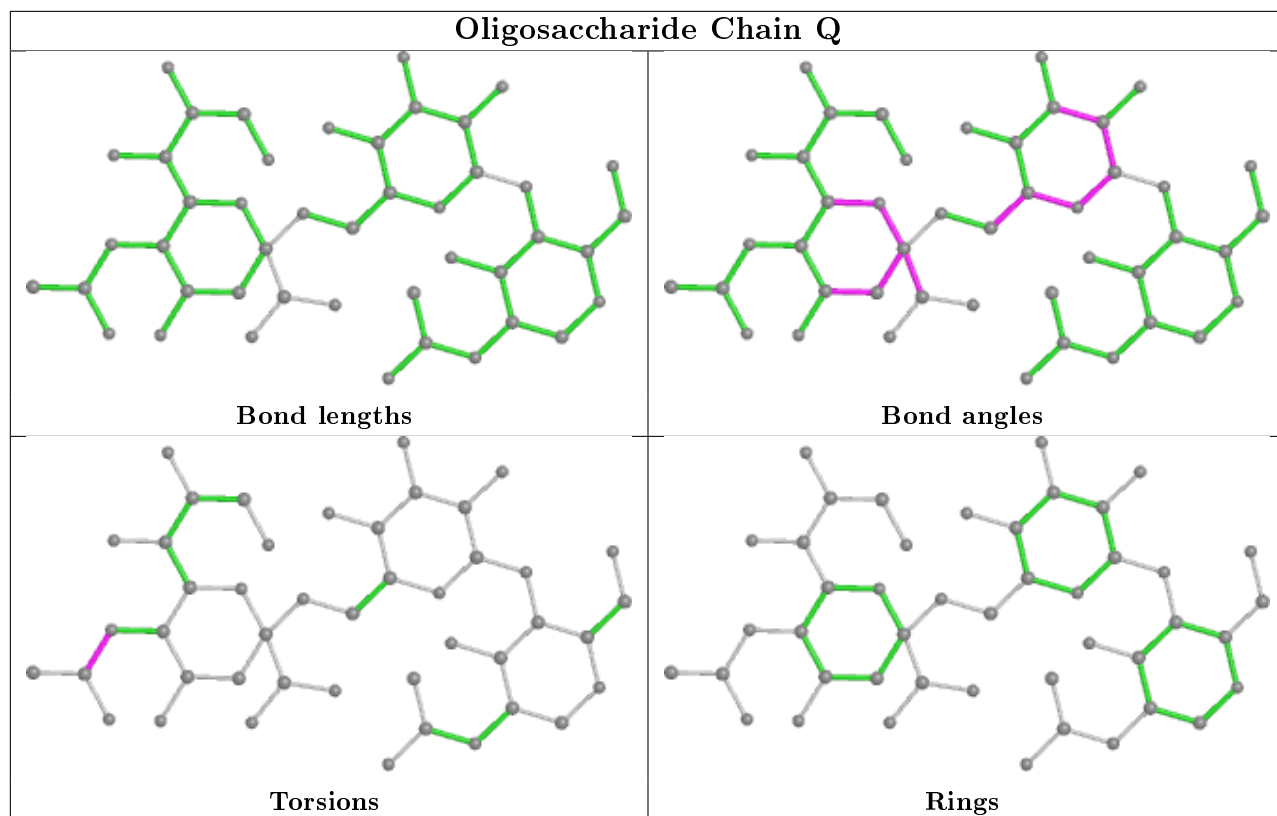
There are no ring outliers.

