wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 18, 2024 – 07:25 PM EDT

PDB ID	:	4JU0
Title	:	Crystal structure of 2009 pandemic influenza virus hemagglutinin mutant D225E 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	:	2.91 Å(reported)

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

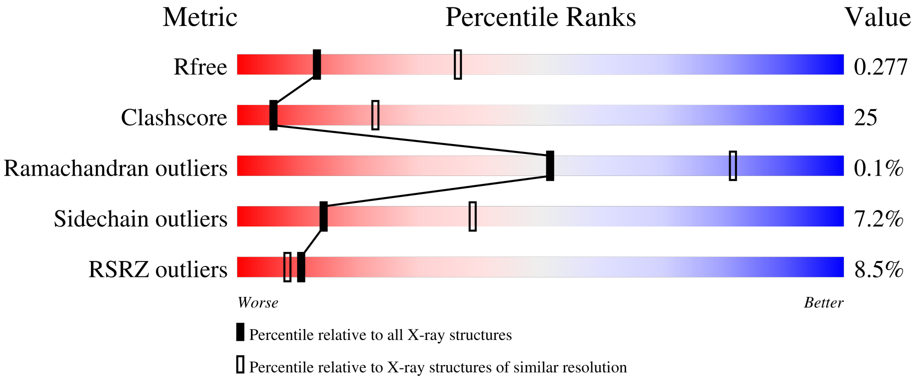
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

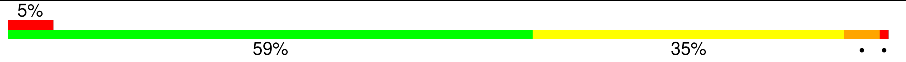
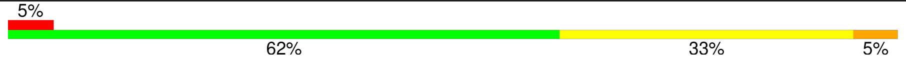
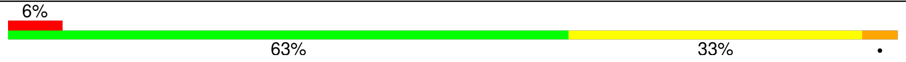
The reported resolution of this entry is 2.91 Å.

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.



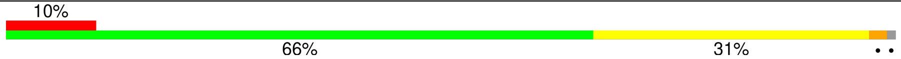
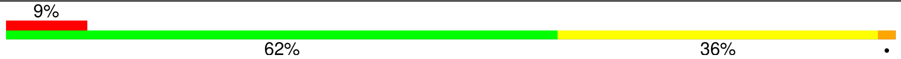
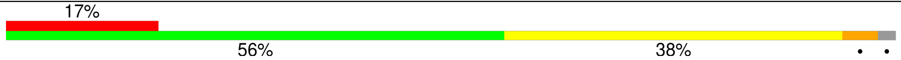
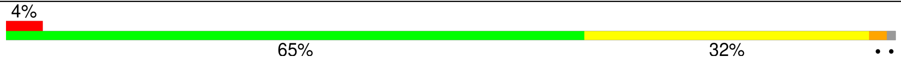
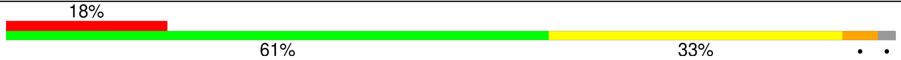
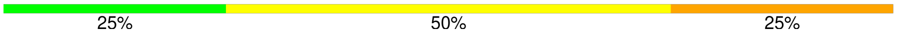
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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 >=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 <=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	322	
1	C	322	
1	E	322	
1	G	322	
1	I	322	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	322	
2	B	164	
2	D	164	
2	F	164	
2	H	164	
2	J	164	
2	L	164	
3	M	4	
3	N	4	
3	P	4	
3	Q	4	
3	R	4	
4	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
3	GAL	P	1	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 23507 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Hemagglutinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	322	2513	1590	434	478	11	0	0	0
1	C	321	2510	1587	433	479	11	0	0	0
1	E	321	2510	1587	433	479	11	0	0	0
1	G	322	2513	1590	434	478	11	0	0	0
1	I	321	2510	1587	433	479	11	0	0	0
1	K	321	2510	1587	433	479	11	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	228	GLU	ASP	engineered mutation	UNP C3W5S1
C	228	GLU	ASP	engineered mutation	UNP C3W5S1
E	228	GLU	ASP	engineered mutation	UNP C3W5S1
G	228	GLU	ASP	engineered mutation	UNP C3W5S1
I	228	GLU	ASP	engineered mutation	UNP C3W5S1
K	228	GLU	ASP	engineered mutation	UNP C3W5S1

- Molecule 2 is a protein called Hemagglutinin.

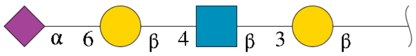
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	162	1305	822	220	257	6	0	0	0
2	D	164	1315	826	221	262	6	0	0	0
2	F	161	1302	821	219	256	6	0	0	0

Continued on next page...

Continued from previous page...

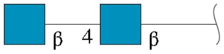
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	162	Total 1305	C 822	N 220	O 257	S 6	0	0	0
2	J	164	Total 1315	C 826	N 221	O 262	S 6	0	0	0
2	L	161	Total 1302	C 821	N 219	O 256	S 6	0	0	0

- 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-(1-3)-beta-D-galactopyranose.



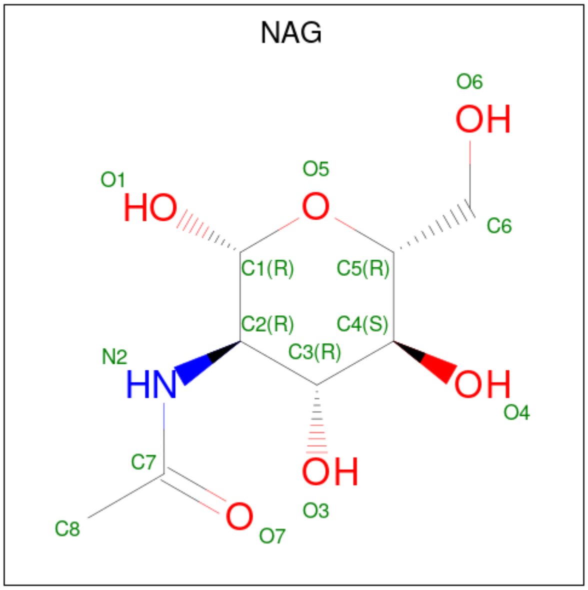
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	M	4	Total 56	C 31	N 2	O 23	0	0	0
3	N	4	Total 56	C 31	N 2	O 23	0	0	0
3	P	4	Total 56	C 31	N 2	O 23	0	0	0
3	Q	4	Total 56	C 31	N 2	O 23	0	0	0
3	R	4	Total 56	C 31	N 2	O 23	0	0	0

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



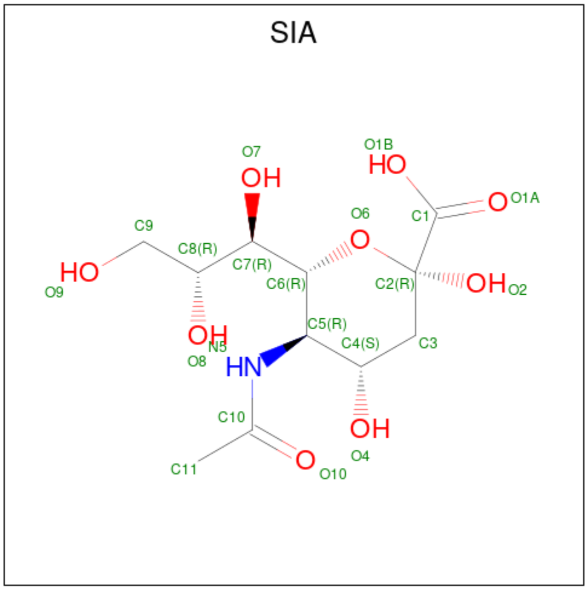
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	O	2	Total 28	C 16	N 2	O 10	0	0	0

- 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	C	1	Total C N O 14 8 1 5	0	0
5	C	1	Total C N O 14 8 1 5	0	0
5	G	1	Total C N O 14 8 1 5	0	0
5	I	1	Total C N O 14 8 1 5	0	0
5	K	1	Total C N O 14 8 1 5	0	0
5	K	1	Total C N O 14 8 1 5	0	0

- Molecule 6 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	K	1	Total C N O 20 11 1 8	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	22	Total O 22 22	0	0
7	B	6	Total O 6 6	0	0
7	C	20	Total O 20 20	0	0
7	D	5	Total O 5 5	0	0
7	E	18	Total O 18 18	0	0
7	F	11	Total O 11 11	0	0
7	G	16	Total O 16 16	0	0
7	H	6	Total O 6 6	0	0
7	I	25	Total O 25 25	0	0
7	J	22	Total O 22 22	0	0
7	K	26	Total O 26 26	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	L	8	Total O 8 8	0	0

3 Residue-property plots ⓘ

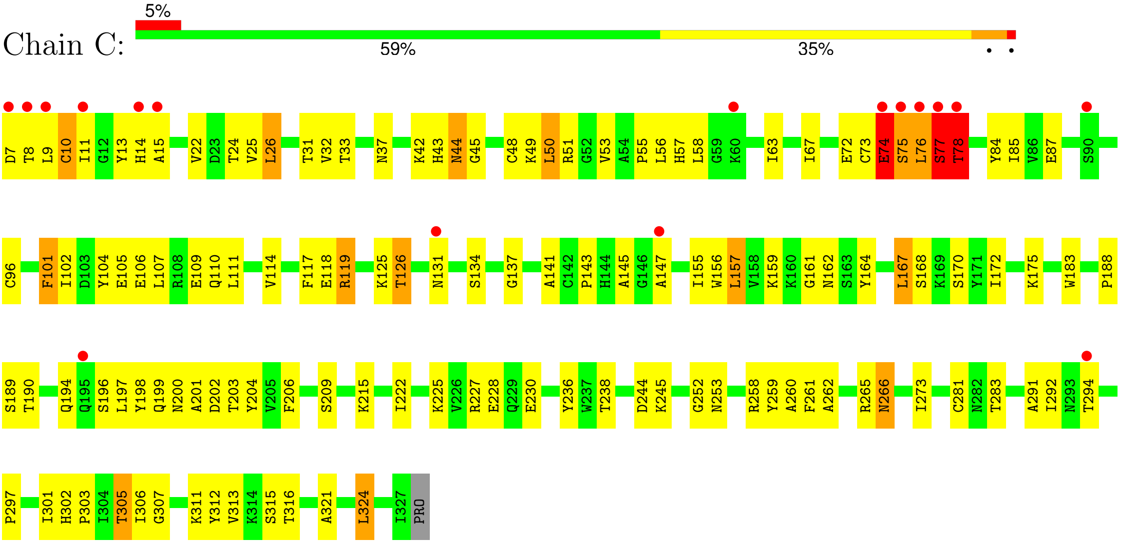
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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

