Full wwPDB X-ray Structure Validation Report (i)

Nov 21, 2023 – 05:28 PM JST

PDB ID	:	6JG7
Title	:	Crystal structure of barley exohydrolaseI W286F in complex with methyl 2-thio-beta-sophoroside
Authors	:	Luang, S.; Streltsov, V.A.; Hrmova, M.
Deposited on	:	2019-02-13
Resolution	:	2.16 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

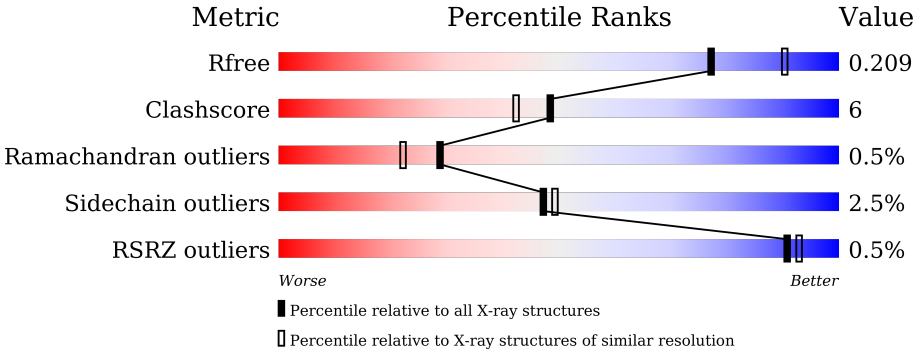
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.16 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	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	609	87% 11% •
2	B	2	50% 50%

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	1PE	A	709	-	-	X	-

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 5140 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 BETA-D-GLUCAN GLUCOHYDROLASE ISOENZYME EXO1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	606	4609	2915	800	868	26	0	5	0

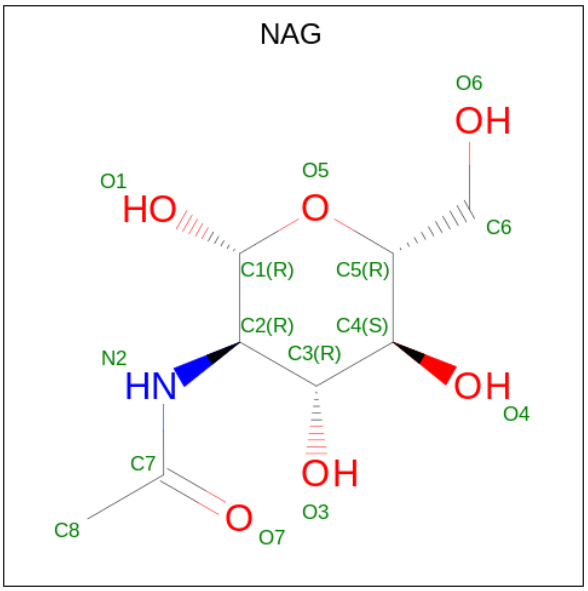
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	HIS	-	expression tag	UNP A0A287SCR5
A	-2	HIS	-	expression tag	UNP A0A287SCR5
A	-1	ALA	-	expression tag	UNP A0A287SCR5
A	0	ALA	-	expression tag	UNP A0A287SCR5
A	286	PHE	TRP	engineered mutation	UNP A0A287SCR5
A	320	LYS	ASN	conflict	UNP A0A287SCR5

- Molecule 2 is an oligosaccharide called beta-D-glucopyranose-(1-2)-methyl 2-thio-beta-D-glucopyranoside.

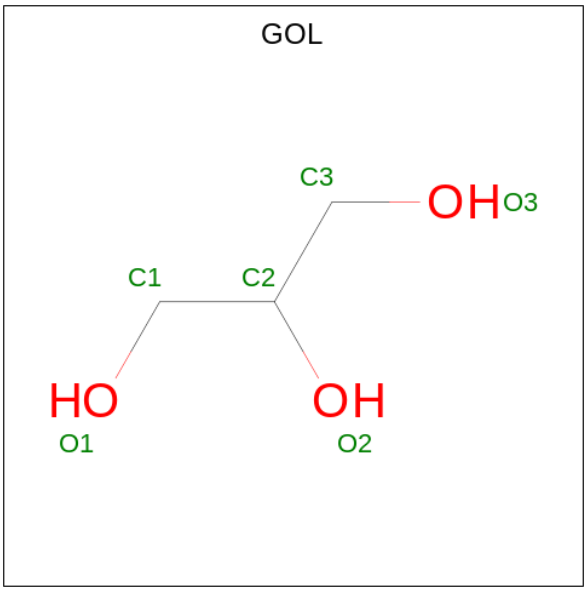
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	O	S			
2	B	2	24	13	10	1	0	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$) (labeled as "Ligand of Interest" by depositor).



