wwPDB X-ray Structure Validation Summary Report (i)

Oct 13, 2020 – 10:15 AM BST

PDB ID	:	6TVT
Title	:	Crystal structure of the haemagglutinin mutant (Gln226Leu, Del228) from an H10N7 seal influenza virus isolated in Germany in complex with human receptor analogue 6'-SLN
Authors	:	Zhang, J.; Xiong, X.; Purkiss, A.; Walker, P.; Gamblin, S.; Skehel, J.J.
Deposited on	:	2020-01-10
Resolution	:	2.20 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

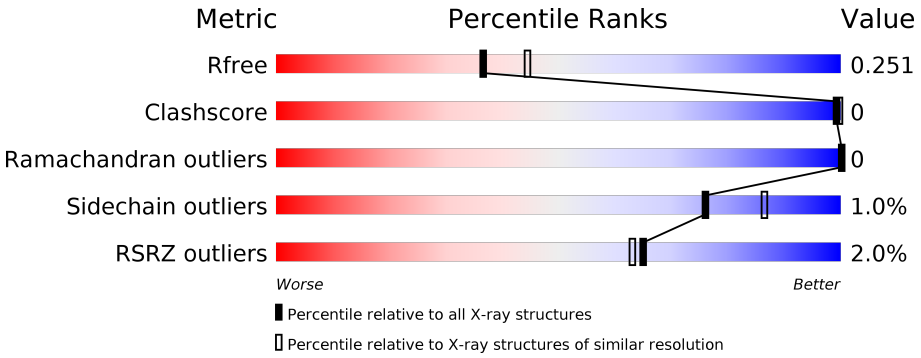
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.6
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.6

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 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 >=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	324	3% 94%
1	C	324	% 94%
1	E	324	5% 95%
2	B	177	% 96%
2	D	177	95%
2	F	177	96%

Continued on next page...

Continued from previous page...

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

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 11831 atoms, of which 0 are hydrogens and 0 are deuteriums.

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

- Molecule 1 is a protein called Hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	312	Total	C	N	O	S	0	0	0
			2390	1485	430	459	16			
1	C	311	Total	C	N	O	S	0	0	0
			2386	1483	429	458	16			
1	E	313	Total	C	N	O	S	0	0	0
			2394	1487	431	460	16			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	ASP	-	expression tag	UNP A0A0A7HR51
A	0	PRO	-	expression tag	UNP A0A0A7HR51
A	?	-	GLY	deletion	UNP A0A0A7HR51
C	-1	ASP	-	expression tag	UNP A0A0A7HR51
C	0	PRO	-	expression tag	UNP A0A0A7HR51
C	?	-	GLY	deletion	UNP A0A0A7HR51
E	-1	ASP	-	expression tag	UNP A0A0A7HR51
E	0	PRO	-	expression tag	UNP A0A0A7HR51
E	?	-	GLY	deletion	UNP A0A0A7HR51

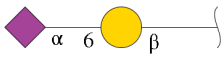
- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			
2	D	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			
2	F	172	Total	C	N	O	S	0	0	0
			1386	857	241	280	8			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	177	LYS	-	expression tag	UNP A0A0A7HR51
D	177	LYS	-	expression tag	UNP A0A0A7HR51
F	177	LYS	-	expression tag	UNP A0A0A7HR51

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

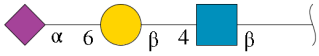


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	G	2	Total	C	N	O	0	0	0
			32	17	1	14			
3	K	2	Total	C	N	O	0	0	0
			32	17	1	14			

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

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

- Molecule 5 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galacto pyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

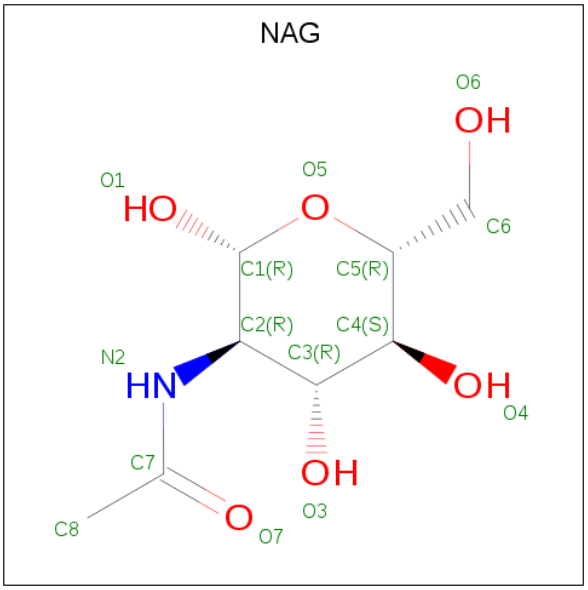


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	I	3	Total	C	N	O	0	0	0
			46	25	2	19			

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total Ca 1 1	0	0
6	A	1	Total Ca 1 1	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$) (labeled as "Ligand of Interest" by author).



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

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	33	Total O 33 33	0	0
8	B	64	Total O 64 64	0	0
8	C	68	Total O 68 68	0	0
8	D	57	Total O 57 57	0	0
8	E	35	Total O 35 35	0	0

Continued on next page...

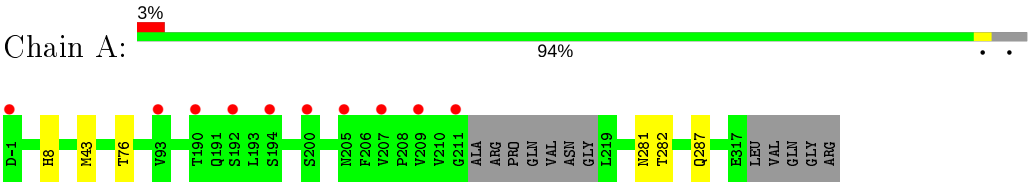
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	F	50	Total O 50 50	0	0

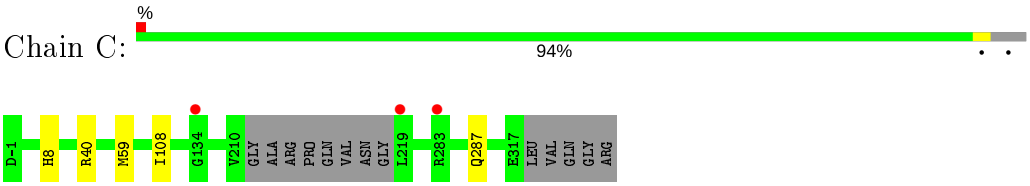
3 Residue-property plots (i)

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.

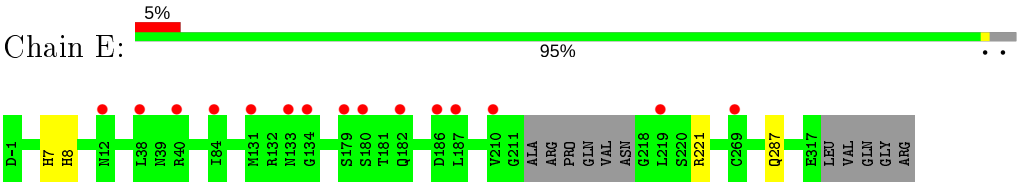
- Molecule 1: Hemagglutinin HA1



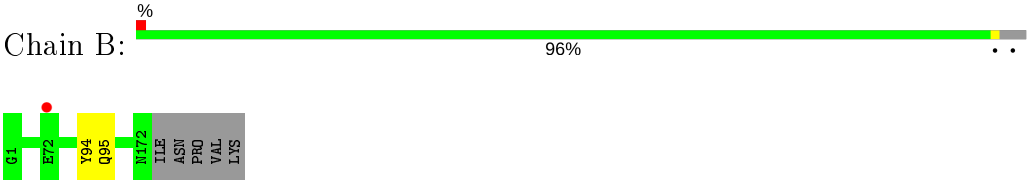
- Molecule 1: Hemagglutinin HA1



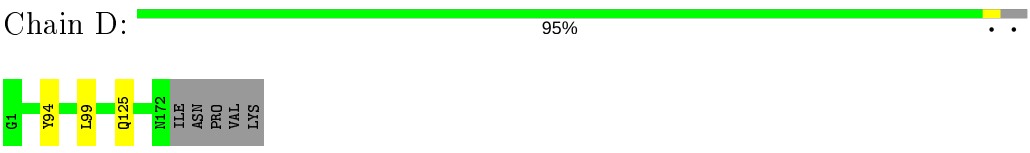
- Molecule 1: Hemagglutinin HA1



- Molecule 2: Hemagglutinin HA2

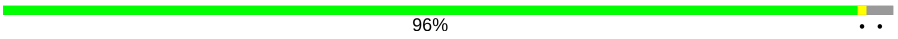


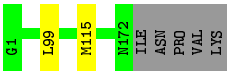
- Molecule 2: Hemagglutinin HA2



- Molecule 2: Hemagglutinin HA2

Chain F:





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

Chain G:

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

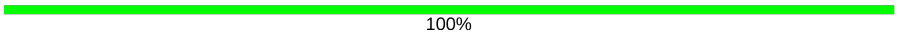
Chain K:

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

Chain H:

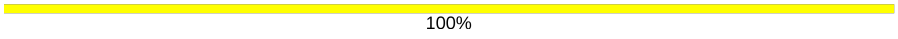
• Molecule 4: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain J:



• Molecule 5: N-acetyl-alpha-neuraminic acid-(2-6)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain I:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.01Å 211.83Å 82.22Å 90.00° 104.70° 90.00°	Depositor
Resolution (Å)	79.53 – 2.20 67.72 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.5 (79.53-2.20) 99.5 (67.72-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.80 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.235 , 0.256 0.232 , 0.251	Depositor DCC
R_{free} test set	5861 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	41.7	Xtriage
Anisotropy	0.661	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 28.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	11831	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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

5.1 Standard geometry

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

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	0.34	0/2438	0.54	0/3301
1	C	0.33	0/2434	0.53	0/3296
1	E	0.34	0/2442	0.53	0/3306
2	B	0.35	0/1411	0.51	0/1903
2	D	0.35	0/1411	0.50	0/1903
2	F	0.35	0/1411	0.50	0/1903
All	All	0.34	0/11547	0.52	0/15612

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	2390	0	2353	1	0
1	C	2386	0	2350	1	0
1	E	2394	0	2359	1	0
2	B	1386	0	1293	1	0
2	D	1386	0	1293	2	0
2	F	1386	0	1291	2	0
3	G	32	0	28	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	K	32	0	28	0	0
4	H	28	0	25	0	0
4	J	28	0	25	0	0
5	I	46	0	40	0	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	E	14	0	13	0	0
7	F	14	0	13	0	0
8	A	33	0	0	0	0
8	B	64	0	0	0	0
8	C	68	0	0	0	0
8	D	57	0	0	0	0
8	E	35	0	0	0	0
8	F	50	0	0	0	0
All	All	11831	0	11111	5	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 0.

All (5) 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)
2:B:94:TYR:OH	2:D:99:LEU:HD22	2.09	0.52
2:D:94:TYR:OH	2:F:99:LEU:HD22	2.18	0.43
1:C:59:MET:HE2	1:C:108:ILE:HD12	2.01	0.42
1:A:43:MET:HE1	1:A:76:THR:HG21	2.01	0.41
1:E:7:HIS:HB3	2:F:115:MET:CE	2.51	0.41

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	308/324 (95%)	297 (96%)	11 (4%)	0	100	100
1	C	307/324 (95%)	303 (99%)	4 (1%)	0	100	100
1	E	309/324 (95%)	300 (97%)	9 (3%)	0	100	100
2	B	170/177 (96%)	166 (98%)	4 (2%)	0	100	100
2	D	170/177 (96%)	166 (98%)	4 (2%)	0	100	100
2	F	170/177 (96%)	165 (97%)	5 (3%)	0	100	100
All	All	1434/1503 (95%)	1397 (97%)	37 (3%)	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	266/275 (97%)	262 (98%)	4 (2%)	65	78
1	C	266/275 (97%)	263 (99%)	3 (1%)	73	85
1	E	266/275 (97%)	263 (99%)	3 (1%)	73	85
2	B	146/151 (97%)	145 (99%)	1 (1%)	84	91
2	D	146/151 (97%)	145 (99%)	1 (1%)	84	91
2	F	146/151 (97%)	146 (100%)	0	100	100
All	All	1236/1278 (97%)	1224 (99%)	12 (1%)	76	86

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	8	HIS
1	C	40	ARG
1	E	8	HIS
2	B	95	GLN
2	D	125	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	105	GLN
2	F	95	GLN
1	E	112	ASN
1	C	267	ASN
1	E	287	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

11 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	G	1	3	12,12,12	0.46	0	17,17,17	0.26	0
3	SIA	G	2	3	17,20,21	0.38	0	21,28,31	1.10	2 (9%)
4	NAG	H	1	2,4,6	14,14,15	0.31	0	17,19,21	0.78	1 (5%)
4	NAG	H	2	4	14,14,15	0.29	0	17,19,21	0.82	0
5	NAG	I	1	5	15,15,15	0.41	0	21,21,21	1.15	2 (9%)
5	GAL	I	2	5	11,11,12	0.32	0	15,15,17	0.62	1 (6%)
5	SIA	I	3	5	17,20,21	0.39	0	21,28,31	1.10	2 (9%)
4	NAG	J	1	2,4	14,14,15	0.37	0	17,19,21	0.83	0
4	NAG	J	2	4	14,14,15	0.29	0	17,19,21	0.74	0
3	GAL	K	1	3	12,12,12	0.45	0	17,17,17	0.32	0
3	SIA	K	2	3	17,20,21	0.36	0	21,28,31	1.11	1 (4%)

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	G	1	3	-	1/2/22/22	0/1/1/1
3	SIA	G	2	3	-	0/14/34/38	0/1/1/1
4	NAG	H	1	2,4,6	-	0/6/23/26	0/1/1/1
4	NAG	H	2	4	-	2/6/23/26	0/1/1/1
5	NAG	I	1	5	-	2/6/26/26	0/1/1/1
5	GAL	I	2	5	-	1/2/19/22	0/1/1/1
5	SIA	I	3	5	-	0/14/34/38	0/1/1/1
4	NAG	J	1	2,4	-	0/6/23/26	0/1/1/1
4	NAG	J	2	4	-	2/6/23/26	0/1/1/1
3	GAL	K	1	3	-	1/2/22/22	0/1/1/1
3	SIA	K	2	3	-	0/14/34/38	0/1/1/1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	I	1	NAG	C1-C2-N2	-3.51	106.66	110.73
5	I	3	SIA	C6-O6-C2	3.32	118.45	111.34
3	G	2	SIA	C6-O6-C2	3.20	118.19	111.34
3	K	2	SIA	C6-O6-C2	3.15	118.08	111.34
5	I	3	SIA	C4-C5-N5	-2.43	105.57	110.38

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

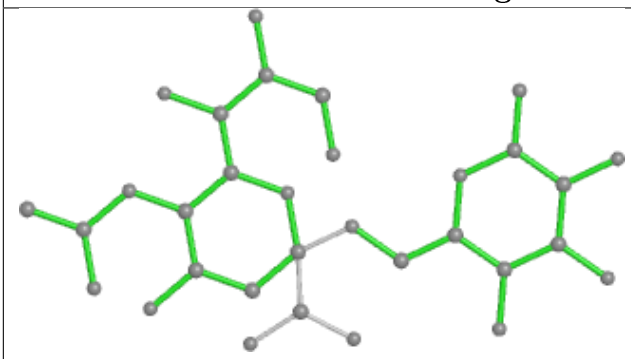
Mol	Chain	Res	Type	Atoms
4	J	2	NAG	C4-C5-C6-O6
4	H	2	NAG	C4-C5-C6-O6
4	J	2	NAG	O5-C5-C6-O6
5	I	1	NAG	C8-C7-N2-C2
5	I	1	NAG	O7-C7-N2-C2

There are no ring outliers.