Chain A:

- Molecule 1: Hemagglutinin

Chain C:

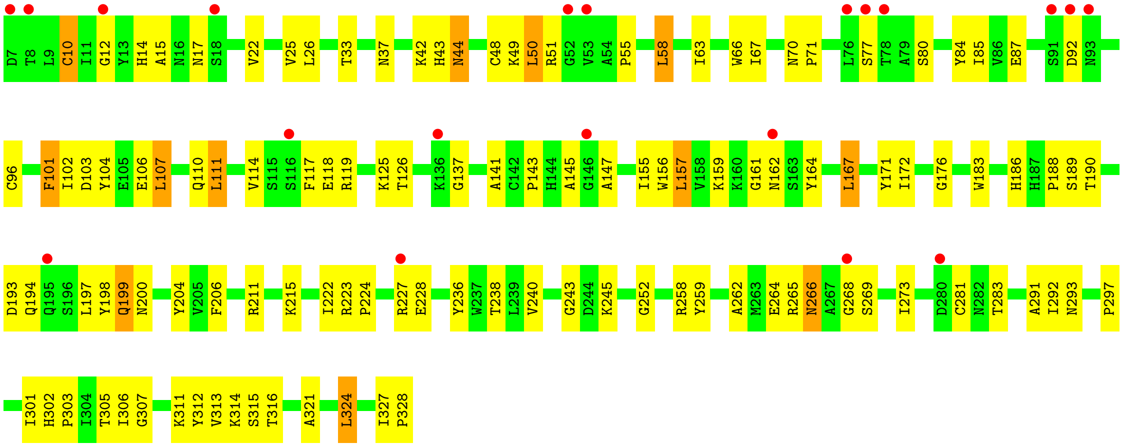


- Molecule 1: Hemagglutinin

Chain E:

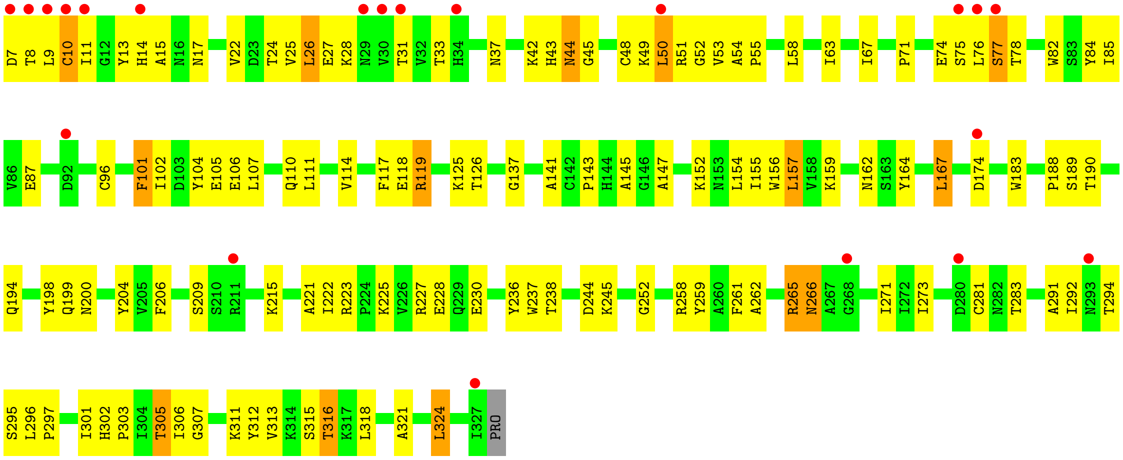
- Molecule 1: Hemagglutinin

Chain G:

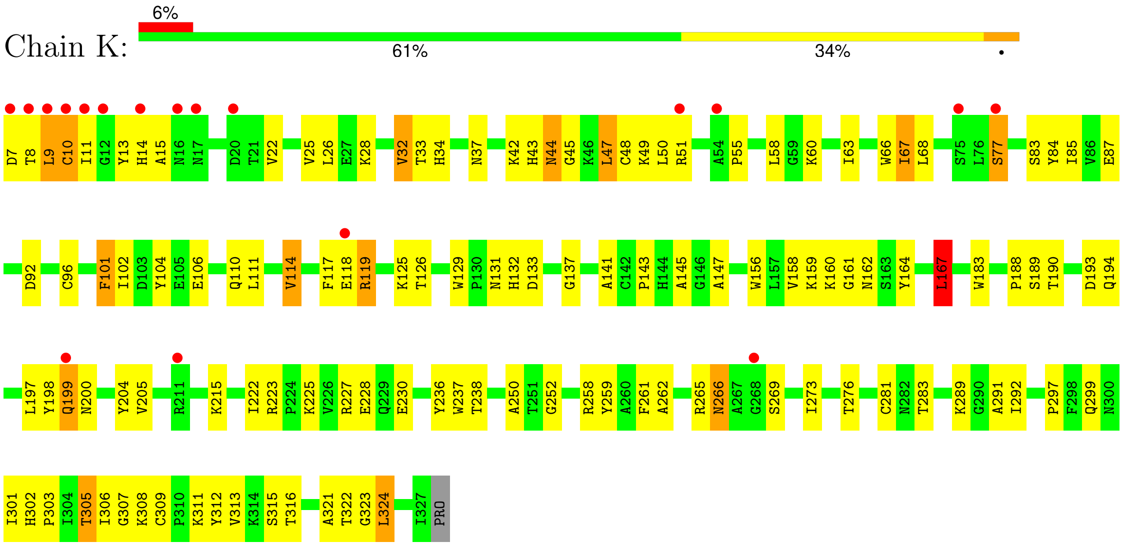


- Molecule 1: Hemagglutinin

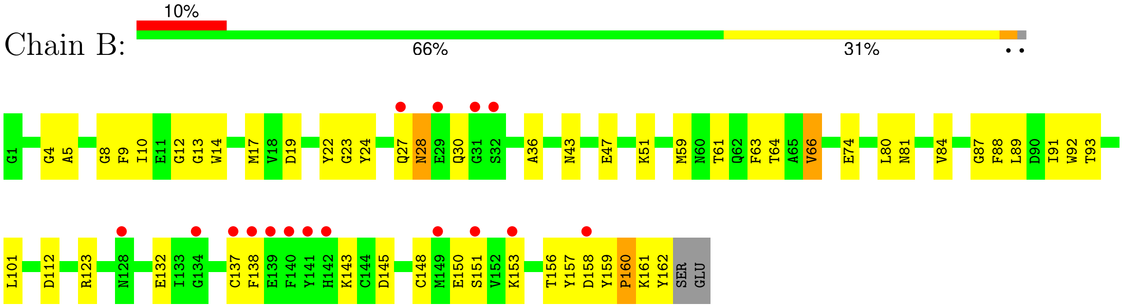
Chain I:



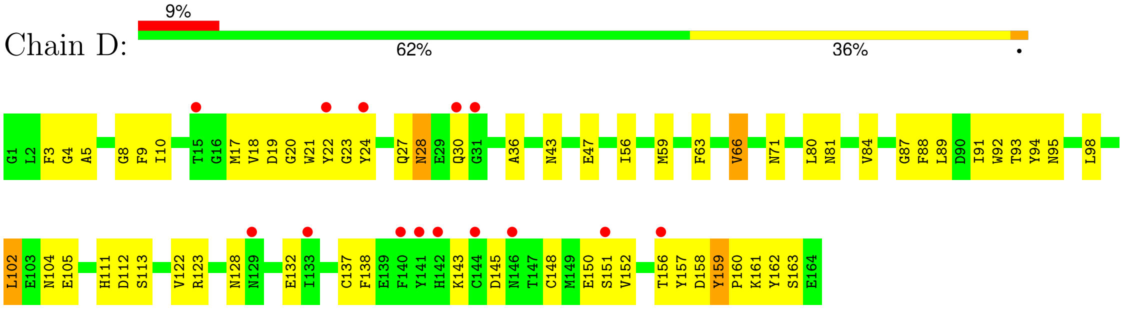
- Molecule 1: Hemagglutinin



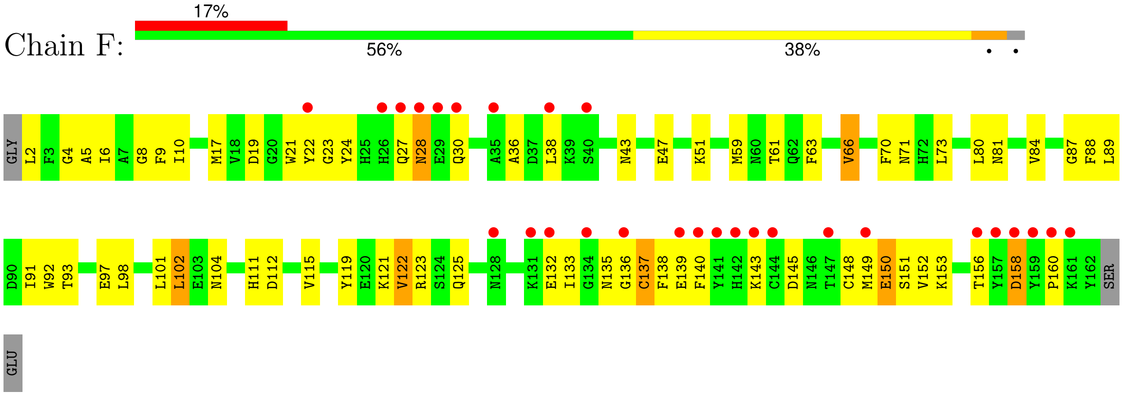
- Molecule 2: Hemagglutinin



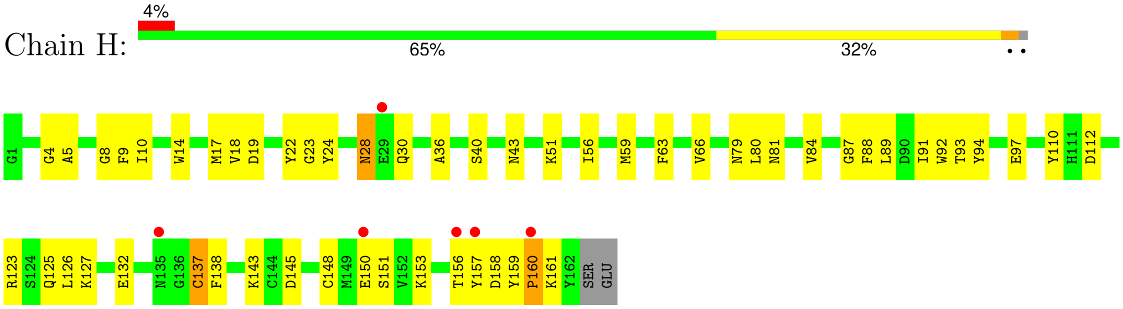
- Molecule 2: Hemagglutinin



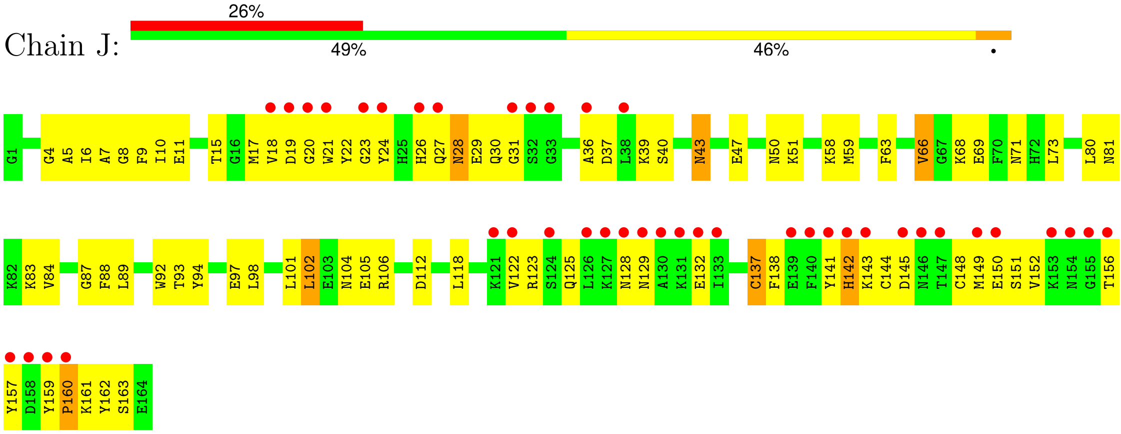
- Molecule 2: Hemagglutinin



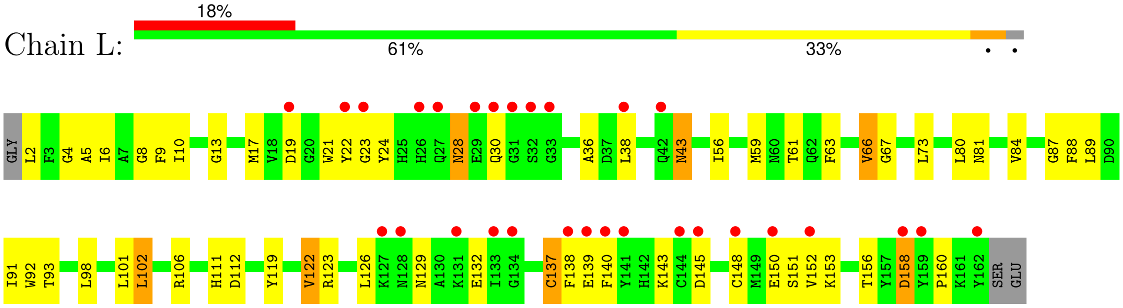
- Molecule 2: Hemagglutinin



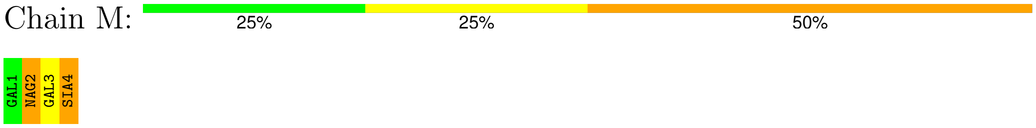
- Molecule 2: Hemagglutinin



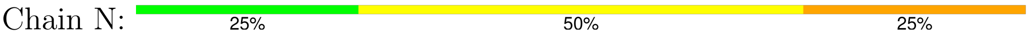
- Molecule 2: Hemagglutinin



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

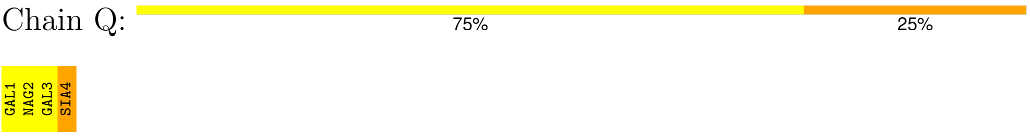


- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose

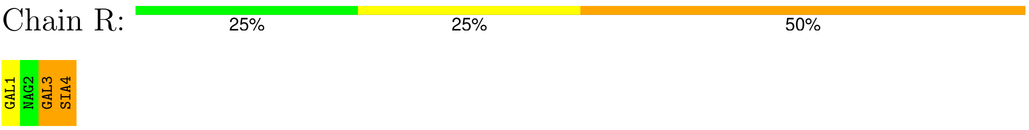


• Molecule 3: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose

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



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



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

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	66.87Å 116.81Å 116.50Å 62.06° 77.76° 81.54°	Depositor
Resolution (Å)	40.22 – 2.91 40.23 – 2.91	Depositor EDS
% Data completeness (in resolution range)	88.2 (40.22-2.91) 94.3 (40.23-2.91)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.59 (at 2.90Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.240 , 0.283 0.238 , 0.277	Depositor DCC
R_{free} test set	3175 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	54.6	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 54.0	EDS
L-test for twinning[2]	$< \|L\| >$ = 0.49, $< L^2 >$ = 0.32	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	23507	wwPDB-VP
Average B, all atoms (Å^2)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.32% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.25	0/2577	0.71	8/3501 (0.2%)
1	C	0.29	1/2573 (0.0%)	0.68	9/3494 (0.3%)
1	E	0.24	0/2573	0.60	8/3494 (0.2%)
1	G	0.26	0/2577	0.70	6/3501 (0.2%)
1	I	0.25	0/2573	0.60	6/3494 (0.2%)
1	K	0.24	0/2573	0.65	8/3494 (0.2%)
2	B	0.24	0/1333	0.54	1/1797 (0.1%)
2	D	0.26	0/1343	0.48	1/1811 (0.1%)
2	F	0.24	0/1330	0.49	0/1794
2	H	0.24	0/1333	0.50	1/1797 (0.1%)
2	J	0.24	0/1343	0.47	1/1811 (0.1%)
2	L	0.24	0/1330	0.49	0/1794
All	All	0.25	1/23458 (0.0%)	0.61	49/31782 (0.2%)

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	E	0	1
1	K	0	1
2	J	0	1
All	All	0	3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	73	CYS	CB-SG	-5.76	1.72	1.81

The worst 5 of 49 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	119	ARG	NE-CZ-NH1	-16.32	112.14	120.30
1	A	119	ARG	NE-CZ-NH1	-16.20	112.20	120.30
1	A	119	ARG	NE-CZ-NH2	16.09	128.34	120.30
1	G	119	ARG	NE-CZ-NH2	15.81	128.21	120.30
1	K	199	GLN	CB-CA-C	-14.58	81.24	110.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	77	SER	Peptide
2	J	141	TYR	Peptide
1	K	77	SER	Peptide

5.2 Too-close contacts ⓘ

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	2513	0	2457	118	0
1	C	2510	0	2452	160	0
1	E	2510	0	2453	130	0
1	G	2513	0	2455	117	0
1	I	2510	0	2452	166	0
1	K	2510	0	2452	150	0
2	B	1305	0	1228	54	0
2	D	1315	0	1227	70	0
2	F	1302	0	1226	84	0
2	H	1305	0	1228	59	0
2	J	1315	0	1227	155	0
2	L	1302	0	1226	81	0
3	M	56	0	47	3	0
3	N	56	0	47	4	0
3	P	56	0	47	3	0
3	Q	56	0	47	5	0
3	R	56	0	47	2	0
4	O	28	0	25	1	0
5	C	28	0	26	5	0
5	G	14	0	13	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	I	14	0	13	0	0
5	K	28	0	26	9	0
6	K	20	0	17	1	0
7	A	22	0	0	32	0
7	B	6	0	0	3	0
7	C	20	0	0	40	0
7	D	5	0	0	14	0
7	E	18	0	0	32	0
7	F	11	0	0	23	0
7	G	16	0	0	15	0
7	H	6	0	0	12	0
7	I	25	0	0	50	0
7	J	22	0	0	54	0
7	K	26	0	0	51	0
7	L	8	0	0	17	0
All	All	23507	0	22438	1137	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 25.

The worst 5 of 1137 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:56:LEU:CD1	1:C:76:LEU:HD12	1.48	1.42
1:C:56:LEU:HD12	1:C:76:LEU:CD1	1.49	1.40
1:C:8:THR:HB	7:D:204:HOH:O	1.18	1.31
1:C:110:GLN:HB2	7:C:714:HOH:O	1.20	1.29
1:G:245:LYS:HA	7:G:516:HOH:O	1.14	1.28

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	320/322 (99%)	296 (92%)	24 (8%)	0	100	100
1	C	319/322 (99%)	297 (93%)	22 (7%)	0	100	100
1	E	319/322 (99%)	297 (93%)	22 (7%)	0	100	100
1	G	320/322 (99%)	298 (93%)	22 (7%)	0	100	100
1	I	319/322 (99%)	298 (93%)	21 (7%)	0	100	100
1	K	319/322 (99%)	297 (93%)	22 (7%)	0	100	100
2	B	160/164 (98%)	149 (93%)	10 (6%)	1 (1%)	25	58
2	D	162/164 (99%)	150 (93%)	11 (7%)	1 (1%)	25	58
2	F	159/164 (97%)	148 (93%)	11 (7%)	0	100	100
2	H	160/164 (98%)	149 (93%)	10 (6%)	1 (1%)	25	58
2	J	162/164 (99%)	149 (92%)	12 (7%)	1 (1%)	25	58
2	L	159/164 (97%)	147 (92%)	12 (8%)	0	100	100
All	All	2878/2916 (99%)	2675 (93%)	199 (7%)	4 (0%)	51	82

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	160	PRO
2	H	160	PRO
2	J	160	PRO
2	D	159	TYR

5.3.2 Protein sidechains ⓘ

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	282/283 (100%)	259 (92%)	23 (8%)	11	32
1	C	282/283 (100%)	260 (92%)	22 (8%)	12	34
1	E	282/283 (100%)	262 (93%)	20 (7%)	14	40
1	G	282/283 (100%)	263 (93%)	19 (7%)	16	43
1	I	282/283 (100%)	261 (93%)	21 (7%)	13	38

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	282/283 (100%)	263 (93%)	19 (7%)	16	43
2	B	139/141 (99%)	132 (95%)	7 (5%)	24	57
2	D	140/141 (99%)	131 (94%)	9 (6%)	17	45
2	F	139/141 (99%)	126 (91%)	13 (9%)	8	26
2	H	139/141 (99%)	132 (95%)	7 (5%)	24	57
2	J	140/141 (99%)	131 (94%)	9 (6%)	17	45
2	L	139/141 (99%)	126 (91%)	13 (9%)	8	26
All	All	2528/2544 (99%)	2346 (93%)	182 (7%)	14	39

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	43	ASN
2	J	43	ASN
2	H	158	ASP
1	I	111	LEU
1	K	9	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 96 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	266	ASN
1	I	229	GLN
2	H	30	GLN
1	I	44	ASN
2	J	30	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates