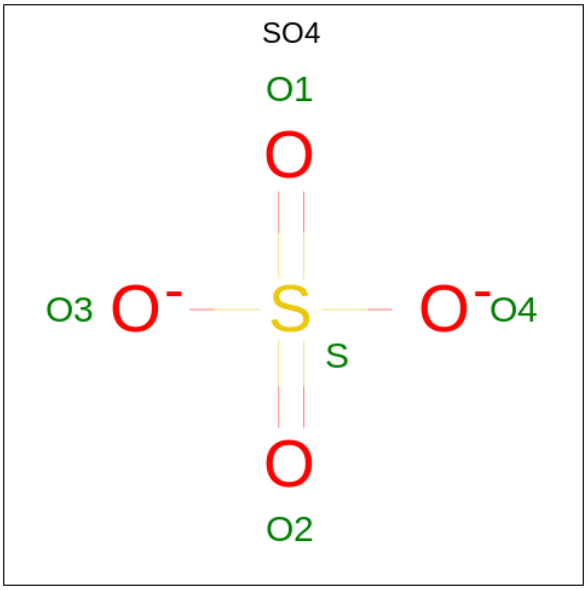
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	A	1	14	8	1	5	0	0
			Total	C	N	O		
3	A	1	14	8	1	5	0	0
			Total	C	N	O		
3	A	1	14	8	1	5	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$) (labeled as "Ligand of Interest" by depositor).



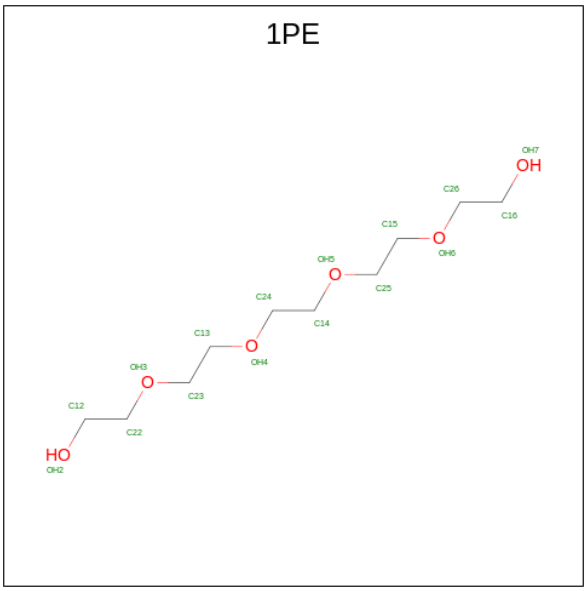
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
4	A	1	6	3	3	0	0
			Total	C	O		
4	A	1	6	3	3	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
5	A	1	5	4	1	0	0
			Total	O	S		
5	A	1	5	4	1	0	0

- Molecule 6 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
6	A	1	16	10	6	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	427	427	427	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: BETA-D-GLUCAN GLUCOHYDROLASE ISOENZYME EXO1

Chain A:

• Molecule 2: beta-D-glucopyranose-(1-2)-methyl 2-thio-beta-D-glucopyranoside

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	99.97Å 99.97Å 182.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.68 – 2.16 87.68 – 2.16	Depositor EDS
% Data completeness (in resolution range)	99.8 (87.68-2.16) 99.8 (87.68-2.16)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.97 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.159 , 0.205 0.169 , 0.209	Depositor DCC
R_{free} test set	2565 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	24.1	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 41.1	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	0.96	EDS
Total number of atoms	5140	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% 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: SO4, U2A, GOL, BGC, NAG, 1PE

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	0.95	1/4727 (0.0%)	0.99	14/6419 (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	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	360	GLU	CD-OE2	-5.63	1.19	1.25