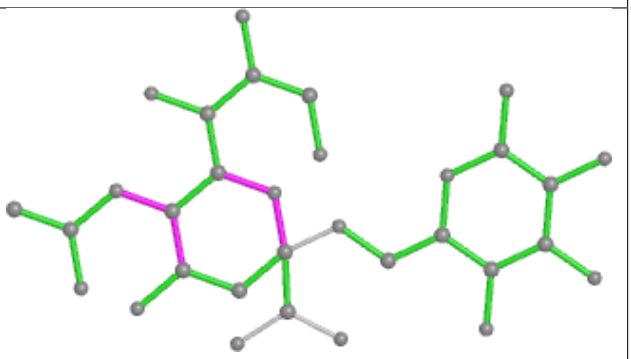
No monomer is involved in short contacts.

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

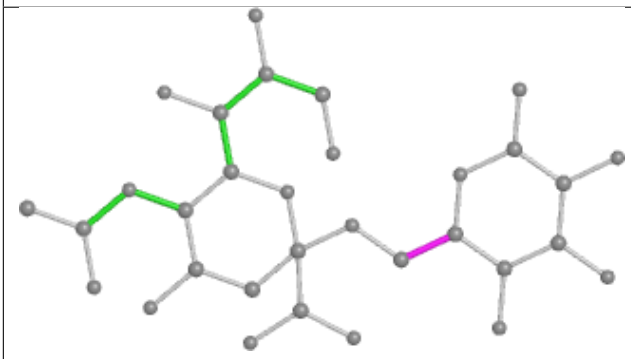
Oligosaccharide Chain G



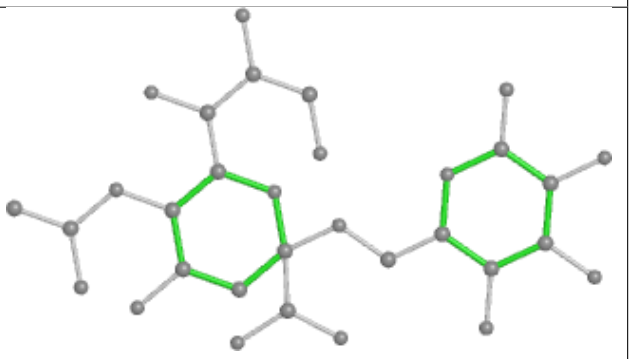
Bond lengths



Bond angles

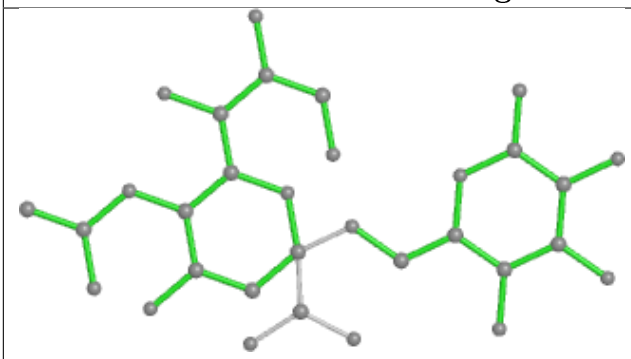


Torsions

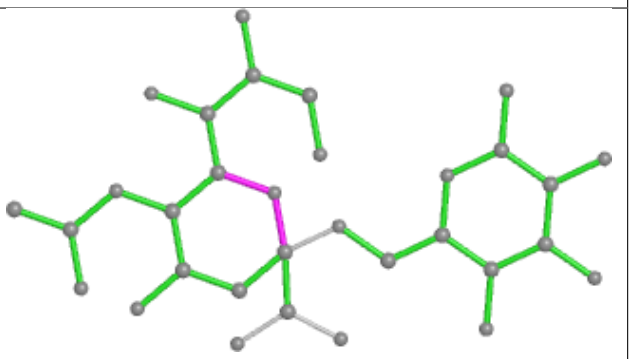


Rings

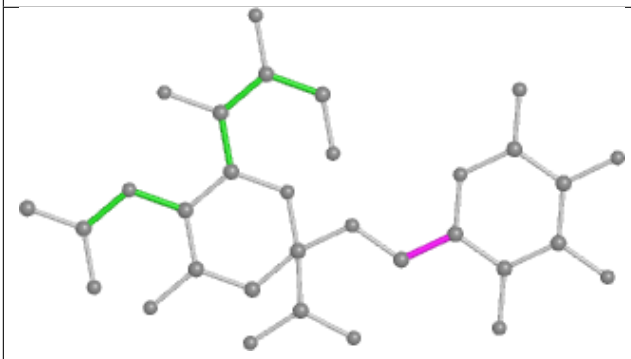
Oligosaccharide Chain K



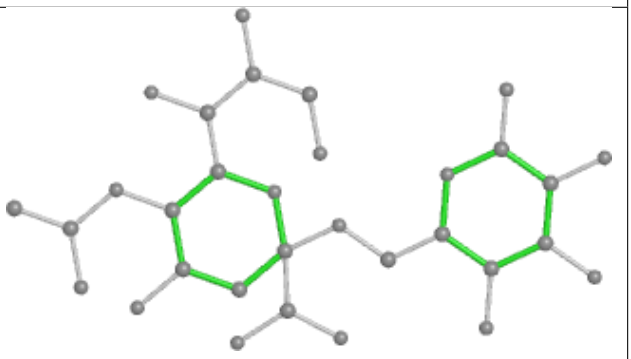
Bond lengths



Bond angles

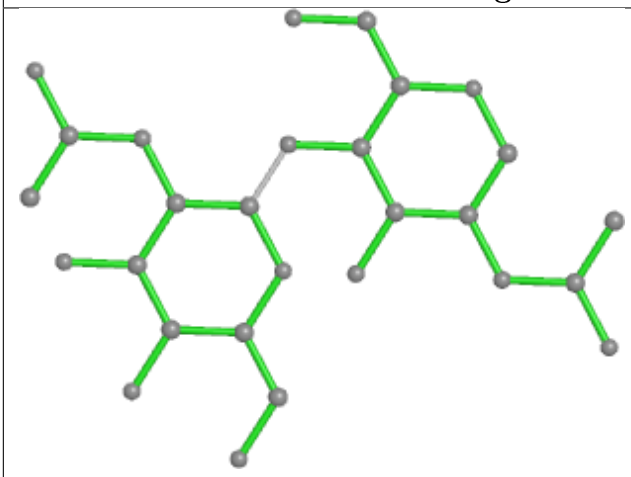


Torsions

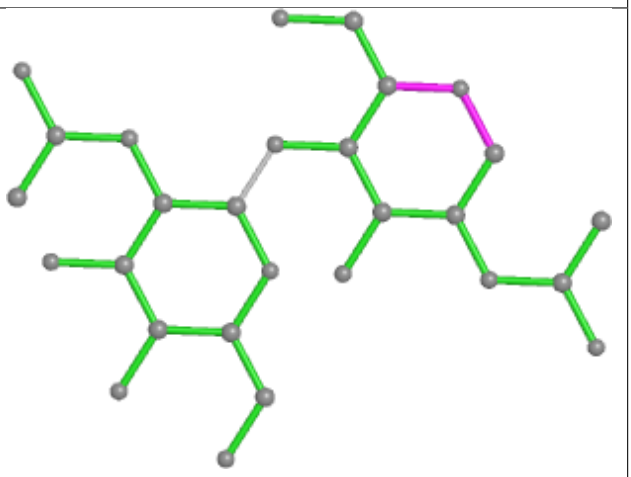


Rings

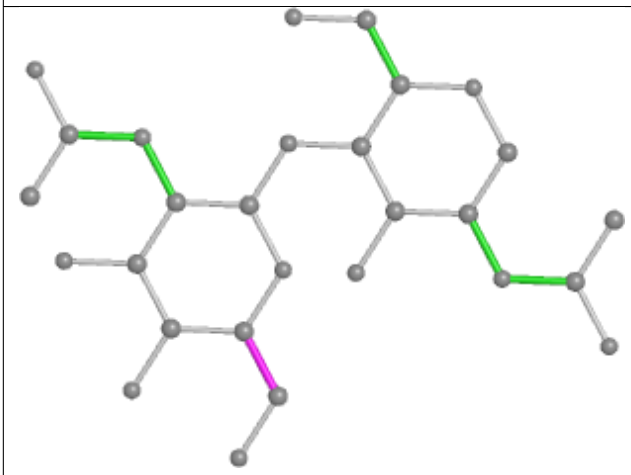
Oligosaccharide Chain H



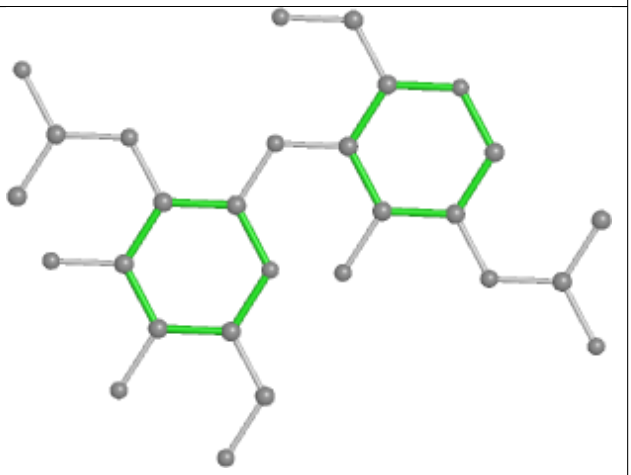
Bond lengths



Bond angles

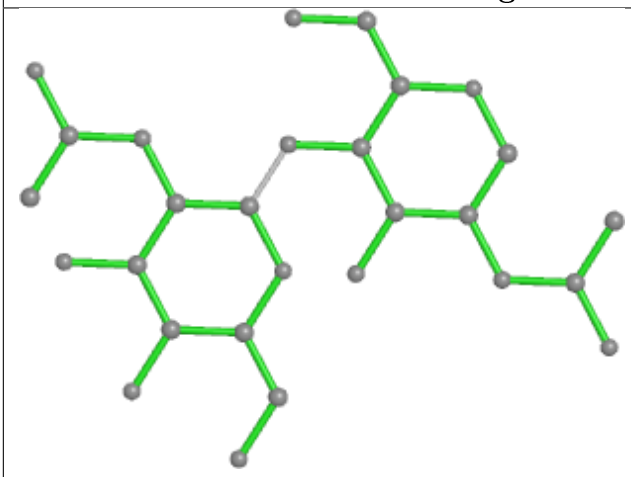


Torsions

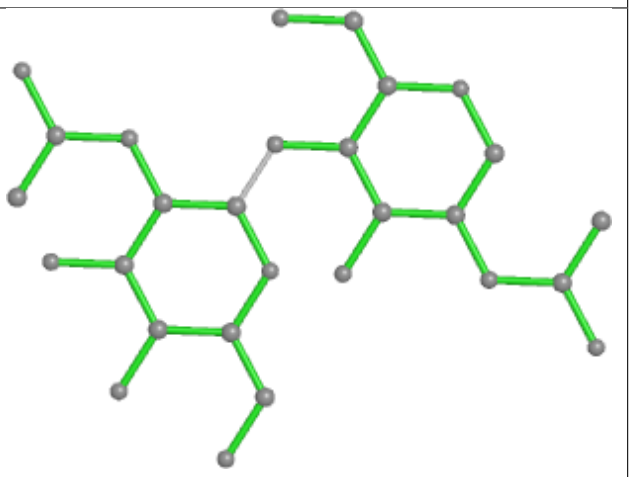


Rings