22 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	GAL	M	1	3	11,11,12	0.63	0	15,15,17	0.58	0
3	NAG	M	2	3	14,14,15	0.65	0	17,19,21	1.03	1 (5%)
3	GAL	M	3	3	11,11,12	0.65	0	15,15,17	0.93	1 (6%)
3	SIA	M	4	3	20,20,21	0.69	0	21,28,31	3.44	3 (14%)
3	GAL	N	1	3	11,11,12	0.64	0	15,15,17	0.49	0
3	NAG	N	2	3	14,14,15	0.58	0	17,19,21	0.81	0
3	GAL	N	3	3	11,11,12	0.62	0	15,15,17	0.78	1 (6%)
3	SIA	N	4	3	20,20,21	0.70	0	21,28,31	3.43	3 (14%)
4	NAG	O	1	1,4	14,14,15	0.57	0	17,19,21	0.67	0
4	NAG	O	2	4	14,14,15	0.47	0	17,19,21	0.71	0
3	GAL	P	1	3	11,11,12	0.61	0	15,15,17	0.55	0
3	NAG	P	2	3	14,14,15	0.62	0	17,19,21	0.73	0
3	GAL	P	3	3	11,11,12	0.71	0	15,15,17	0.83	0
3	SIA	P	4	3	20,20,21	0.59	0	21,28,31	3.65	3 (14%)
3	GAL	Q	1	3	11,11,12	0.71	0	15,15,17	1.13	1 (6%)
3	NAG	Q	2	3	14,14,15	0.59	0	17,19,21	0.96	0
3	GAL	Q	3	3	11,11,12	0.72	0	15,15,17	0.92	1 (6%)
3	SIA	Q	4	3	20,20,21	0.67	0	21,28,31	3.37	3 (14%)
3	GAL	R	1	3	11,11,12	0.67	0	15,15,17	1.34	3 (20%)
3	NAG	R	2	3	14,14,15	0.57	0	17,19,21	0.59	0
3	GAL	R	3	3	11,11,12	0.67	0	15,15,17	1.22	2 (13%)
3	SIA	R	4	3	20,20,21	0.70	0	21,28,31	3.22	3 (14%)

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	GAL	M	1	3	-	0/2/19/22	0/1/1/1
3	NAG	M	2	3	-	0/6/23/26	0/1/1/1
3	GAL	M	3	3	-	0/2/19/22	0/1/1/1
3	SIA	M	4	3	-	2/18/34/38	0/1/1/1
3	GAL	N	1	3	-	0/2/19/22	0/1/1/1
3	NAG	N	2	3	-	0/6/23/26	0/1/1/1
3	GAL	N	3	3	-	1/2/19/22	0/1/1/1
3	SIA	N	4	3	-	3/18/34/38	0/1/1/1
4	NAG	O	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	O	2	4	-	0/6/23/26	0/1/1/1
3	GAL	P	1	3	-	0/2/19/22	0/1/1/1
3	NAG	P	2	3	-	0/6/23/26	0/1/1/1
3	GAL	P	3	3	-	0/2/19/22	0/1/1/1
3	SIA	P	4	3	-	3/18/34/38	0/1/1/1
3	GAL	Q	1	3	-	2/2/19/22	0/1/1/1
3	NAG	Q	2	3	-	0/6/23/26	0/1/1/1
3	GAL	Q	3	3	-	0/2/19/22	0/1/1/1
3	SIA	Q	4	3	-	4/18/34/38	0/1/1/1
3	GAL	R	1	3	-	0/2/19/22	0/1/1/1
3	NAG	R	2	3	-	0/6/23/26	0/1/1/1
3	GAL	R	3	3	-	0/2/19/22	0/1/1/1
3	SIA	R	4	3	-	4/18/34/38	0/1/1/1

There are no bond length outliers.

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	P	4	SIA	O6-C2-C3	-14.49	91.08	110.56
3	N	4	SIA	O6-C2-C3	-13.24	92.76	110.56
3	M	4	SIA	O6-C2-C3	-13.21	92.80	110.56
3	Q	4	SIA	O6-C2-C3	-12.97	93.11	110.56
3	R	4	SIA	O6-C2-C3	-11.68	94.85	110.56

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	Q	1	GAL	O5-C5-C6-O6
3	Q	1	GAL	C4-C5-C6-O6
3	Q	4	SIA	O8-C8-C9-O9
3	M	4	SIA	C11-C10-N5-C5

Continued on next page...

Continued from previous page...

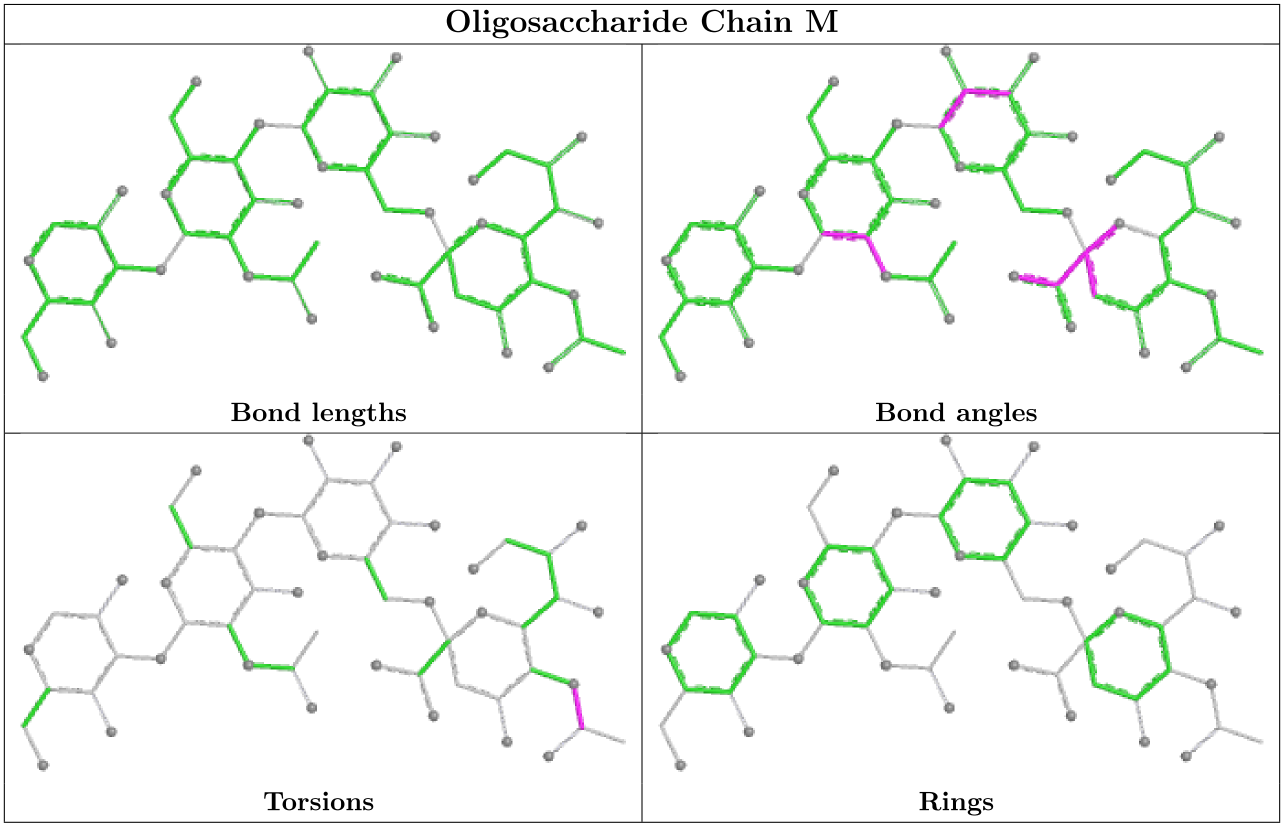
Mol	Chain	Res	Type	Atoms
3	M	4	SIA	O10-C10-N5-C5

There are no ring outliers.

11 monomers are involved in 18 short contacts:

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

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


Oligosaccharide Chain M

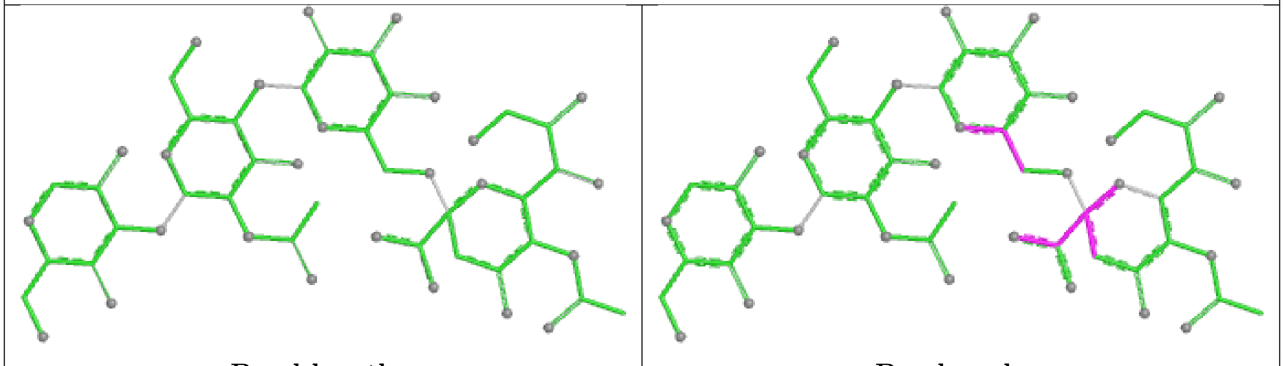
Bond lengths

Bond angles

Torsions

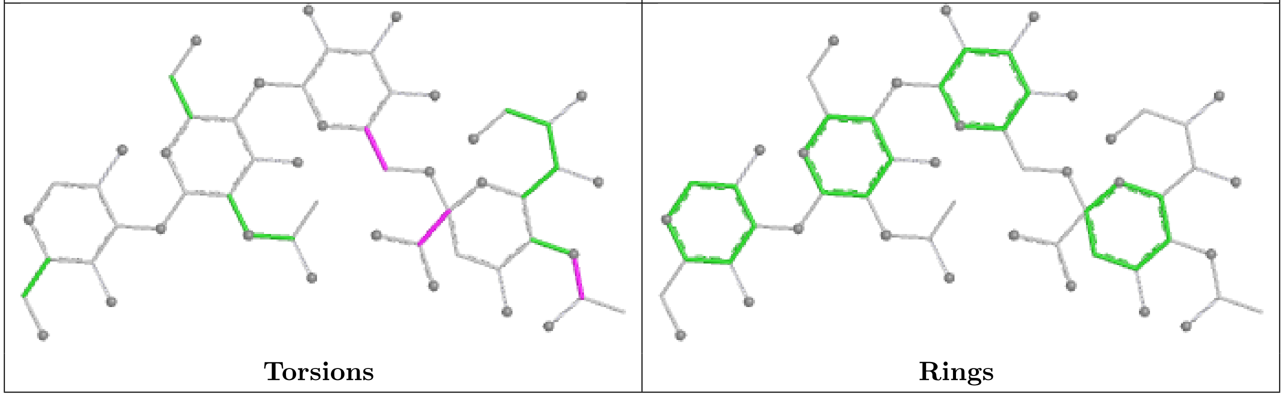
Rings

Oligosaccharide Chain N



Bond lengths

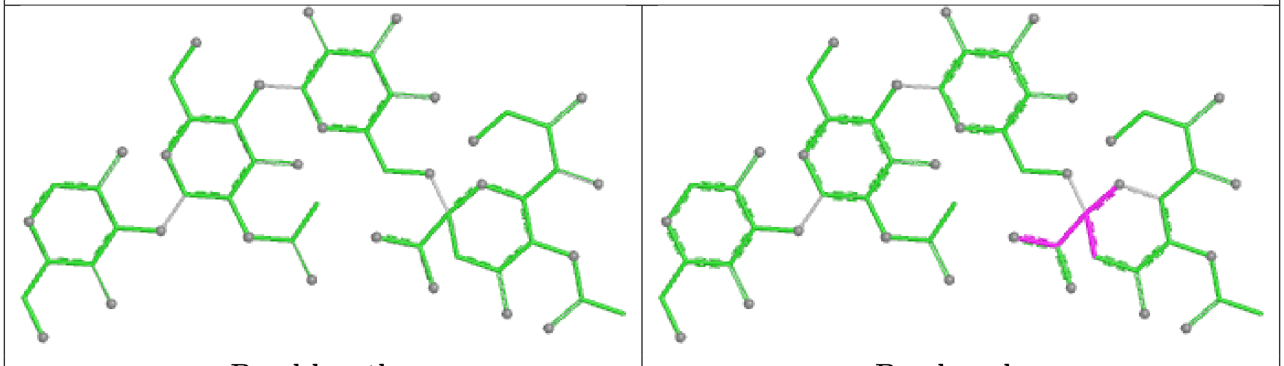
Bond angles



Torsions

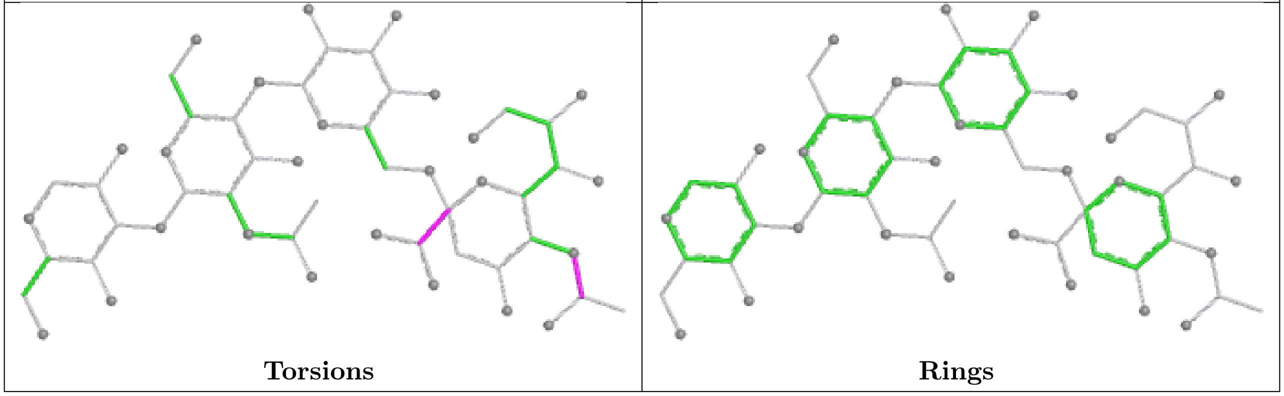
Rings

Oligosaccharide Chain P



Bond lengths

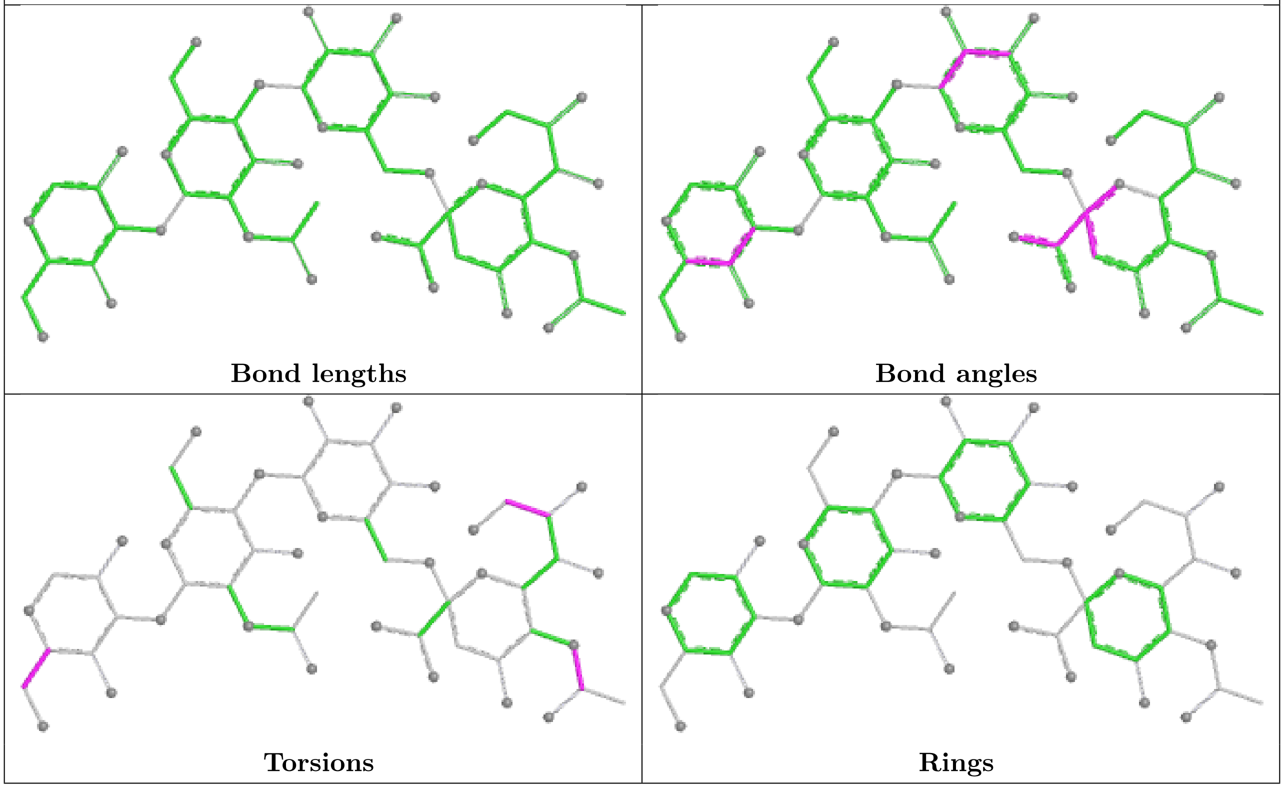
Bond angles



Torsions

Rings

Oligosaccharide Chain Q



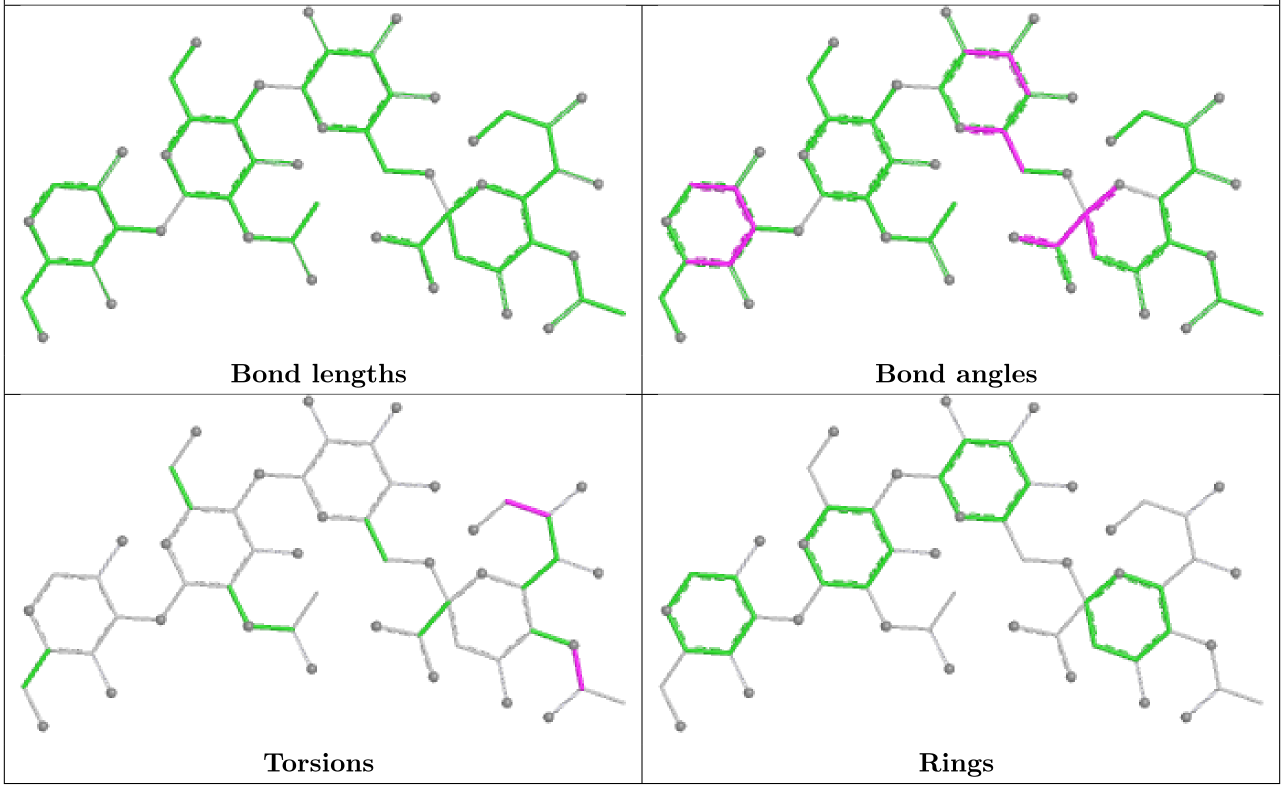
Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain R