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	126	ARG	NE-CZ-NH1	13.03	126.82	120.30
1	A	126	ARG	NE-CZ-NH2	-11.98	114.31	120.30
1	A	46	ARG	NE-CZ-NH2	8.35	124.47	120.30
1	A	468	ASP	CB-CG-OD1	7.31	124.88	118.30
1	A	468	ASP	CB-CG-OD2	-6.36	112.58	118.30
1	A	79	LYS	CD-CE-NZ	6.30	126.18	111.70
1	A	168	ARG	NE-CZ-NH1	6.19	123.39	120.30
1	A	62	ARG	NE-CZ-NH2	-6.08	117.26	120.30
1	A	440	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	A	291	ARG	NE-CZ-NH2	-5.45	117.57	120.30
1	A	440	ARG	NE-CZ-NH2	-5.27	117.67	120.30
1	A	126	ARG	CD-NE-CZ	5.12	130.77	123.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	211	ASP	CB-CG-OD1	5.12	122.90	118.30
1	A	351	ARG	NE-CZ-NH1	5.11	122.86	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	-3	HIS	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	4609	0	4587	52	0
2	B	24	0	10	1	0
3	A	42	0	39	1	0
4	A	12	0	16	3	0
5	A	10	0	0	0	0
6	A	16	0	22	7	0
7	A	427	0	0	13	0
All	All	5140	0	4674	56	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 6.

All (56) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:466[A]:ASN:ND2	1:A:508:THR:OG1	1.88	1.06
1:A:399:THR:HG23	7:A:840:HOH:O	1.77	0.84
1:A:394:ASN:HD21	1:A:404:LEU:H	1.22	0.84
1:A:122:TYR:CE2	1:A:126:ARG:HD2	2.21	0.75
1:A:156:TRP:HE1	1:A:219:ASN:HD22	1.37	0.73
1:A:566:ARG:NH2	4:A:705:GOL:O2	2.20	0.71
1:A:100:GLN:HE21	1:A:100:GLN:HA	1.56	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:397:THR:OG1	1:A:399:THR:HG22	1.95	0.66
1:A:517[A]:ARG:NH2	7:A:802:HOH:O	2.25	0.66
1:A:181:GLN:HE21	1:A:203:ALA:H	1.45	0.64
1:A:134:GLU:OE2	1:A:377:HIS:HD2	1.82	0.62
1:A:97:VAL:H	1:A:101:ASN:HD21	1.45	0.62
1:A:600:ASN:HD22	3:A:702:NAG:H83	1.64	0.62
1:A:167:ARG:HH21	1:A:171:GLN:HE22	1.46	0.61
1:A:58:GLY:H	1:A:102:ASN:ND2	1.99	0.60
1:A:377:HIS:CE1	4:A:706:GOL:H2	2.36	0.60
6:A:709:1PE:C13	7:A:1140:HOH:O	2.50	0.59
6:A:709:1PE:H131	7:A:1140:HOH:O	2.03	0.58
1:A:374[B]:LYS:NZ	7:A:805:HOH:O	2.36	0.57
6:A:709:1PE:H142	6:A:709:1PE:H162	1.89	0.54
1:A:100:GLN:HE21	1:A:100:GLN:CA	2.21	0.54
1:A:471:PHE:CA	6:A:709:1PE:H121	2.38	0.53
1:A:156:TRP:HE1	1:A:219:ASN:ND2	2.06	0.53
1:A:167:ARG:HH21	1:A:171:GLN:NE2	2.07	0.52
1:A:466[A]:ASN:ND2	1:A:508:THR:HG1	2.01	0.51
1:A:490:THR:HG22	1:A:491:GLU:HG3	1.92	0.51
1:A:517[B]:ARG:HG2	7:A:1066:HOH:O	2.11	0.50
1:A:262:HIS:CE1	1:A:288:GLY:HA3	2.46	0.50
1:A:97:VAL:H	1:A:101:ASN:ND2	2.08	0.50
1:A:377:HIS:HE1	4:A:706:GOL:H2	1.76	0.50
1:A:331:HIS:HD2	7:A:1154:HOH:O	1.95	0.49
1:A:68[B]:LYS:HE3	7:A:888:HOH:O	2.13	0.48
1:A:187:ASP:OD1	1:A:187:ASP:N	2.47	0.48
1:A:471:PHE:HA	6:A:709:1PE:H121	1.95	0.47
1:A:531:GLN:HB2	1:A:532:PRO:HD3	1.97	0.47
6:A:709:1PE:H142	6:A:709:1PE:C16	2.45	0.47
1:A:-3:HIS:HB3	7:A:945:HOH:O	2.15	0.46
1:A:188:PHE:CZ	1:A:192:MET:HB2	2.51	0.45
1:A:437:ASP:HB3	1:A:441:THR:HG21	1.97	0.45
1:A:593:LEU:HD23	1:A:594:GLY:N	2.32	0.45
1:A:57:GLY:N	2:B:1:U2A:C7	2.80	0.45
1:A:593:LEU:HD23	1:A:593:LEU:C	2.37	0.44
1:A:234:HIS:HE1	7:A:1147:HOH:O	2.00	0.44
1:A:167:ARG:NH2	1:A:171:GLN:HE22	2.14	0.44
1:A:262:HIS:HE1	1:A:285:ASP:H	1.66	0.44
1:A:592:ARG:NH2	7:A:816:HOH:O	2.46	0.43
1:A:134:GLU:OE2	1:A:377:HIS:CD2	2.68	0.43
1:A:517[A]:ARG:HB2	7:A:1066:HOH:O	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:524:SER:O	1:A:544:TRP:HA	2.19	0.42
1:A:404:LEU:HD23	1:A:406:LEU:HG	2.02	0.42
1:A:63:LYS:HE3	7:A:1028:HOH:O	2.19	0.41
1:A:471:PHE:N	6:A:709:1PE:H121	2.35	0.41
1:A:125:LYS:O	1:A:129:GLU:HG3	2.20	0.41
1:A:235:MET:N	1:A:236:PRO:CD	2.84	0.41
1:A:262:HIS:CE1	1:A:285:ASP:H	2.39	0.40
1:A:71:GLN:HE22	1:A:366:PRO:HA	1.85	0.40