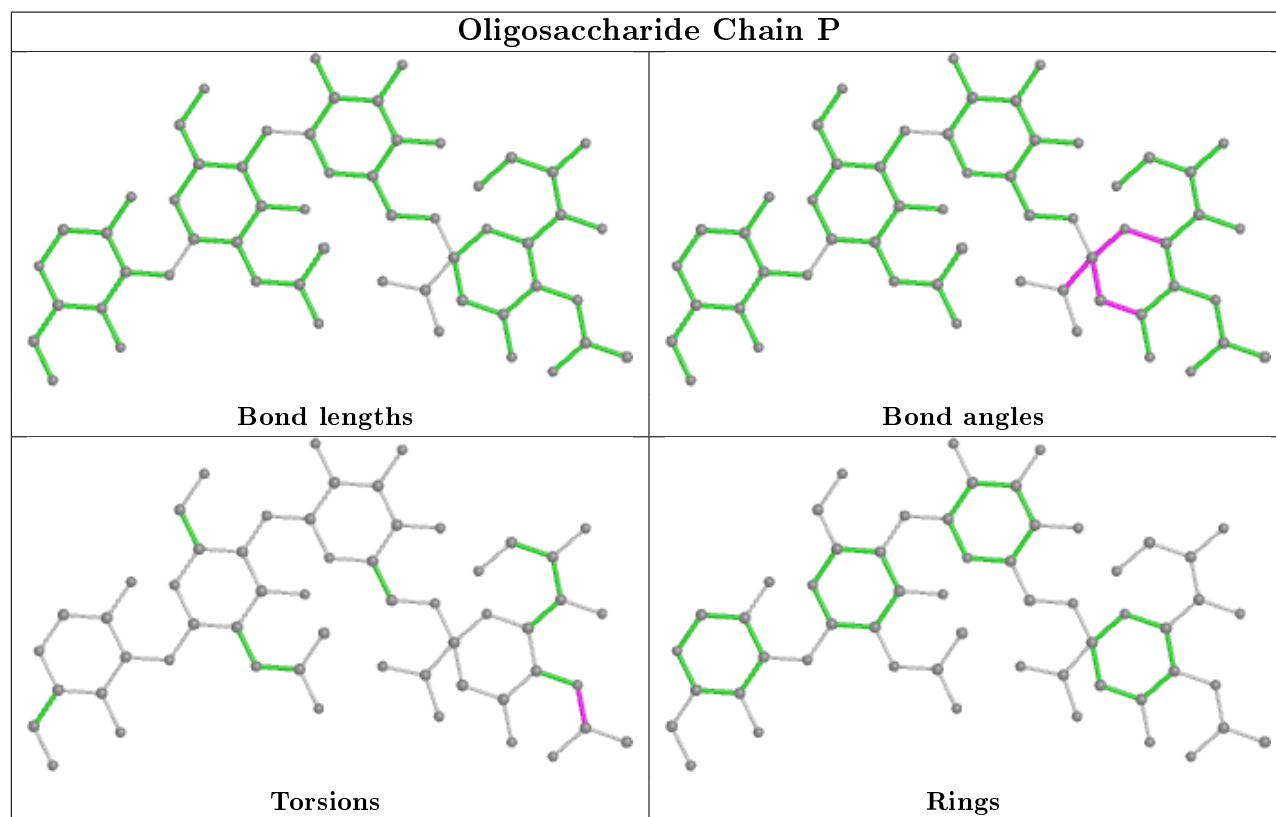
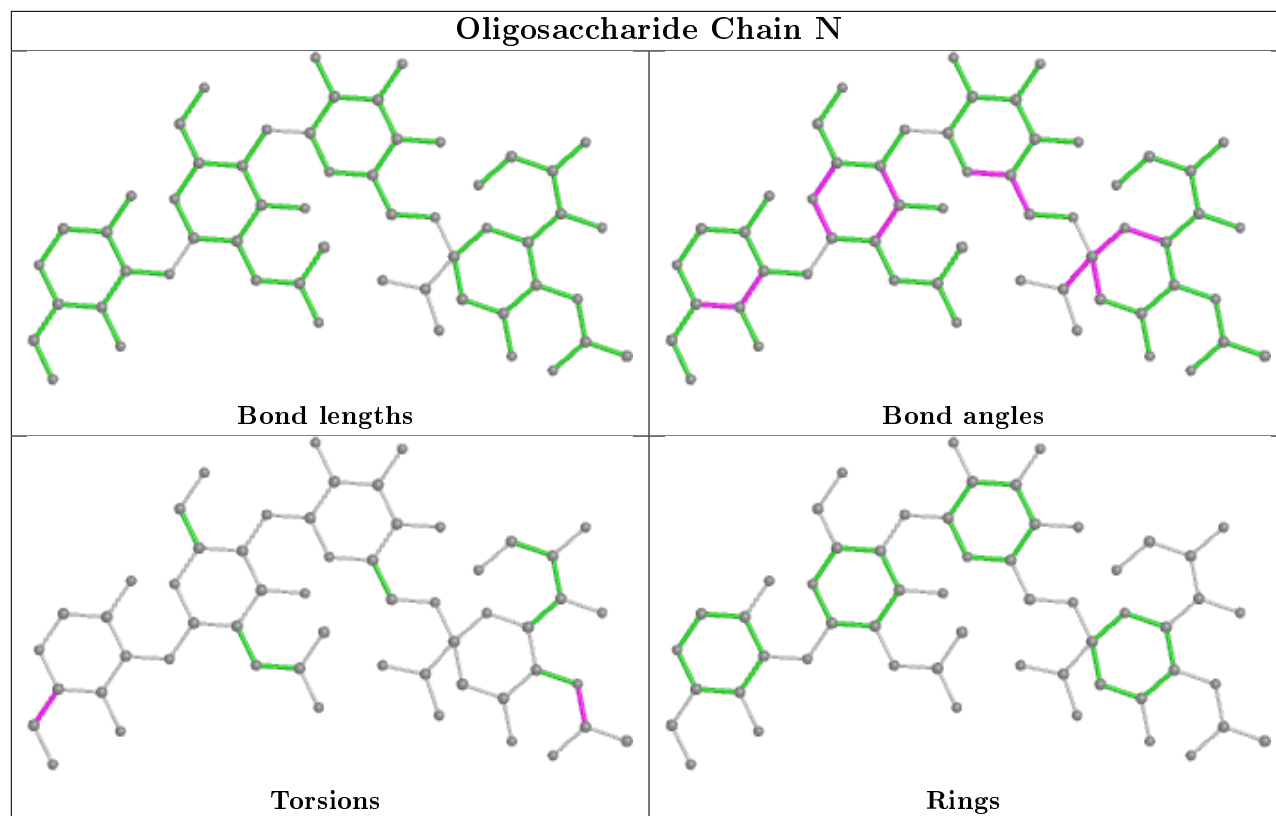
9 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	M	2	GAL	1	0
5	O	1	NAG	3	0
5	O	2	NAG	1	0
3	M	3	SIA	1	0
4	N	4	SIA	1	0
4	P	4	SIA	2	0
3	Q	1	NAG	1	0
3	M	1	NAG	2	0
3	Q	3	SIA	3	0