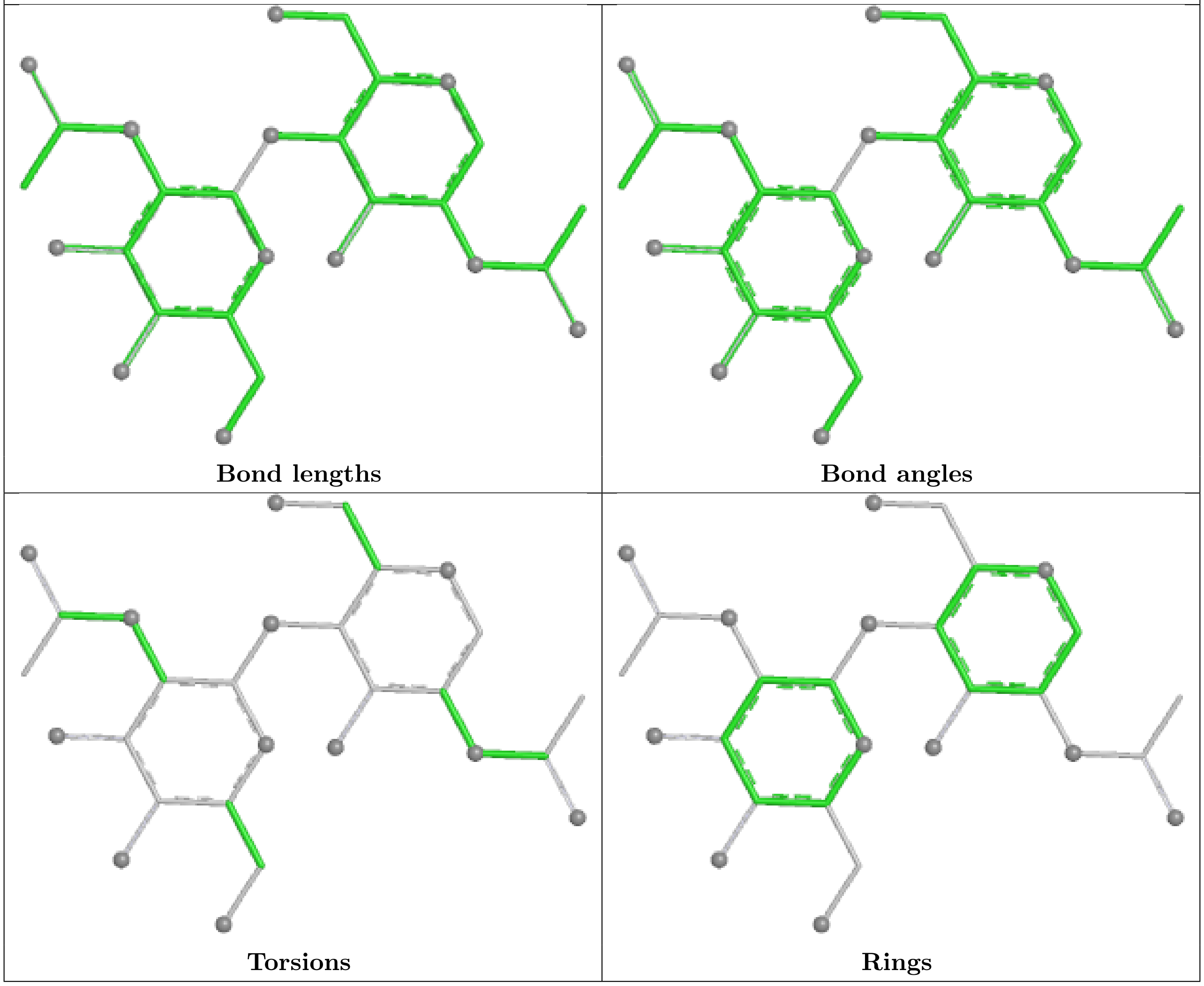
Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain O



Bond lengths

Bond angles

Torsions

Rings

5.6 Ligand geometry (i)

7 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	SIA	K	603	-	20,20,21	0.63	0	21,28,31	1.08	3 (14%)
5	NAG	K	602	1	14,14,15	0.67	0	17,19,21	0.68	0
5	NAG	C	601	1	14,14,15	0.80	0	17,19,21	1.25	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	NAG	K	601	1	14,14,15	0.46	0	17,19,21	1.82	4 (23%)
5	NAG	G	401	1	14,14,15	0.47	0	17,19,21	1.03	1 (5%)
5	NAG	I	601	1	14,14,15	0.59	0	17,19,21	0.90	1 (5%)
5	NAG	C	602	1	14,14,15	0.53	0	17,19,21	0.80	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	SIA	K	603	-	-	2/18/34/38	0/1/1/1
5	NAG	K	602	1	-	2/6/23/26	0/1/1/1
5	NAG	C	601	1	-	3/6/23/26	0/1/1/1
5	NAG	K	601	1	-	0/6/23/26	0/1/1/1
5	NAG	G	401	1	-	2/6/23/26	0/1/1/1
5	NAG	I	601	1	-	0/6/23/26	0/1/1/1
5	NAG	C	602	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	K	601	NAG	C1-O5-C5	5.31	119.30	112.19
5	K	601	NAG	C6-C5-C4	-3.03	105.58	113.02
6	K	603	SIA	O6-C2-C3	2.89	114.45	110.56
5	G	401	NAG	C1-O5-C5	2.87	116.04	112.19
5	C	601	NAG	C4-C3-C2	2.61	114.84	111.02

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	601	NAG	C4-C5-C6-O6
5	C	601	NAG	O5-C5-C6-O6
5	K	602	NAG	O5-C5-C6-O6
5	K	602	NAG	C4-C5-C6-O6
6	K	603	SIA	C11-C10-N5-C5

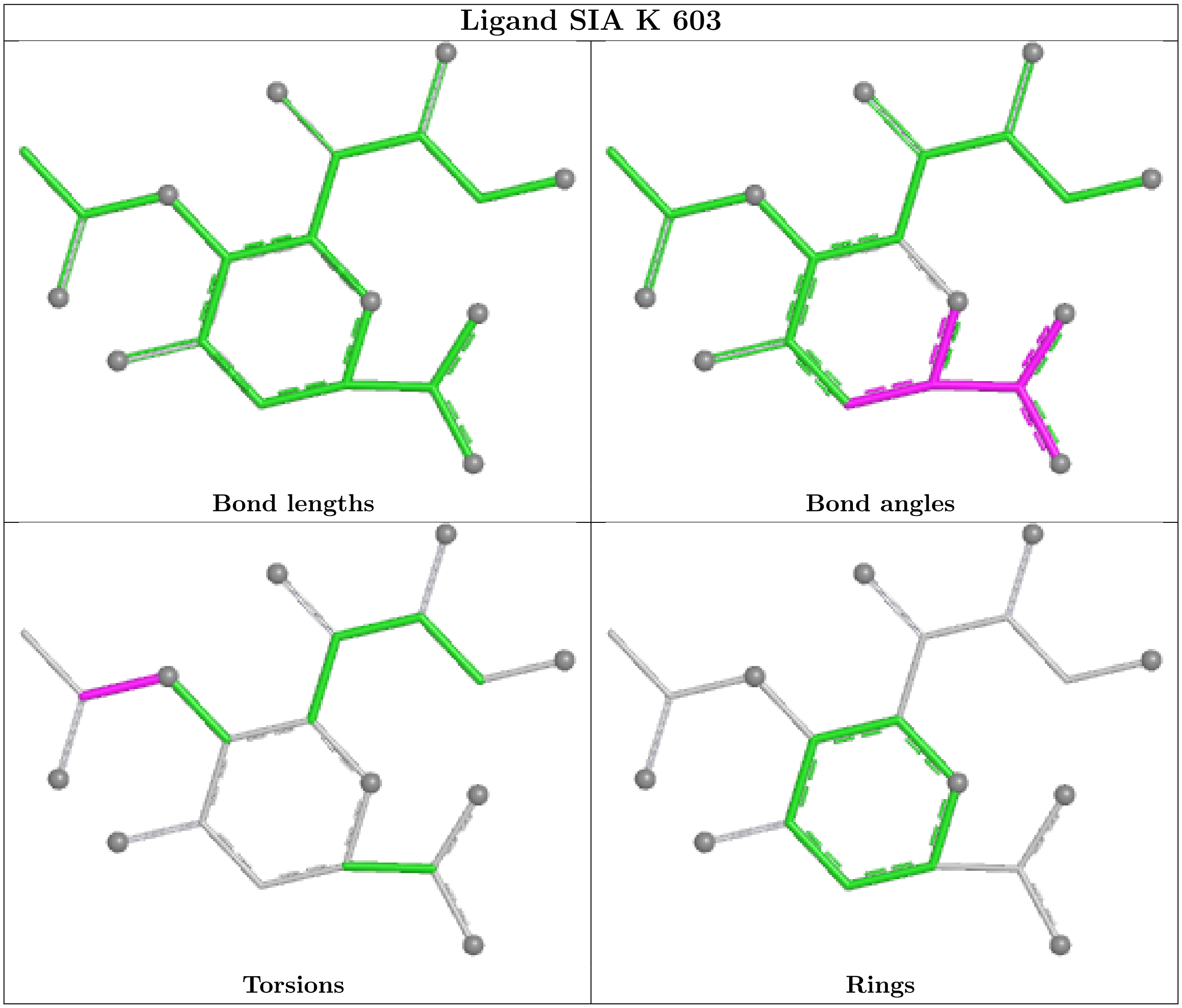
There are no ring outliers.

4 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	K	603	SIA	1	0
5	K	602	NAG	4	0
5	K	601	NAG	5	0
5	C	602	NAG	5	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 then 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.

Ligand SIA K 603



Bond lengths

Bond angles

Torsions

Rings

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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^2)	Q<0.9
1	A	322/322 (100%)	0.36	21 (6%)	18	14	30, 55, 103, 206	0
1	C	321/322 (99%)	0.18	17 (5%)	26	22	29, 55, 102, 249	0
1	E	321/322 (99%)	0.14	15 (4%)	31	28	24, 56, 102, 220	0
1	G	322/322 (100%)	0.22	20 (6%)	20	16	29, 58, 98, 209	0
1	I	321/322 (99%)	0.42	21 (6%)	18	14	32, 64, 124, 236	0
1	K	321/322 (99%)	0.21	18 (5%)	24	20	26, 53, 96, 165	0
2	B	162/164 (98%)	0.55	16 (9%)	7	5	32, 72, 129, 175	0
2	D	164/164 (100%)	0.51	14 (8%)	10	8	36, 74, 130, 191	0
2	F	161/164 (98%)	0.80	28 (17%)	1	1	22, 74, 155, 210	0
2	H	162/164 (98%)	0.24	6 (3%)	41	37	28, 59, 101, 138	0
2	J	164/164 (100%)	1.38	42 (25%)	0	0	29, 81, 203, 324	0
2	L	161/164 (98%)	0.74	29 (18%)	1	1	30, 63, 163, 201	0
All	All	2902/2916 (99%)	0.41	247 (8%)	10	8	22, 60, 129, 324	0

The worst 5 of 247 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	J	158	ASP	11.0
2	F	159	TYR	11.0
1	I	9	LEU	8.6
2	J	129	ASN	8.2
2	J	20	GLY	8.2