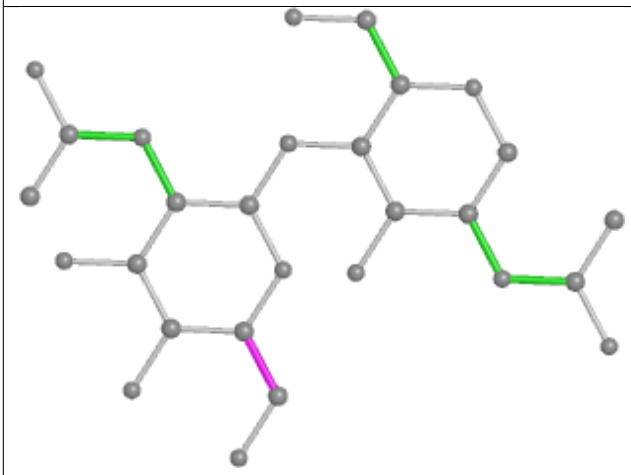
Oligosaccharide Chain J



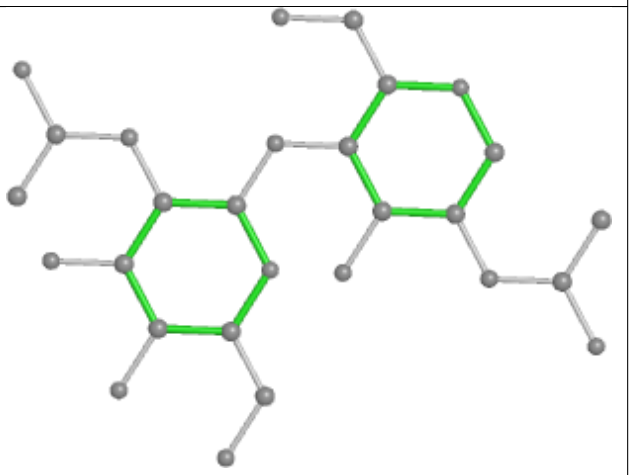
Bond lengths



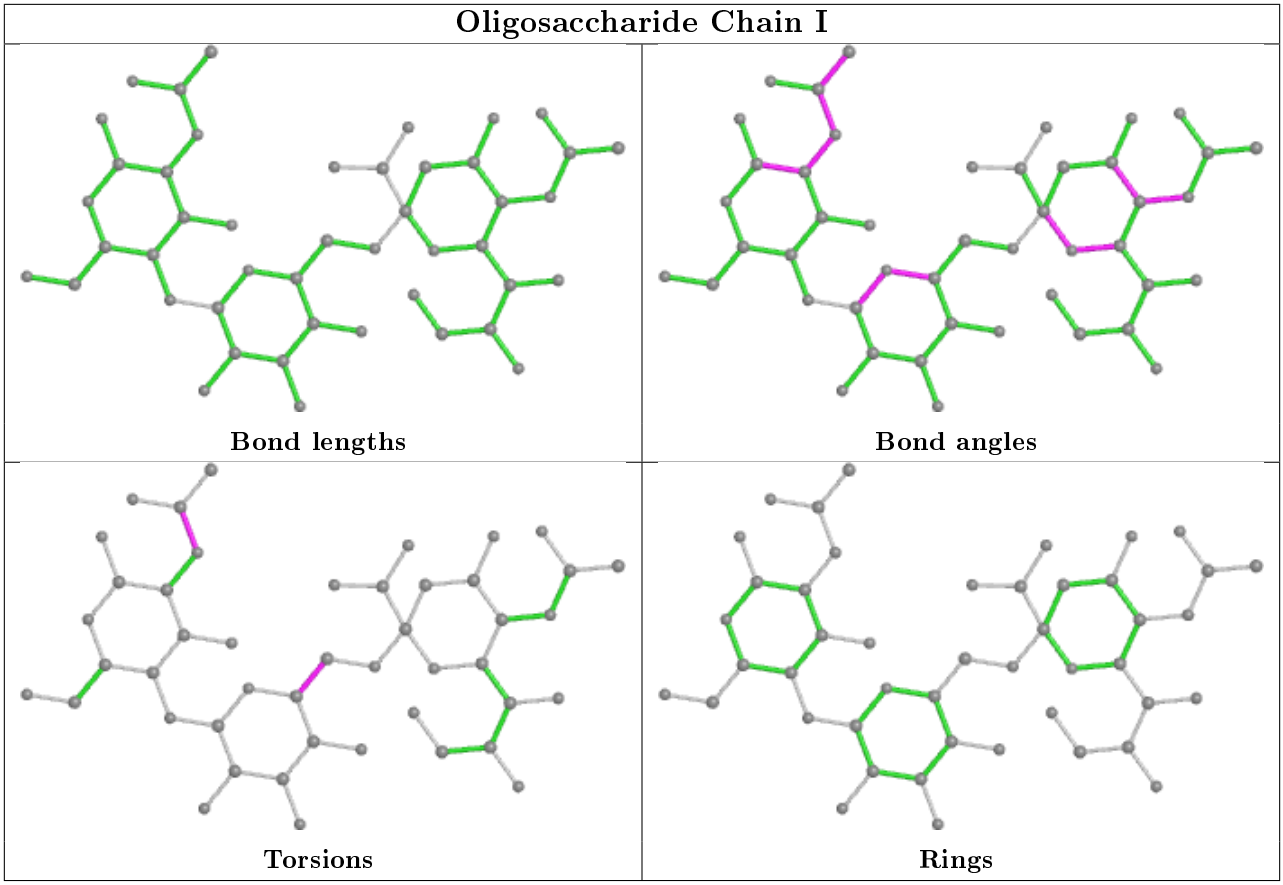
Bond angles



Torsions



Rings



5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
7	NAG	E	403	1	14,14,15	0.51	0	17,19,21	1.10	1 (5%)
7	NAG	F	201	2,6	14,14,15	0.37	0	17,19,21	0.86	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
7	NAG	E	403	1	-	1/6/23/26	0/1/1/1
7	NAG	F	201	2,6	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	E	403	NAG	C1-O5-C5	3.87	117.44	112.19
7	F	201	NAG	C1-O5-C5	2.49	115.56	112.19

There are no chirality outliers.

All (1) torsion outliers are listed below:

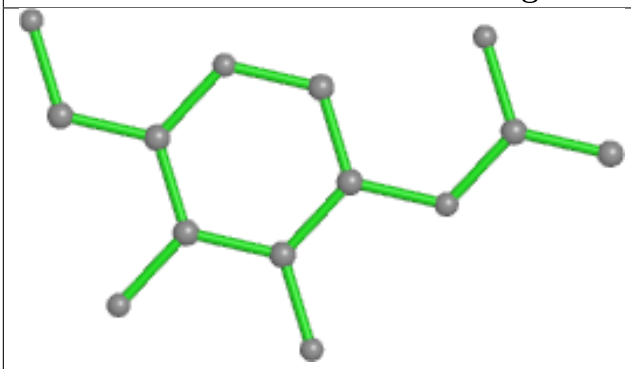
Mol	Chain	Res	Type	Atoms
7	E	403	NAG	O5-C5-C6-O6

There are no ring outliers.

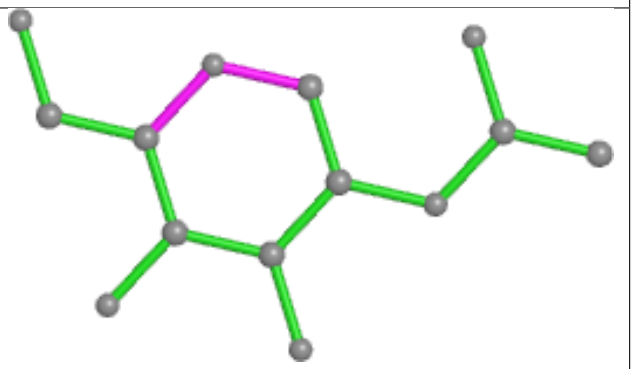
No monomer is involved in short contacts.