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	608/609 (100%)	587 (96%)	18 (3%)	3 (0%)	29 22

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	-2	HIS
1	A	186	LYS
1	A	505	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	491/490 (100%)	478 (97%)	13 (3%)	46	47

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	GLU
1	A	54	LEU
1	A	63	LYS
1	A	79	LYS
1	A	90	MET
1	A	98	HIS
1	A	100	GLN
1	A	186	LYS
1	A	290	ASP
1	A	440	ARG
1	A	517[A]	ARG
1	A	517[B]	ARG
1	A	592	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (16) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	71	GLN
1	A	100	GLN
1	A	101	ASN
1	A	102	ASN
1	A	112	ASN
1	A	171	GLN
1	A	181	GLN
1	A	199	ASN
1	A	219	ASN
1	A	234	HIS
1	A	262	HIS
1	A	265	GLN
1	A	331	HIS
1	A	371	GLN
1	A	377	HIS
1	A	394	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)

3 non-standard protein/DNA/RNA residues are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	NAG	A	701	1	14,14,15	0.72	0	17,19,21	1.48	3 (17%)
3	NAG	A	702	1	14,14,15	0.88	1 (7%)	17,19,21	2.87	6 (35%)
3	NAG	A	703	1	14,14,15	0.63	0	17,19,21	1.83	4 (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	A	701	1	-	0/6/23/26	0/1/1/1
3	NAG	A	702	1	-	2/6/23/26	0/1/1/1
3	NAG	A	703	1	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	702	NAG	C1-C2	2.21	1.55	1.52

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	702	NAG	C4-C3-C2	-5.73	102.62	111.02
3	A	702	NAG	C8-C7-N2	5.32	125.11	116.10
3	A	703	NAG	C1-O5-C5	5.25	119.31	112.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	702	NAG	C1-C2-N2	5.06	119.13	110.49
3	A	702	NAG	C2-N2-C7	3.59	128.01	122.90
3	A	702	NAG	O5-C5-C6	3.56	112.78	107.20
3	A	702	NAG	O7-C7-C8	-3.27	115.98	122.06
3	A	701	NAG	O3-C3-C2	-2.69	103.90	109.47
3	A	701	NAG	O6-C6-C5	-2.49	102.75	111.29
3	A	703	NAG	O3-C3-C4	-2.43	104.74	110.35
3	A	701	NAG	O7-C7-C8	-2.40	117.60	122.06
3	A	703	NAG	O5-C1-C2	2.33	114.97	111.29
3	A	703	NAG	C1-C2-N2	-2.29	106.57	110.49

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	702	NAG	C8-C7-N2-C2
3	A	702	NAG	O7-C7-N2-C2
3	A	703	NAG	C4-C5-C6-O6
3	A	703	NAG	O5-C5-C6-O6

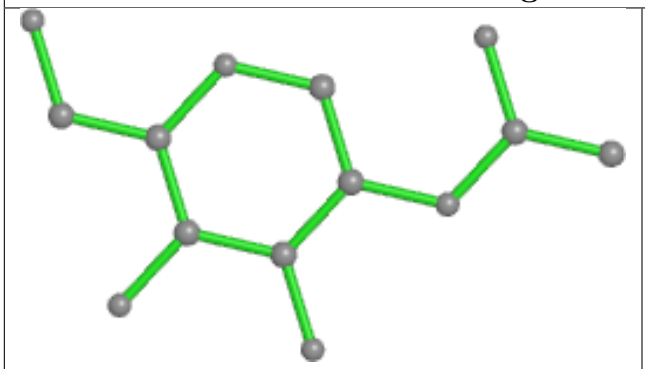
There are no ring outliers.