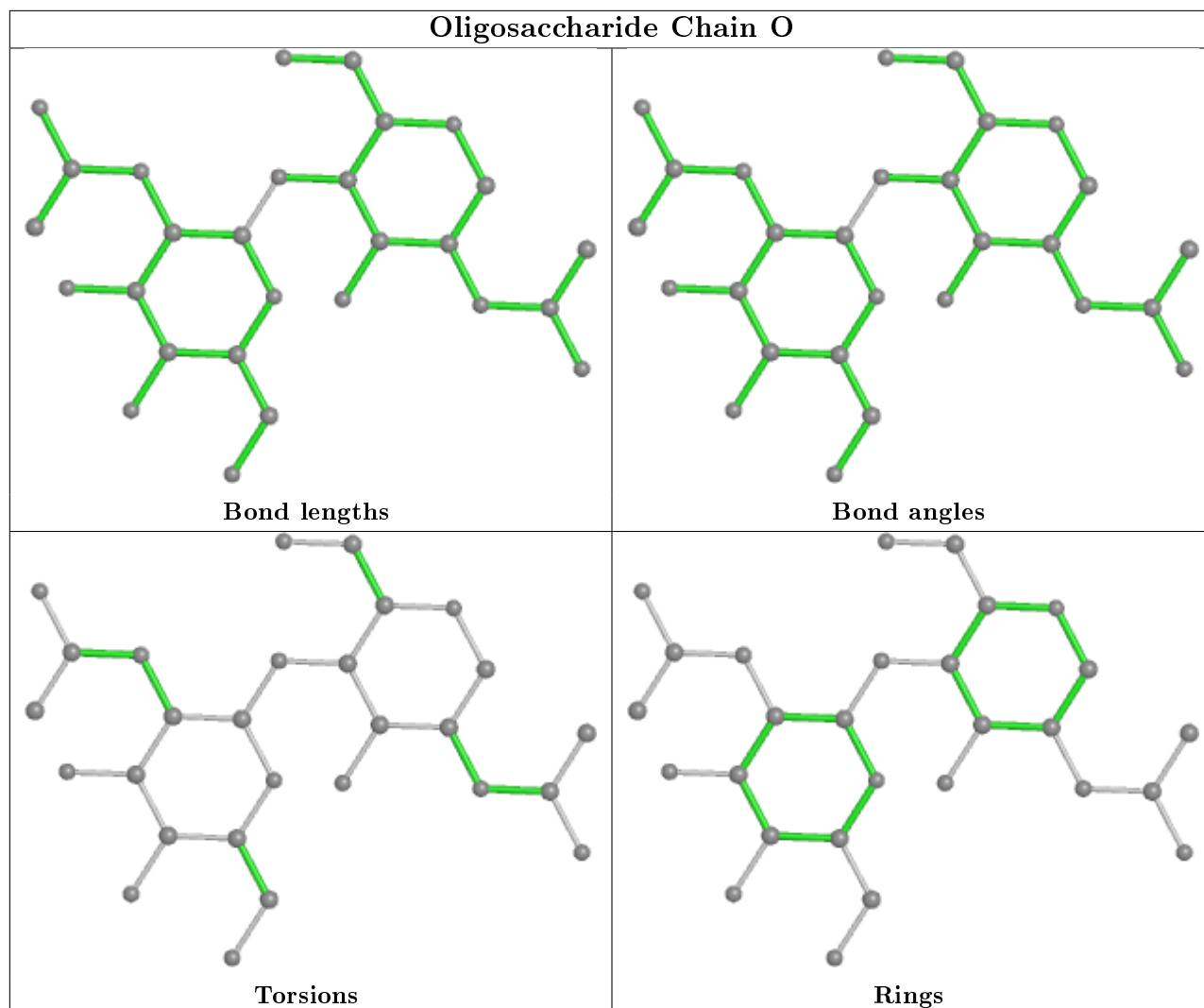
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











## 5.6 Ligand geometry [i](#)

8 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
6	NAG	A	604	1	14,14,15	0.57	0	17,19,21	0.66	0
6	NAG	A	603	1	14,14,15	0.54	0	17,19,21	0.60	0
6	NAG	G	601	1	14,14,15	0.29	0	17,19,21	0.62	0
6	NAG	A	601	1	14,14,15	0.52	0	17,19,21	0.84	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	SIA	G	603	-	17,20,21	0.44	0	21,28,31	1.34	2 (9%)
6	NAG	A	602	1	14,14,15	0.57	0	17,19,21	0.65	0
6	NAG	G	602	1	14,14,15	0.77	0	17,19,21	1.66	3 (17%)
6	NAG	C	601	1	14,14,15	0.46	0	17,19,21	1.04	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	A	604	1	-	2/6/23/26	0/1/1/1
6	NAG	A	603	1	-	0/6/23/26	0/1/1/1
6	NAG	G	601	1	-	2/6/23/26	0/1/1/1
6	NAG	A	601	1	-	2/6/23/26	0/1/1/1
7	SIA	G	603	-	-	3/14/34/38	0/1/1/1
6	NAG	A	602	1	-	1/6/23/26	0/1/1/1
6	NAG	G	602	1	-	1/6/23/26	0/1/1/1
6	NAG	C	601	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	602	NAG	C2-N2-C7	4.05	128.67	122.90
7	G	603	SIA	C6-O6-C2	3.56	118.95	111.34
7	G	603	SIA	O6-C2-C3	3.48	115.98	109.87
6	C	601	NAG	C1-O5-C5	3.17	116.49	112.19
6	G	602	NAG	C1-C2-N2	2.42	114.62	110.49