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

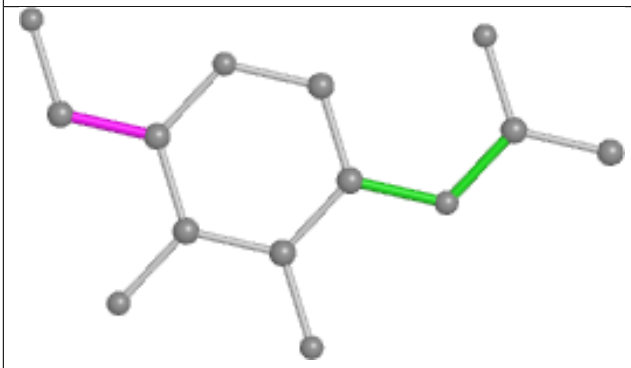
Ligand NAG E 403



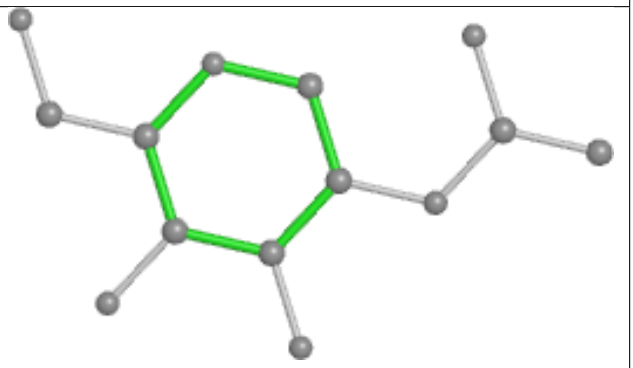
Bond lengths



Bond angles

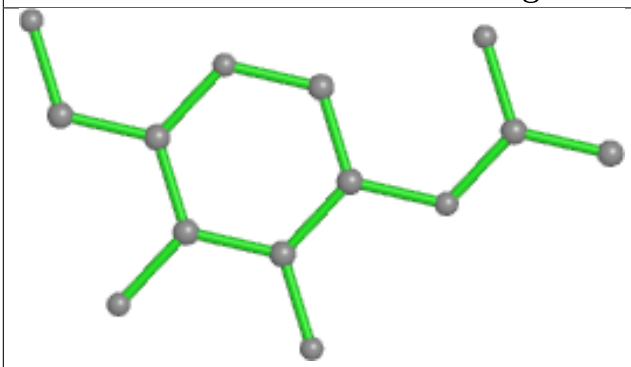


Torsions

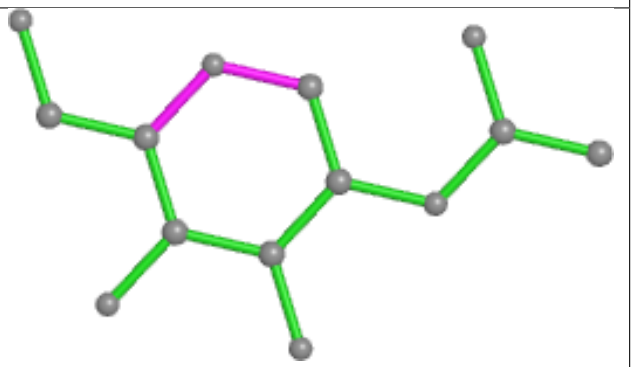


Rings

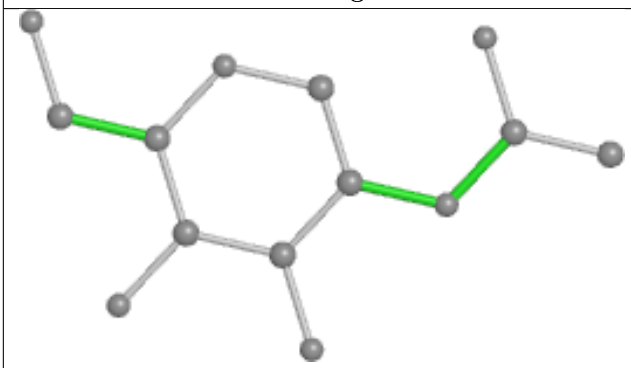
Ligand NAG F 201



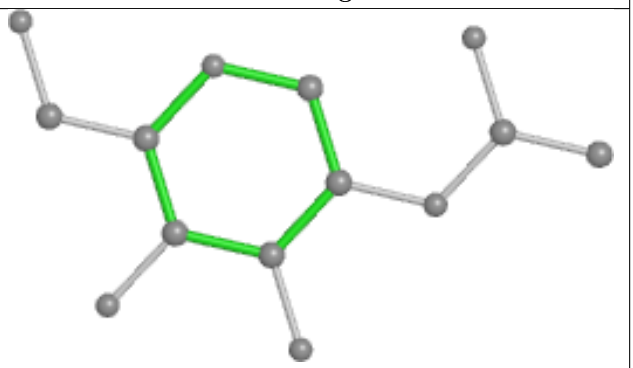
Bond lengths



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	312/324 (96%)	0.42	10 (3%) 47 45	34, 60, 92, 101	0
1	C	311/324 (95%)	0.07	3 (0%) 82 81	34, 46, 65, 76	0
1	E	313/324 (96%)	0.32	15 (4%) 30 29	40, 59, 89, 96	0
2	B	172/177 (97%)	0.14	1 (0%) 89 88	29, 43, 55, 67	0
2	D	172/177 (97%)	0.18	0 100 100	30, 45, 62, 69	0
2	F	172/177 (97%)	0.08	0 100 100	32, 47, 64, 76	0
All	All	1452/1503 (96%)	0.22	29 (1%) 65 63	29, 50, 85, 101	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	133	ASN	4.8
1	E	186	ASP	4.0
1	E	219	LEU	3.7
1	A	190	THR	3.5
1	C	134	GLY	3.1

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

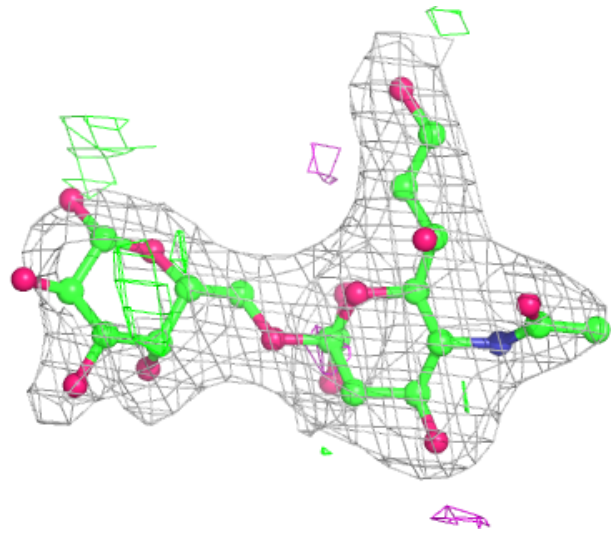