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates

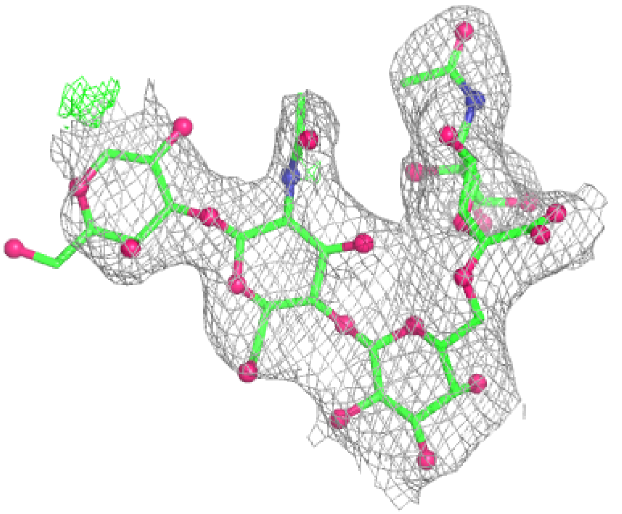
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

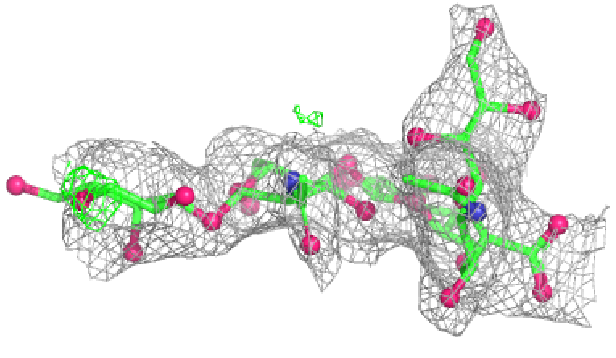
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GAL	Q	1	11/12	0.60	0.34	99,134,164,181	0
3	GAL	P	1	11/12	0.70	0.42	122,137,149,151	0
3	GAL	R	1	11/12	0.79	0.24	62,112,129,131	0
3	GAL	N	1	11/12	0.80	0.29	57,74,100,104	0
3	NAG	R	2	14/15	0.81	0.20	62,83,91,98	0
3	NAG	P	2	14/15	0.84	0.30	77,97,113,115	0
3	GAL	M	1	11/12	0.85	0.27	84,94,133,136	0
3	GAL	Q	3	11/12	0.86	0.19	45,65,85,103	0
3	NAG	Q	2	14/15	0.87	0.26	72,89,109,111	0
3	NAG	M	2	14/15	0.88	0.19	51,70,78,85	0
3	GAL	R	3	11/12	0.89	0.20	78,96,102,114	0
4	NAG	O	1	14/15	0.89	0.18	52,64,82,92	0
3	GAL	N	3	11/12	0.91	0.16	53,69,77,79	0
3	SIA	P	4	20/21	0.91	0.20	47,59,73,76	0
3	GAL	M	3	11/12	0.91	0.13	42,52,72,77	0
4	NAG	O	2	14/15	0.91	0.33	63,93,116,119	0
3	SIA	Q	4	20/21	0.92	0.18	44,68,95,96	0
3	SIA	N	4	20/21	0.93	0.14	41,62,76,87	0
3	GAL	P	3	11/12	0.93	0.13	53,71,89,92	0
3	NAG	N	2	14/15	0.93	0.30	60,71,84,92	0
3	SIA	R	4	20/21	0.94	0.17	38,55,67,73	0
3	SIA	M	4	20/21	0.95	0.14	34,46,60,66	0

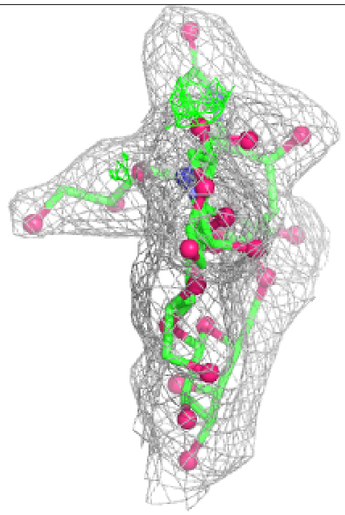
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

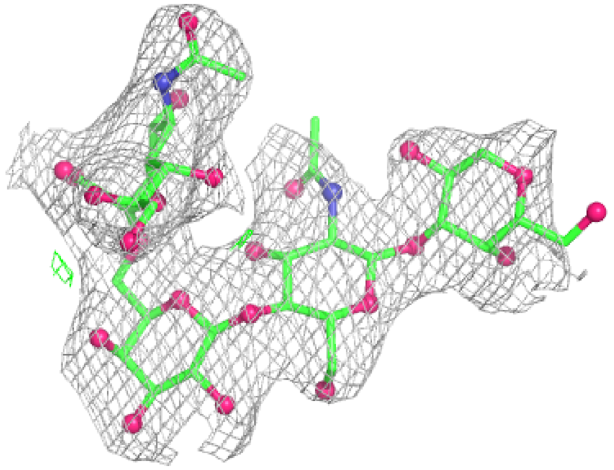


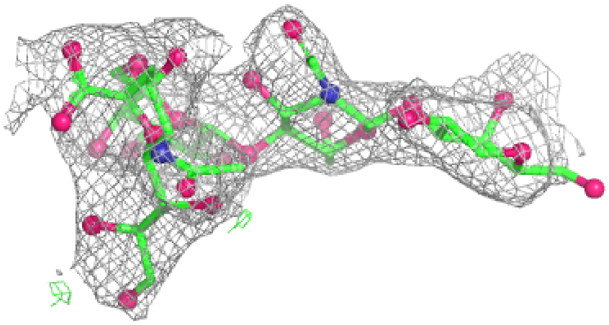


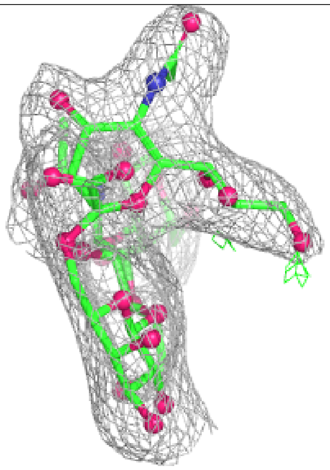


Electron density around Chain N:

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

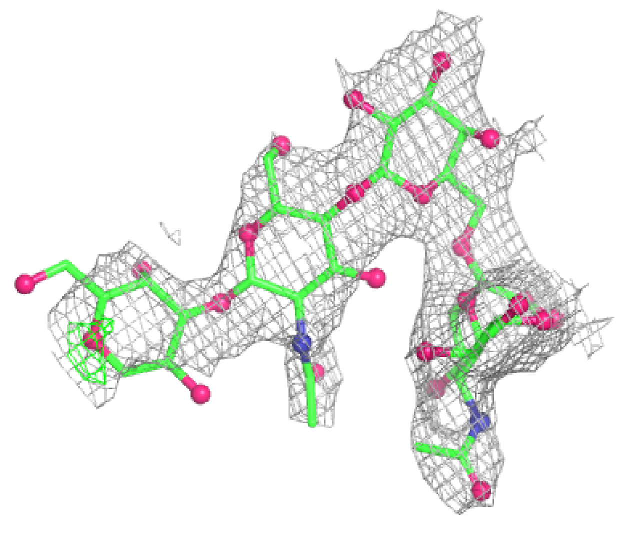


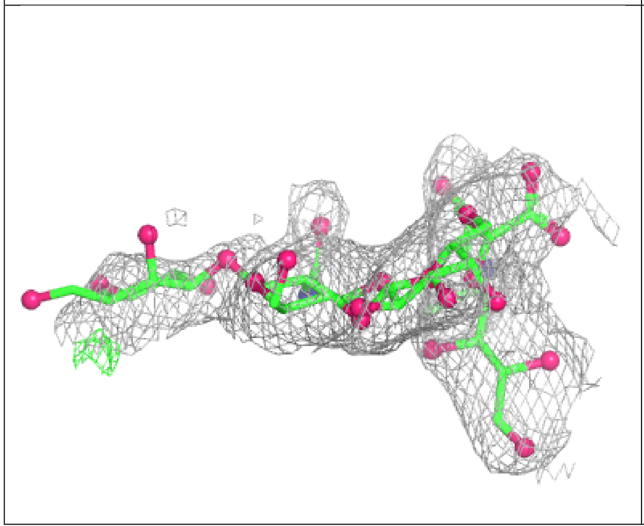


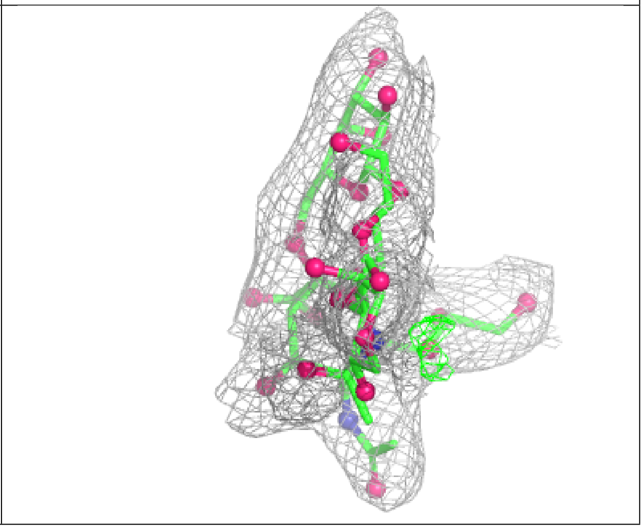


Electron density around Chain P:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