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

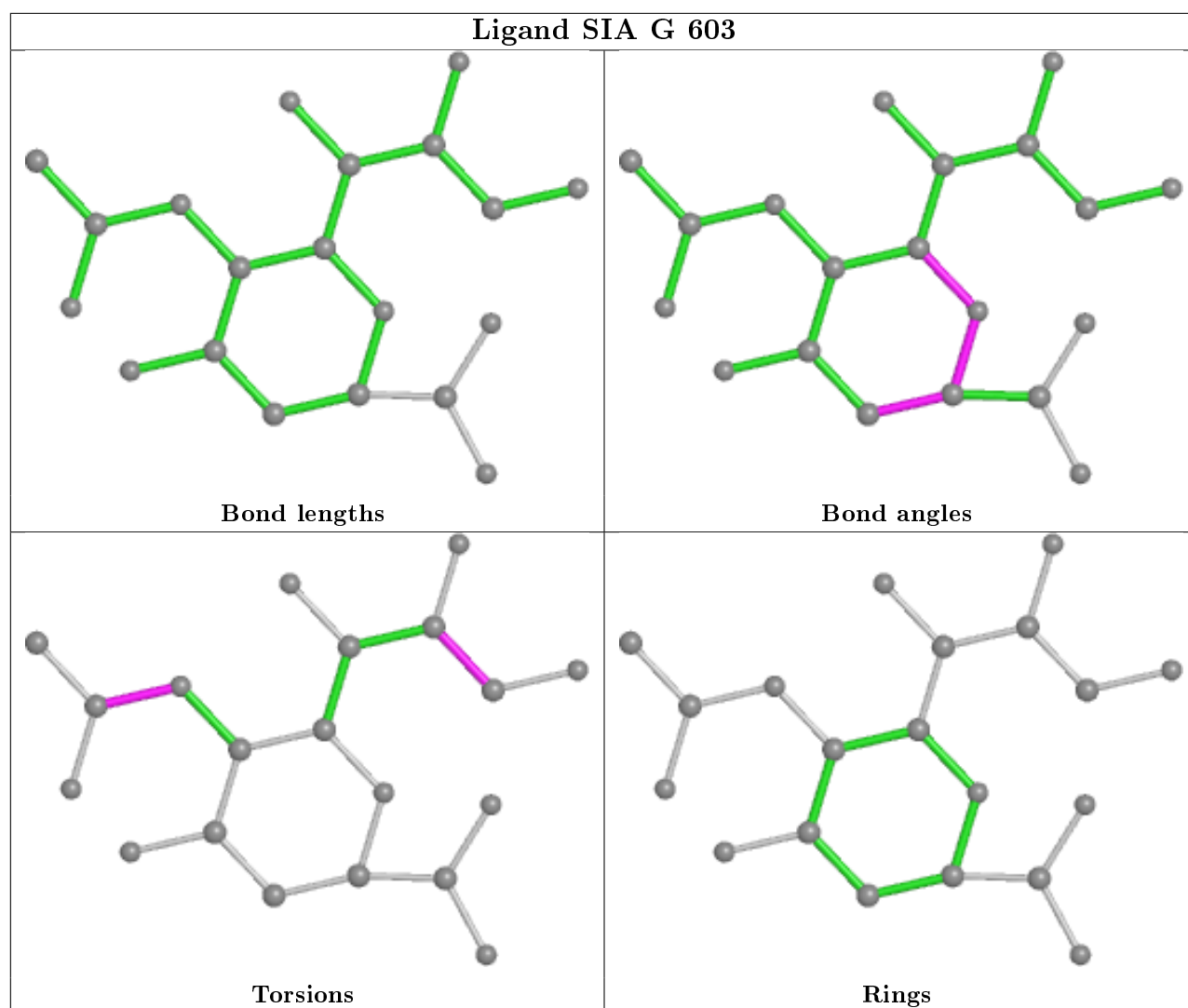
Mol	Chain	Res	Type	Atoms
6	G	602	NAG	C1-C2-N2-C7
6	C	601	NAG	O5-C5-C6-O6
6	A	601	NAG	O5-C5-C6-O6
6	A	601	NAG	C4-C5-C6-O6
6	A	604	NAG	O5-C5-C6-O6

There are no ring outliers.

6 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	603	NAG	2	0
6	G	601	NAG	1	0
7	G	603	SIA	1	0
6	A	602	NAG	4	0
6	G	602	NAG	4	0
6	C	601	NAG	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



## 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	321/321 (100%)	0.14	10 (3%) 49 21	35, 69, 111, 154	0
1	C	321/321 (100%)	0.22	18 (5%) 24 8	36, 68, 118, 223	0
1	E	321/321 (100%)	0.15	7 (2%) 62 33	41, 71, 121, 203	0
1	G	321/321 (100%)	0.29	17 (5%) 26 10	39, 76, 124, 154	0
1	I	321/321 (100%)	0.25	9 (2%) 53 25	39, 75, 117, 169	0
1	K	321/321 (100%)	0.48	20 (6%) 20 7	52, 91, 136, 187	0
2	B	162/162 (100%)	0.41	15 (9%) 8 3	35, 84, 142, 180	0
2	D	162/162 (100%)	0.51	15 (9%) 8 3	39, 86, 163, 220	0
2	F	161/162 (99%)	0.72	27 (16%) 1 0	43, 86, 160, 246	0
2	H	162/162 (100%)	0.46	17 (10%) 6 2	40, 85, 149, 181	0
2	J	162/162 (100%)	0.26	4 (2%) 57 29	47, 80, 129, 156	0
2	L	161/162 (99%)	1.82	56 (34%) 0 0	43, 101, 223, 362	0
All	All	2896/2898 (99%)	0.40	215 (7%) 14 4	35, 78, 142, 362	0