1 monomer is involved in 1 short contact:

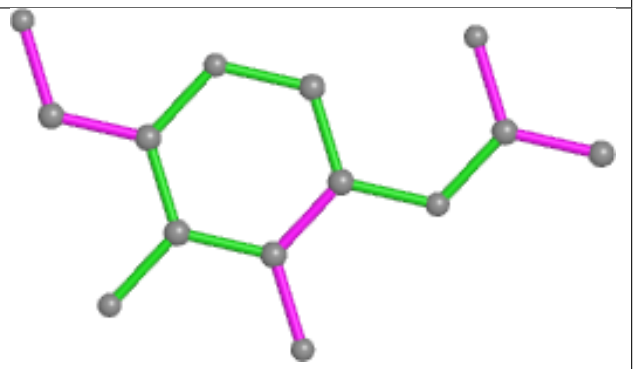
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	702	NAG	1	0

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

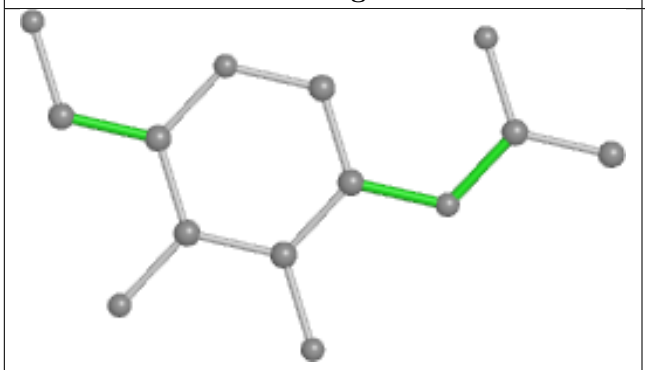
Ligand NAG A 701



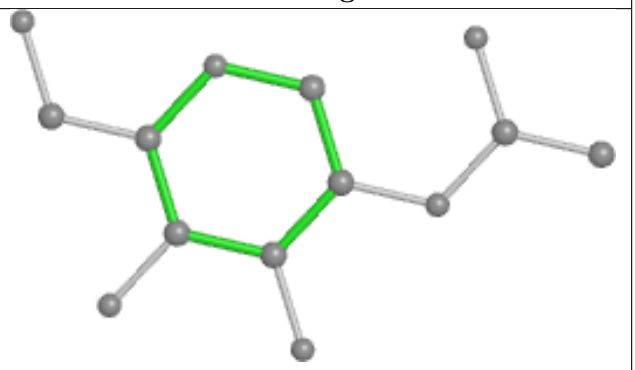
Bond lengths



Bond angles

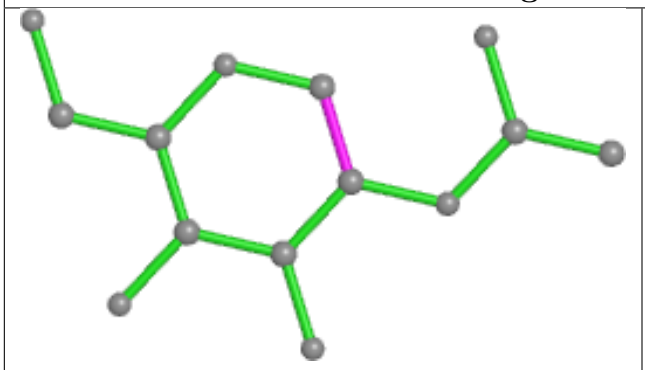


Torsions

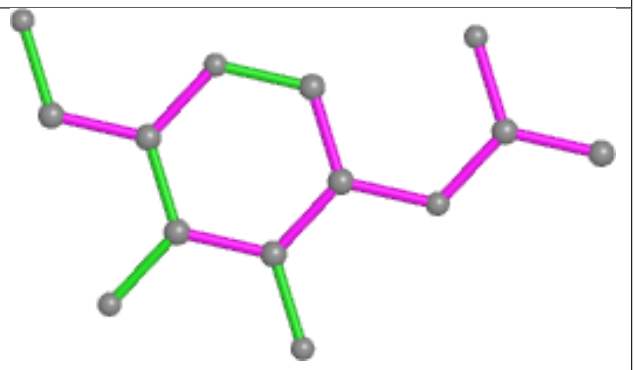


Rings

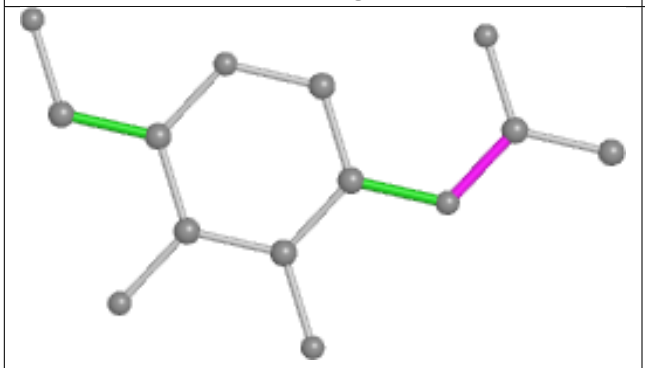
Ligand NAG A 702



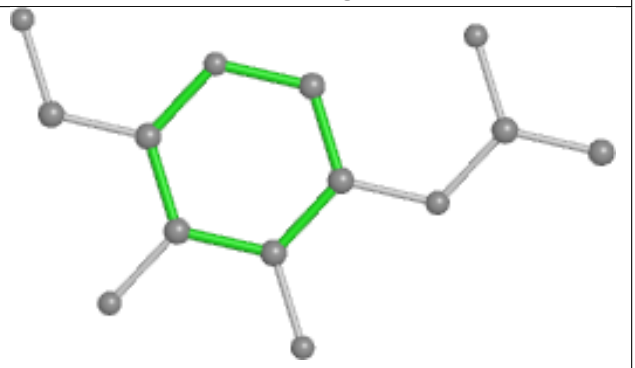
Bond lengths



Bond angles

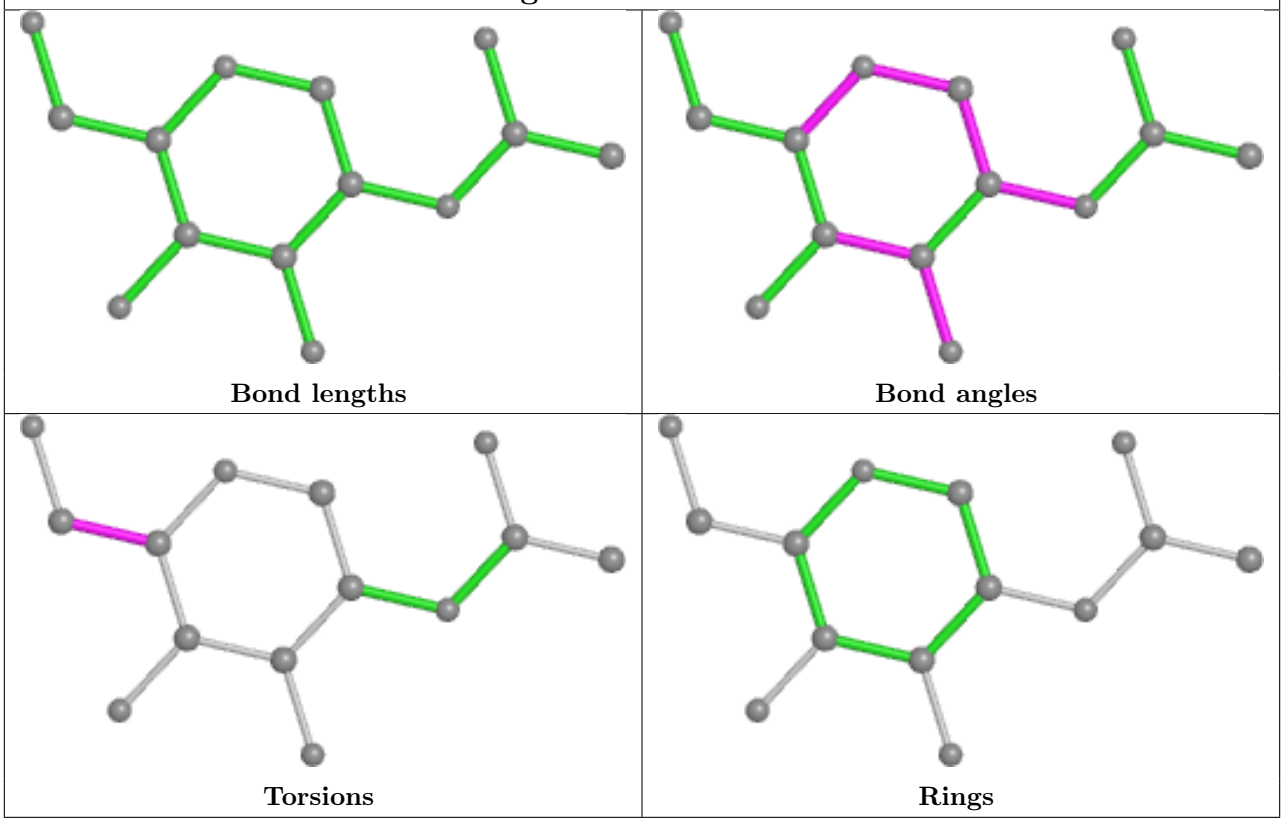