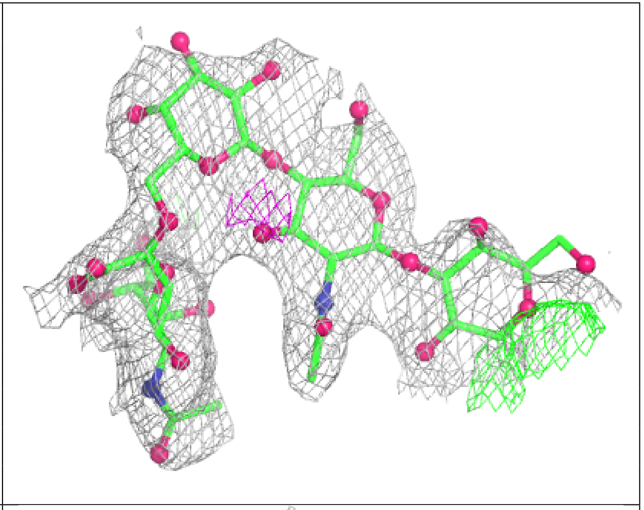


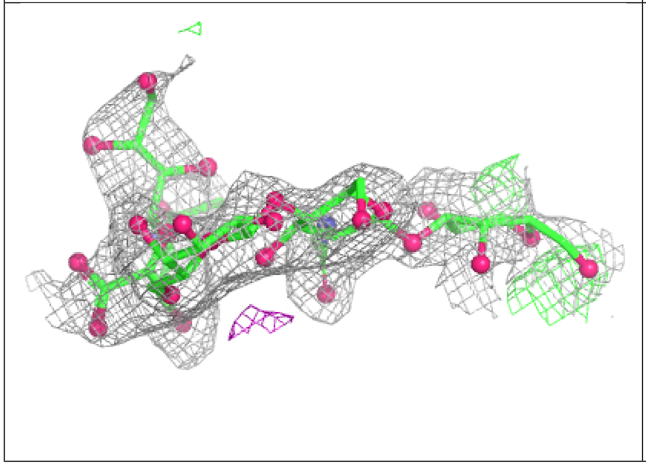


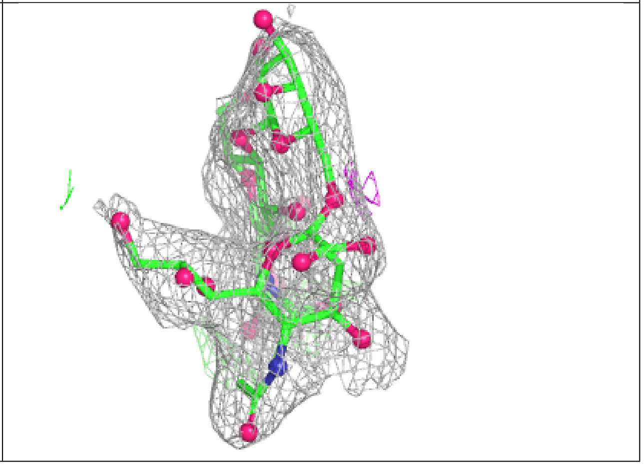


Electron density around Chain Q:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

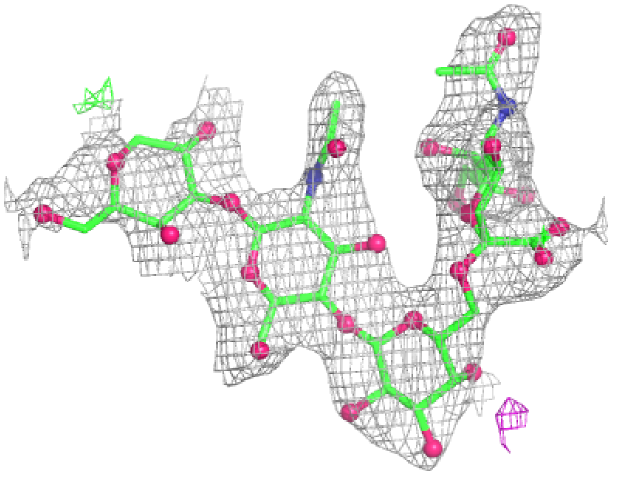


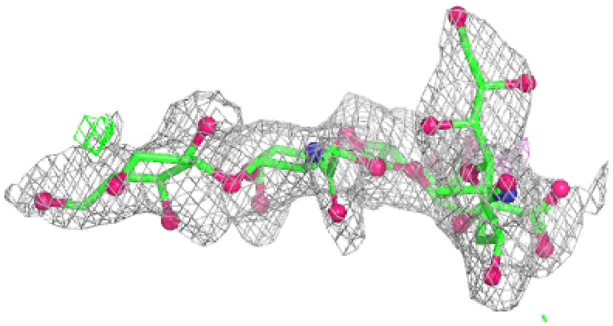


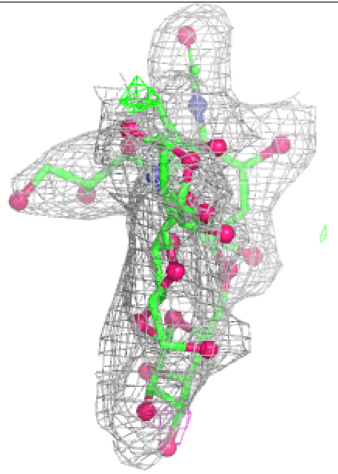


Electron density around Chain R:

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

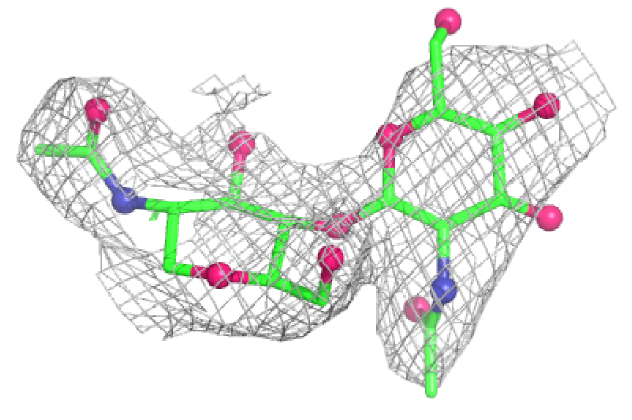


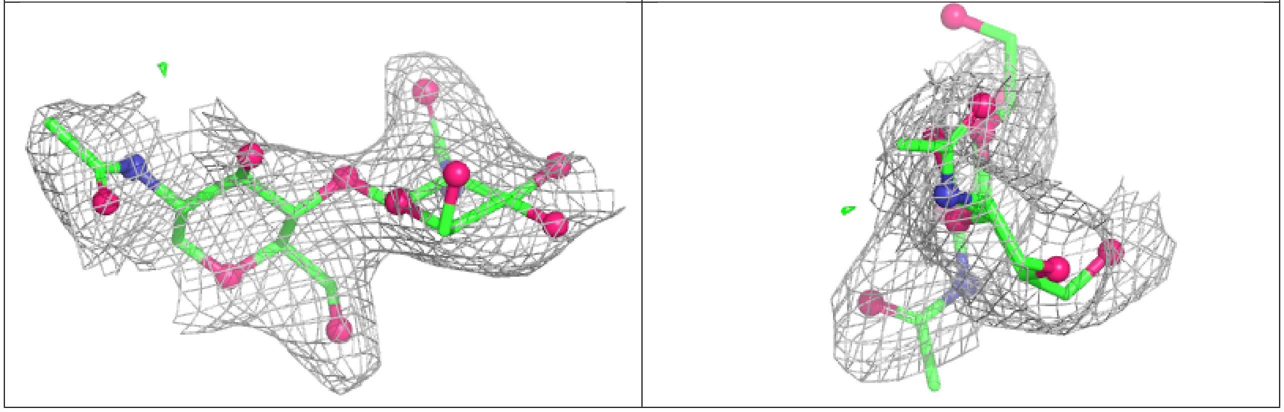




Electron density around Chain O:

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



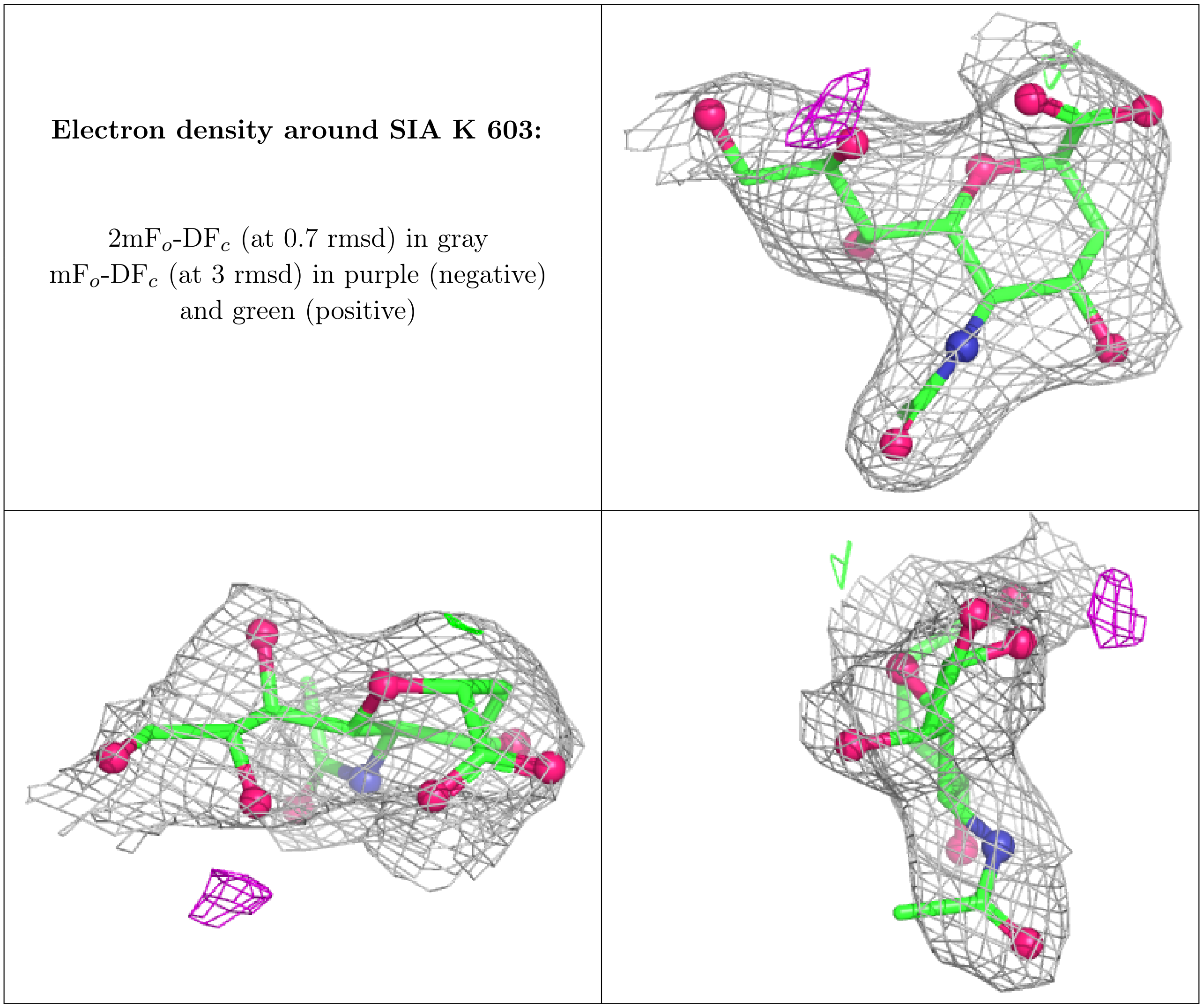


6.4 Ligands ⓘ

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, 95th 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($Å^2$)	Q<0.9
5	NAG	I	601	14/15	0.69	0.32	101,130,147,150	0
5	NAG	K	602	14/15	0.72	0.32	76,105,115,115	0
5	NAG	K	601	14/15	0.77	0.31	119,128,136,137	0
5	NAG	C	601	14/15	0.78	0.21	66,93,106,114	0
5	NAG	G	401	14/15	0.82	0.38	78,108,129,132	0
5	NAG	C	602	14/15	0.92	0.18	25,57,74,79	0
6	SIA	K	603	20/21	0.95	0.16	42,52,69,76	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.



Electron density around SIA K 603:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5 Other polymers ⓘ

There are no such residues in this entry.