The worst 5 of 215 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	L	129	ASN	11.9
2	L	146	ASN	10.3
2	L	130	ALA	9.2
2	L	147	THR	9.0
2	L	149	MET	8.5

### 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

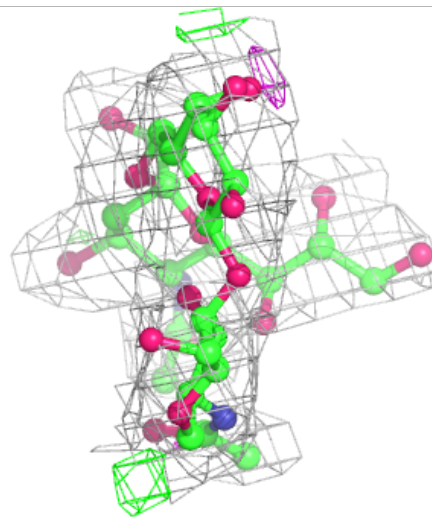
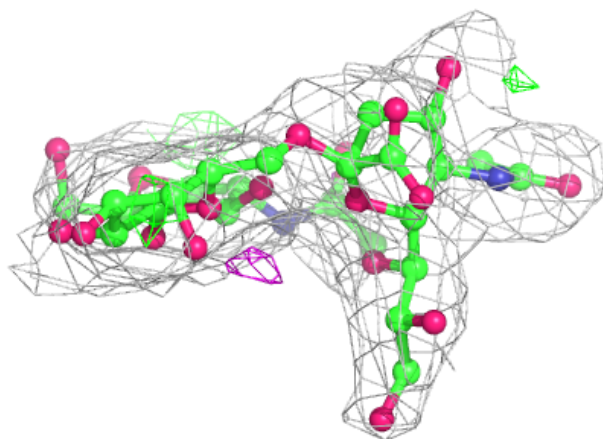
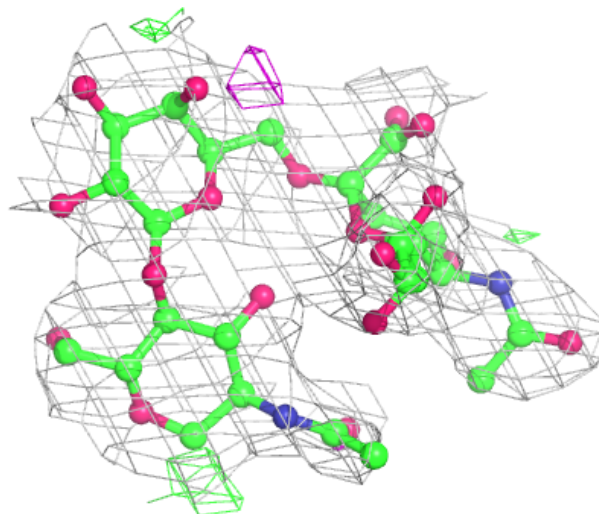
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<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
4	GAL	N	1	11/12	0.64	0.26	104,129,151,153	0
4	GAL	P	1	11/12	0.75	0.26	110,139,162,174	0
5	NAG	O	2	14/15	0.79	0.34	88,113,132,144	0
3	GAL	R	2	11/12	0.80	0.23	96,131,149,169	0
3	NAG	R	1	14/15	0.83	0.23	110,146,159,172	0
4	NAG	N	2	14/15	0.83	0.25	75,106,127,128	0
3	NAG	M	1	14/15	0.85	0.23	65,80,89,96	0
4	NAG	P	2	14/15	0.86	0.32	120,126,146,147	0
3	NAG	Q	1	14/15	0.86	0.27	111,129,140,142	0
3	GAL	Q	2	11/12	0.87	0.23	61,95,108,114	0
4	SIA	P	4	20/21	0.92	0.22	42,65,87,99	0
5	NAG	O	1	14/15	0.92	0.16	74,82,98,102	0
4	SIA	N	4	20/21	0.93	0.20	48,71,90,97	0
3	GAL	M	2	11/12	0.93	0.12	53,93,104,109	0
3	SIA	Q	3	20/21	0.93	0.18	61,77,95,96	0
4	GAL	P	3	11/12	0.93	0.15	78,95,110,122	0
3	SIA	R	3	20/21	0.94	0.21	64,83,91,91	0
4	GAL	N	3	11/12	0.94	0.17	56,84,92,100	0
3	SIA	M	3	20/21	0.94	0.17	37,53,76,80	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.