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

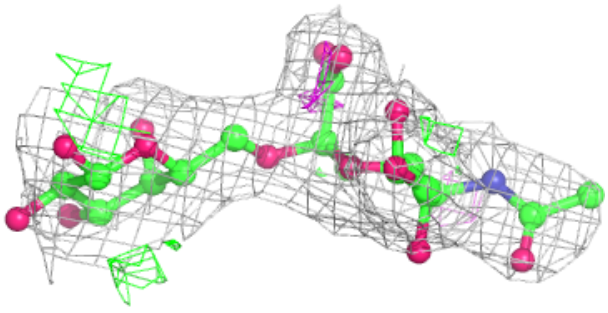
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	J	2	14/15	0.78	0.27	80,83,84,84	0
4	NAG	H	2	14/15	0.80	0.24	61,64,68,68	0
3	GAL	K	1	12/12	0.81	0.26	95,100,101,102	0
3	GAL	G	1	12/12	0.81	0.18	79,87,88,89	0
3	SIA	K	2	20/21	0.88	0.18	87,89,93,93	0
5	NAG	I	1	15/15	0.88	0.22	86,91,92,92	0
3	SIA	G	2	20/21	0.90	0.15	72,74,76,76	0
5	GAL	I	2	11/12	0.90	0.21	71,79,82,83	0
4	NAG	J	1	14/15	0.92	0.13	63,66,70,76	0
5	SIA	I	3	20/21	0.93	0.12	60,63,67,67	0
4	NAG	H	1	14/15	0.95	0.11	46,49,52,56	0

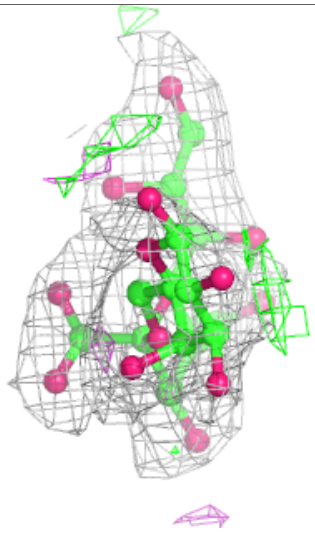
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

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)

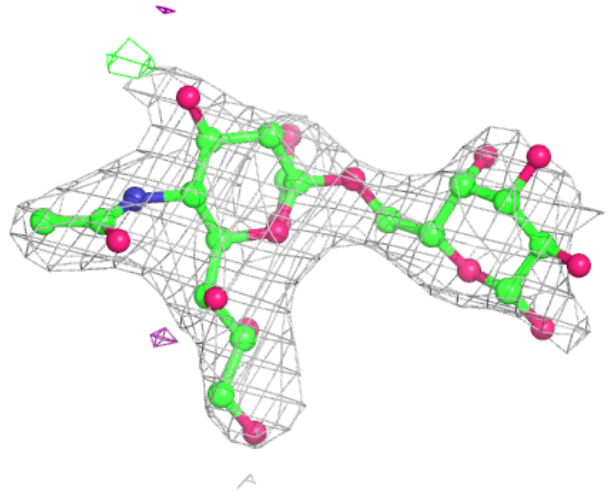


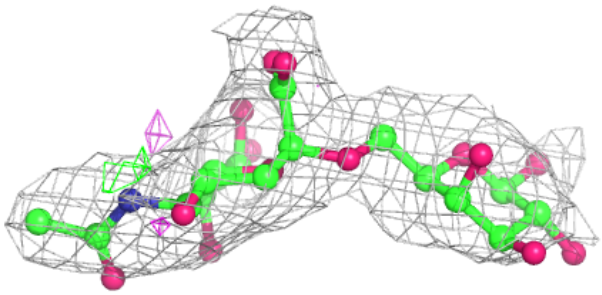


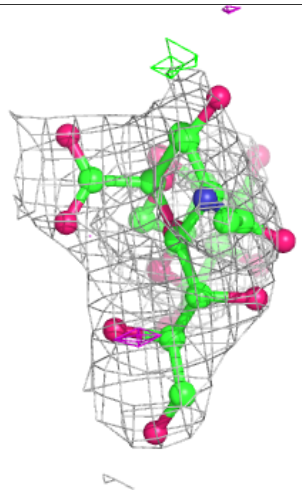


Electron density around Chain K:

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

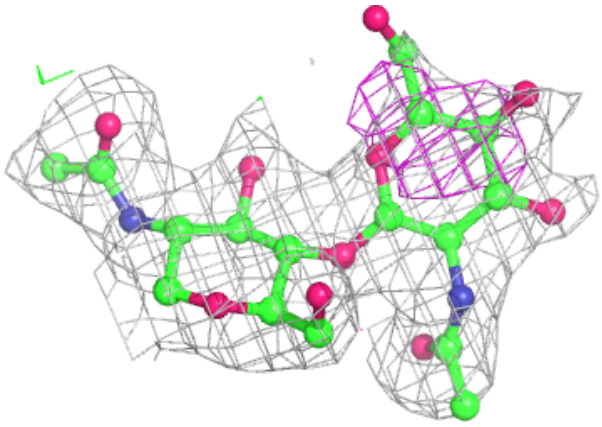


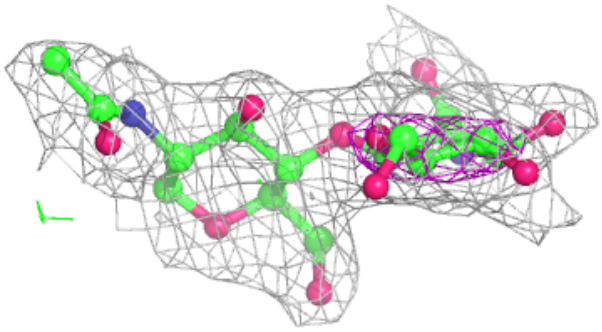




Electron density around Chain H:

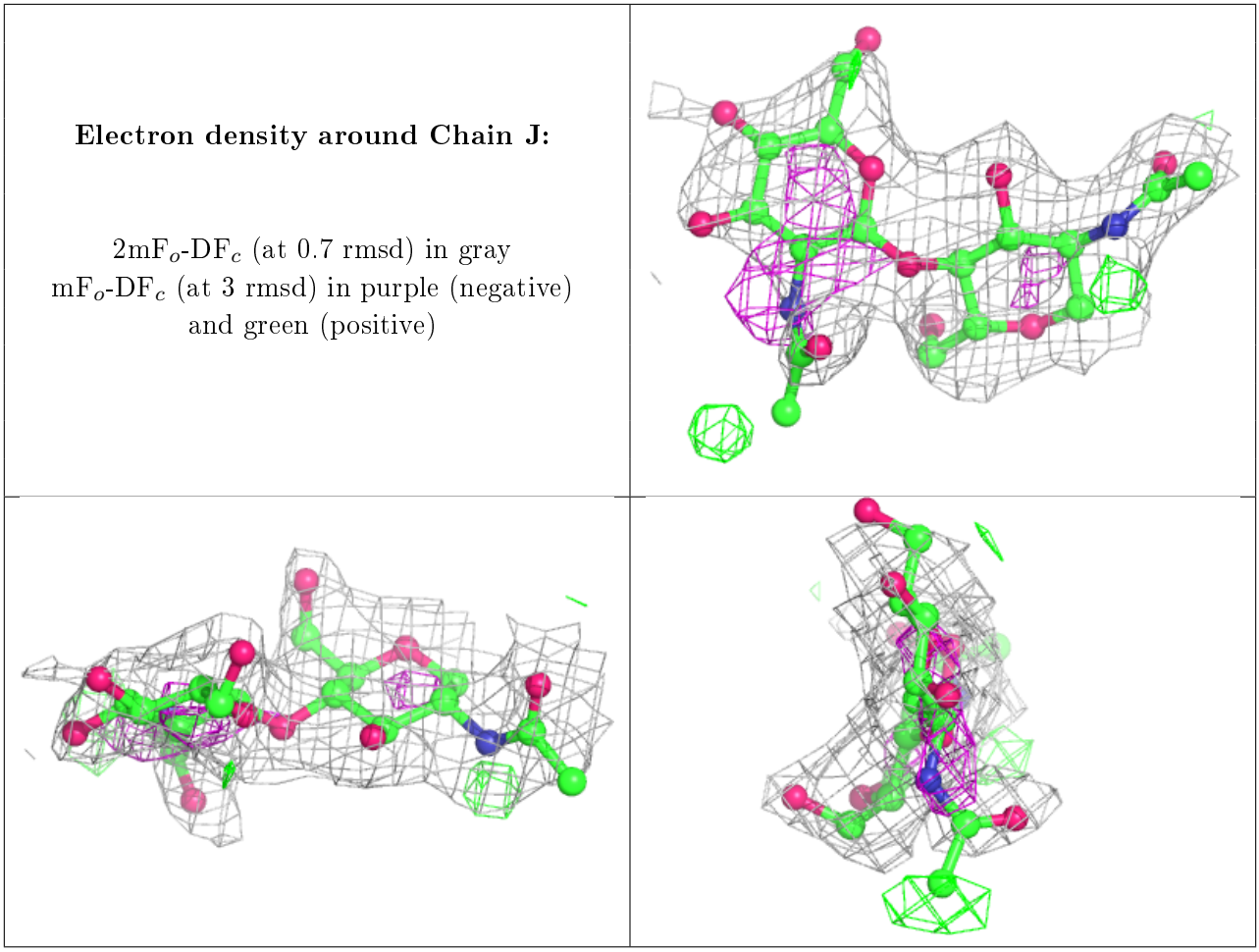
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)





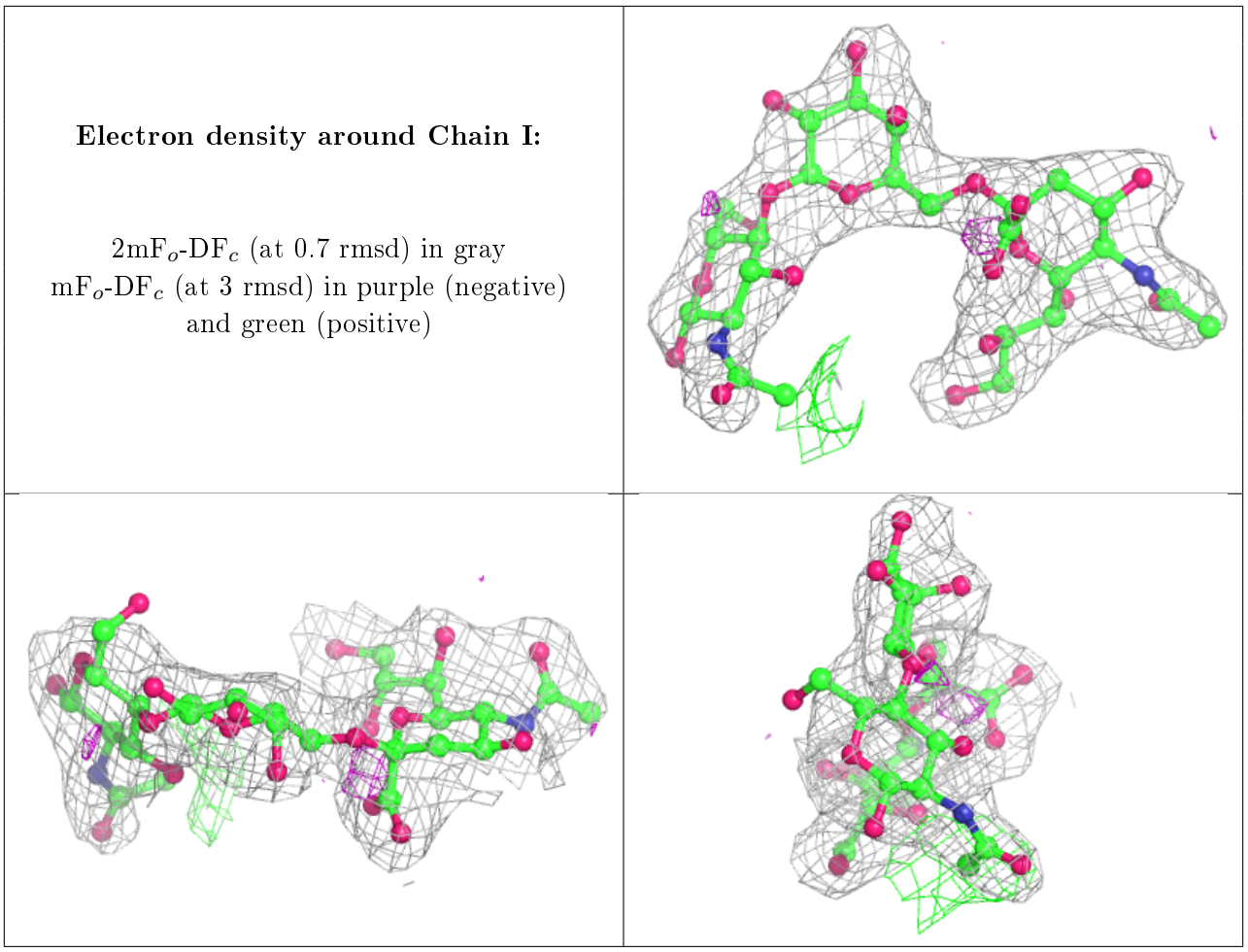
Electron density around Chain J:

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



Electron density around Chain I:

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



6.4 Ligands

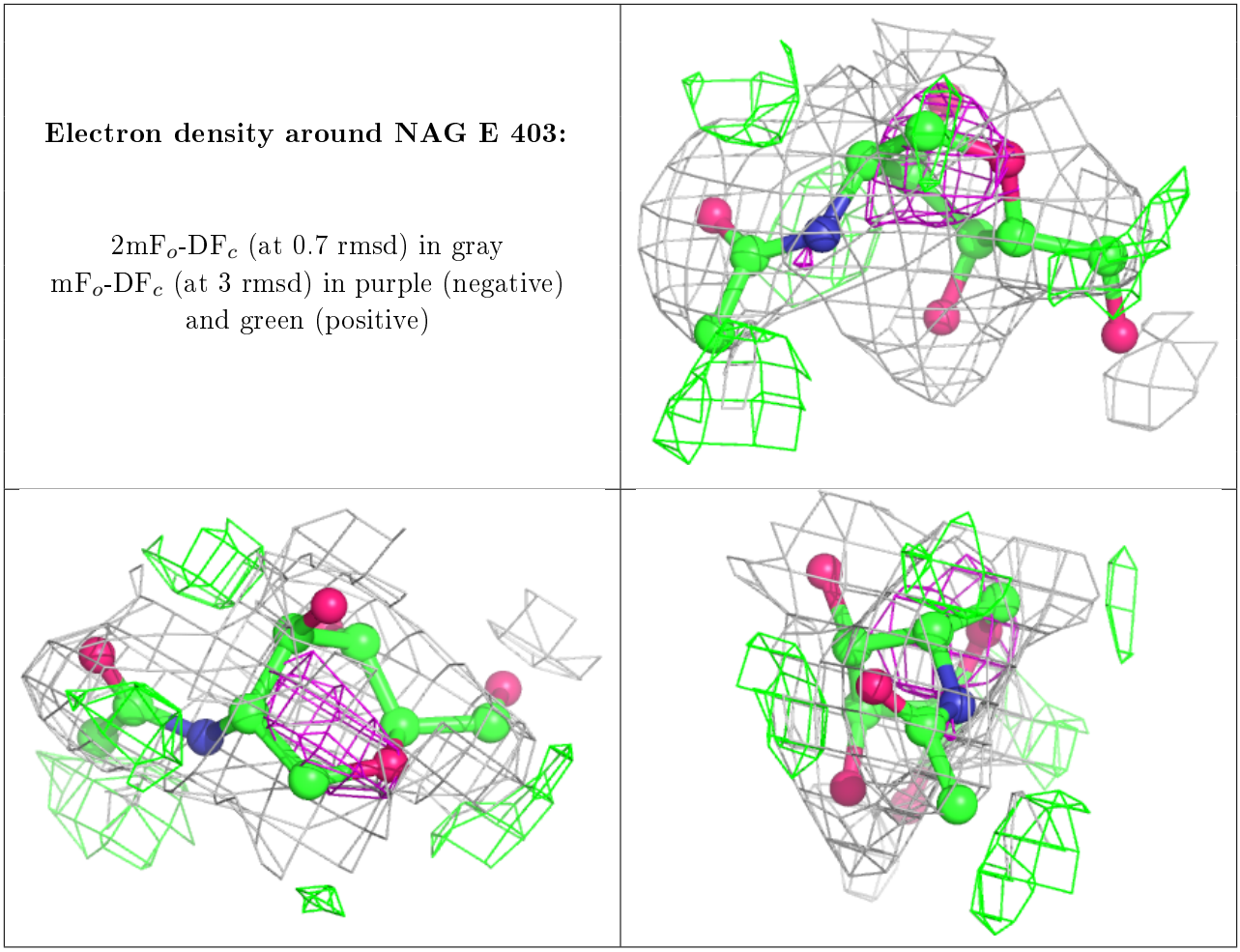
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
6	CA	A	403	1/1	0.44	0.10	79,79,79,79	0
7	NAG	E	403	14/15	0.75	0.31	75,79,81,81	0
7	NAG	F	201	14/15	0.79	0.18	61,65,67,71	0
6	CA	B	203	1/1	0.93	0.14	46,46,46,46	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.

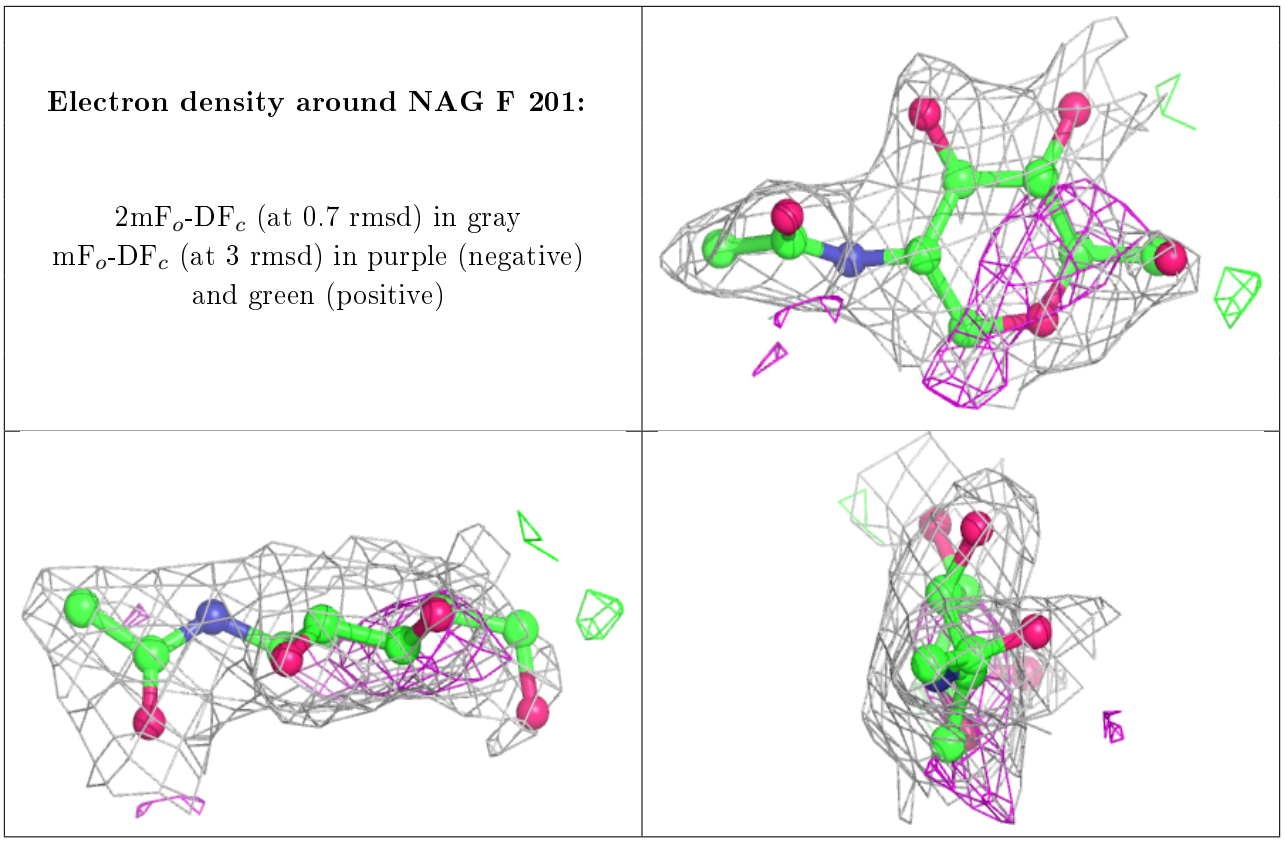
Electron density around NAG E 403:

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



Electron density around NAG F 201:

$2mF_o$-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.