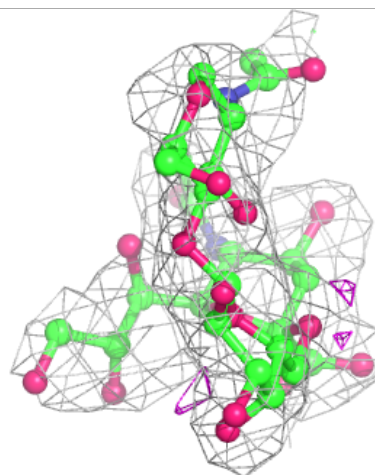
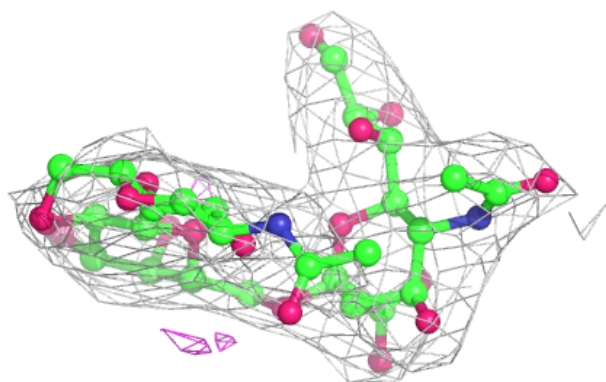
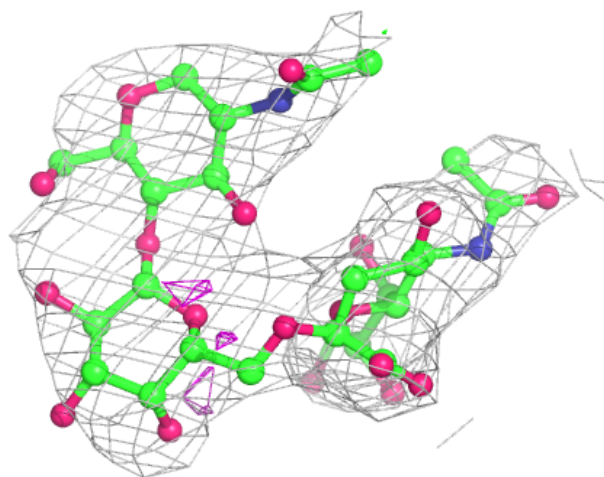
**Electron density around Chain M:**

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



**Electron density around Chain Q:**

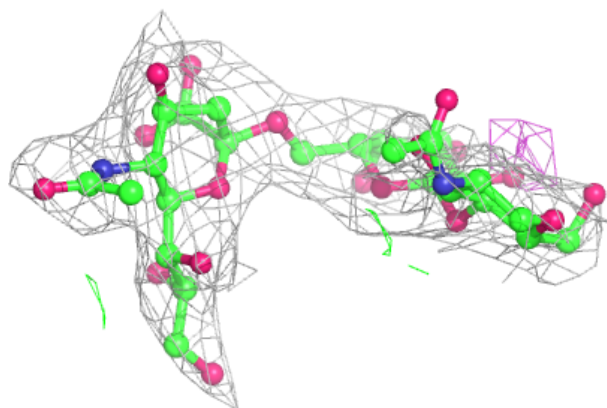
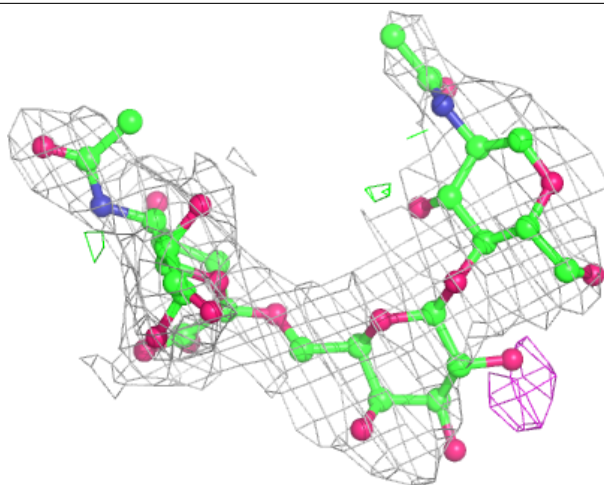
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





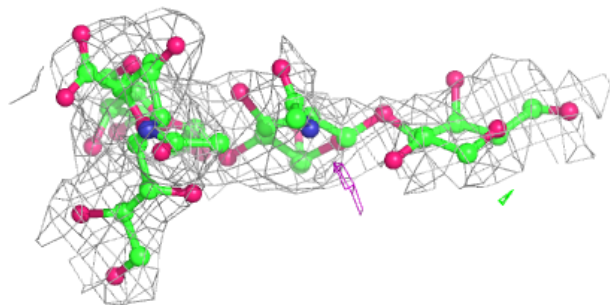
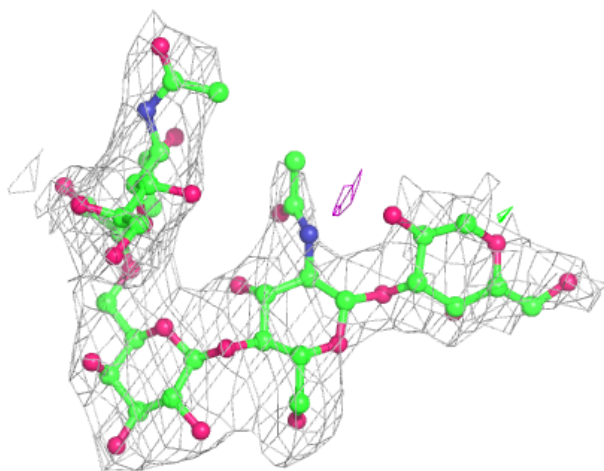
**Electron density around Chain R:**

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



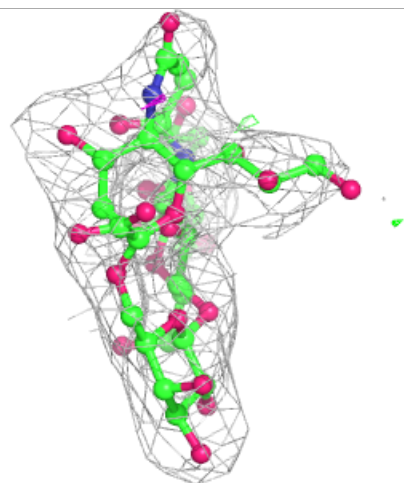
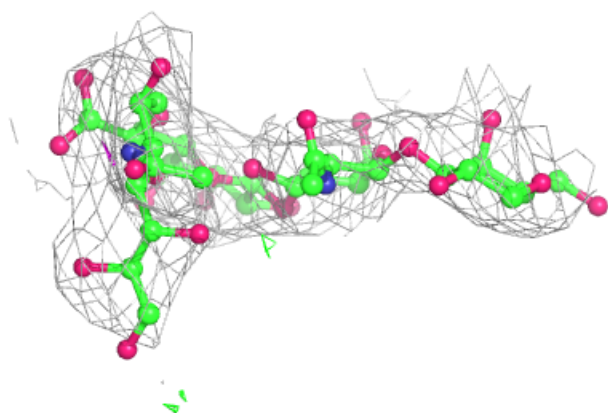
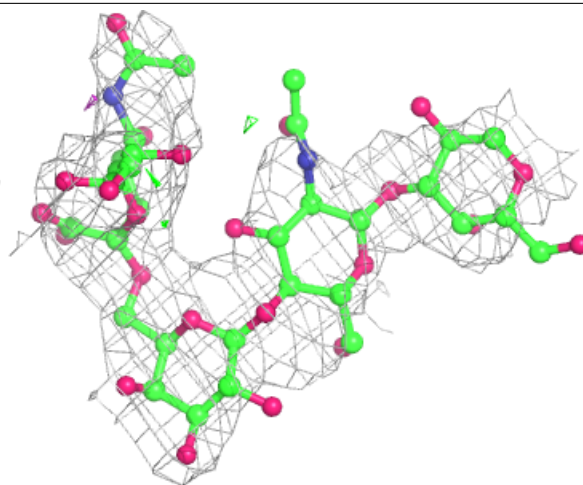
**Electron density around Chain N:**

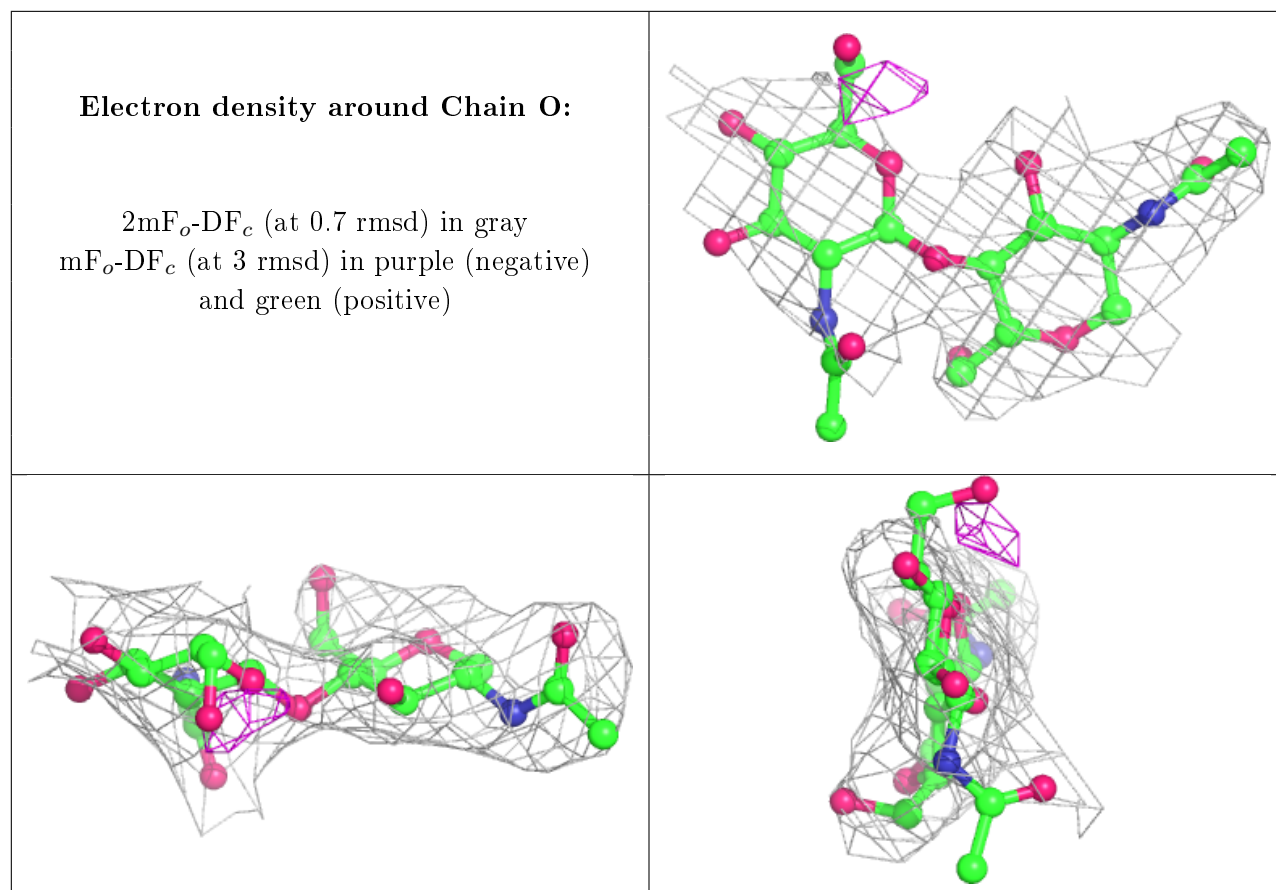
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain P:**

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



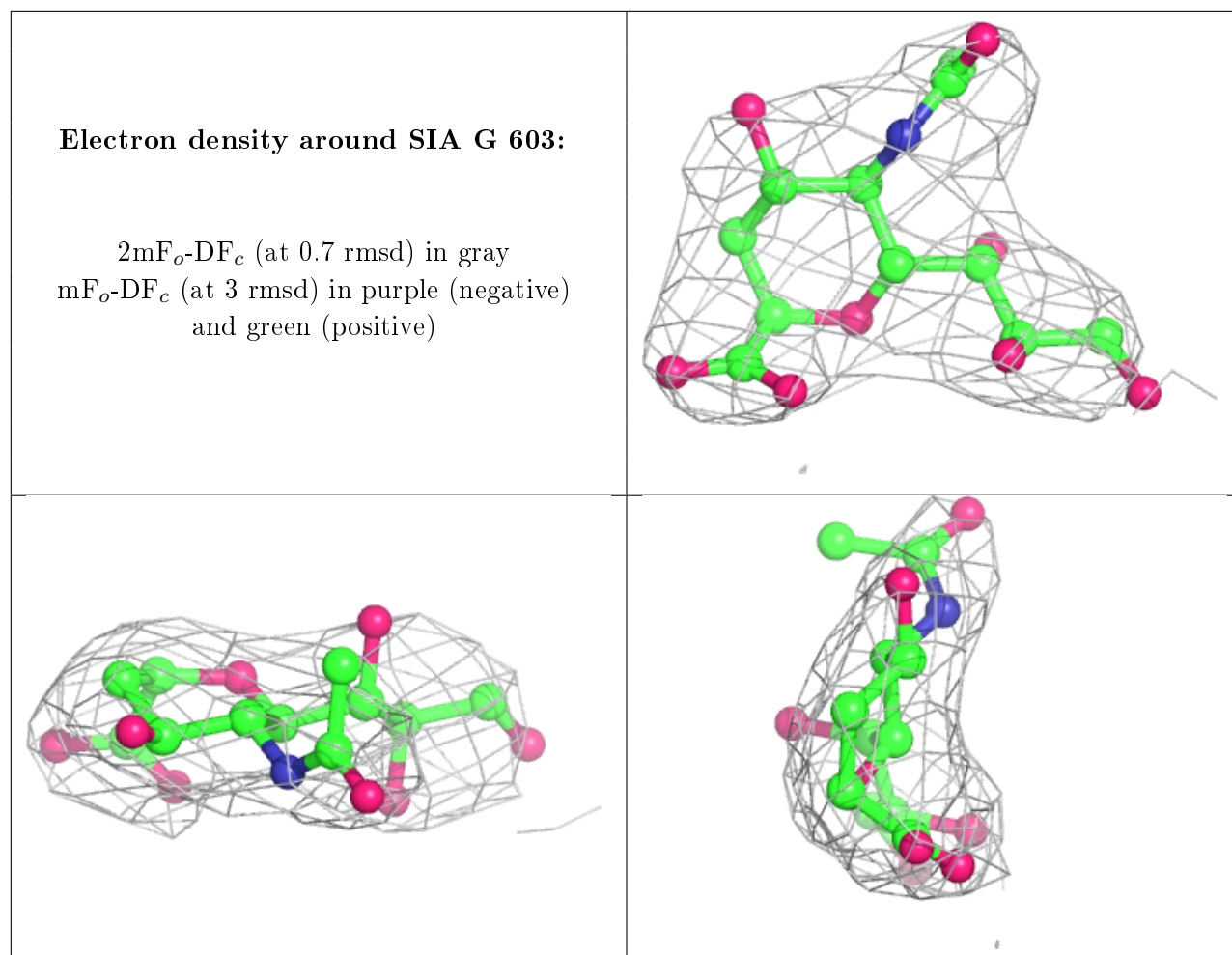


## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	NAG	G	601	14/15	0.52	0.48	119,147,162,173	0
6	NAG	A	601	14/15	0.67	0.35	142,168,177,178	0
6	NAG	A	602	14/15	0.69	0.30	146,167,182,187	0
6	NAG	A	604	14/15	0.72	0.26	99,145,158,161	0
6	NAG	C	601	14/15	0.81	0.23	70,86,97,103	0
6	NAG	G	602	14/15	0.82	0.28	100,121,147,149	0
7	SIA	G	603	20/21	0.89	0.28	65,86,119,122	0
6	NAG	A	603	14/15	0.89	0.27	70,102,144,146	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.