Torsions



Rings

Ligand NAG A 703



5.5 Carbohydrates (i)

2 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	$\#\|Z\| > 2$
2	U2A	B	1	2	12,13,13	1.72	3 (25%)	14,18,18	4.60	10 (71%)
2	BGC	B	2	2	11,11,12	1.60	3 (27%)	15,15,17	1.66	3 (20%)

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
2	U2A	B	1	2	-	4/4/24/24	0/1/1/1
2	BGC	B	2	2	-	2/2/19/22	0/1/1/1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	U2A	O1-C1	3.88	1.46	1.40
2	B	2	BGC	O5-C5	2.96	1.49	1.43
2	B	2	BGC	C1-C2	2.70	1.58	1.52
2	B	1	U2A	O3-C3	2.59	1.49	1.43
2	B	1	U2A	C1-C2	2.41	1.57	1.52
2	B	2	BGC	O5-C1	2.35	1.47	1.43

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1	U2A	C7-O1-C1	-11.35	95.76	113.27
2	B	1	U2A	O1-C1-C2	9.21	122.24	108.10
2	B	1	U2A	O4-C4-C5	4.19	119.70	109.30
2	B	1	U2A	O5-C1-O1	-3.63	102.56	110.97
2	B	2	BGC	C1-C2-C3	3.34	113.78	109.67
2	B	1	U2A	O3-C3-C2	3.33	115.97	109.12
2	B	1	U2A	C3-C4-C5	-3.17	104.58	110.24
2	B	1	U2A	O3-C3-C4	3.12	117.57	110.35
2	B	2	BGC	O5-C1-C2	2.93	115.30	110.77
2	B	2	BGC	C1-O5-C5	2.87	116.08	112.19
2	B	1	U2A	C6-C5-C4	2.83	119.64	113.00
2	B	1	U2A	O6-C6-C5	-2.35	103.24	111.29
2	B	1	U2A	O5-C5-C4	-2.02	106.03	109.69

There are no chirality outliers.

All (6) torsion outliers are listed below:

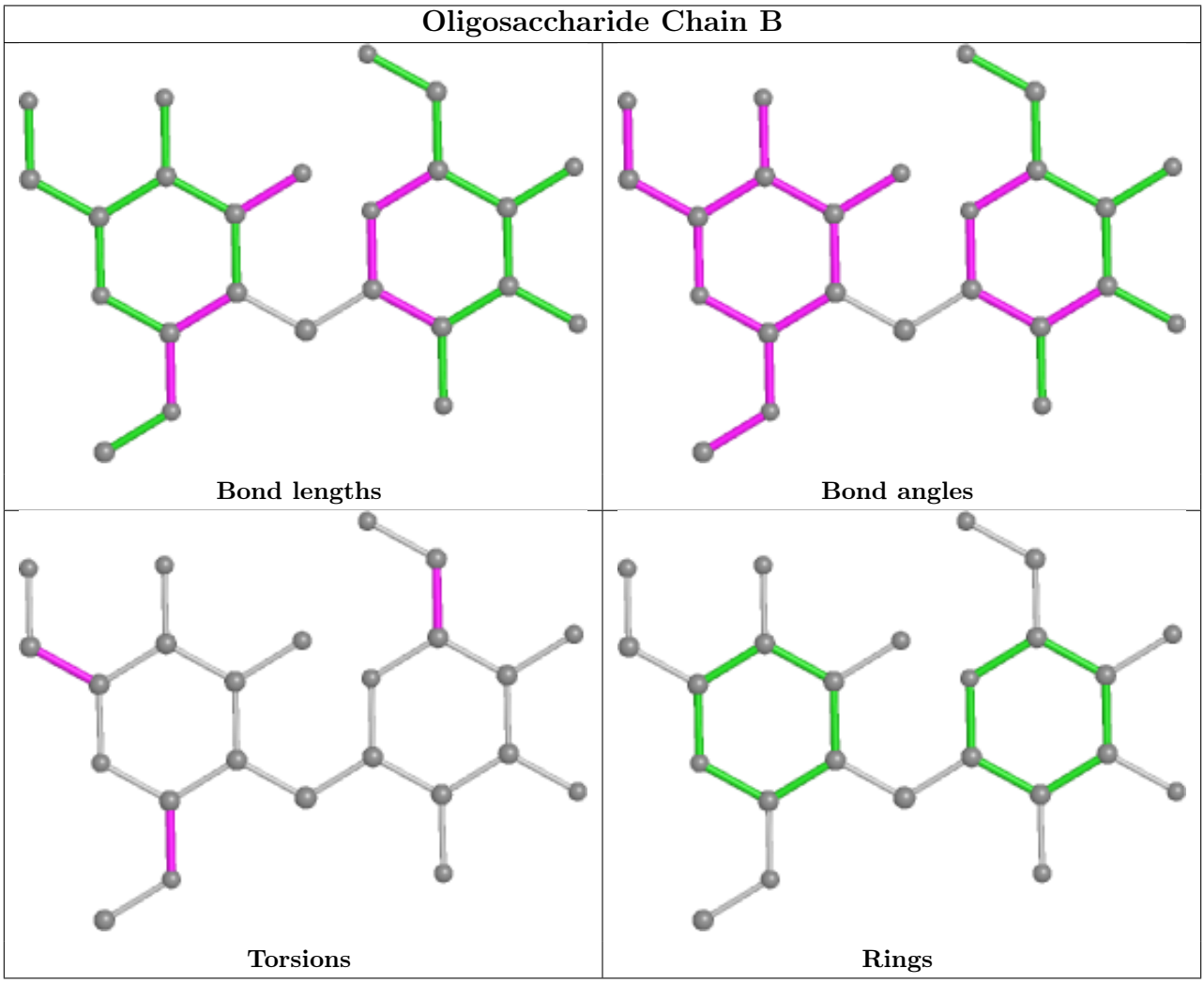
Mol	Chain	Res	Type	Atoms
2	B	1	U2A	O5-C1-O1-C7
2	B	1	U2A	O5-C5-C6-O6
2	B	1	U2A	C4-C5-C6-O6
2	B	2	BGC	O5-C5-C6-O6
2	B	2	BGC	C4-C5-C6-O6
2	B	1	U2A	C2-C1-O1-C7

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1	U2A	1	0

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



5.6 Ligand geometry (i)

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	$\#\|Z\| > 2$
4	GOL	A	706	-	5,5,5	0.93	0	5,5,5	0.78	0
3	NAG	A	702	1	14,14,15	0.88	1 (7%)	17,19,21	2.87	6 (35%)
6	1PE	A	709	-	15,15,15	0.89	0	14,14,14	1.01	0
3	NAG	A	703	1	14,14,15	0.63	0	17,19,21	1.83	4 (23%)
5	SO4	A	707	-	4,4,4	0.41	0	6,6,6	0.60	0
5	SO4	A	708	-	4,4,4	0.51	0	6,6,6	0.95	0
3	NAG	A	701	1	14,14,15	0.72	0	17,19,21	1.48	3 (17%)
4	GOL	A	705	-	5,5,5	0.48	0	5,5,5	0.94	0

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	706	-	-	4/4/4/4	-
3	NAG	A	702	1	-	2/6/23/26	0/1/1/1
6	1PE	A	709	-	-	9/13/13/13	-
3	NAG	A	703	1	-	2/6/23/26	0/1/1/1
3	NAG	A	701	1	-	0/6/23/26	0/1/1/1
4	GOL	A	705	-	-	4/4/4/4	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	702	NAG	C1-C2	2.21	1.55	1.52

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	702	NAG	C4-C3-C2	-5.73	102.62	111.02
3	A	702	NAG	C8-C7-N2	5.32	125.11	116.10
3	A	703	NAG	C1-O5-C5	5.25	119.31	112.19
3	A	702	NAG	C1-C2-N2	5.06	119.13	110.49
3	A	702	NAG	C2-N2-C7	3.59	128.01	122.90
3	A	702	NAG	O5-C5-C6	3.56	112.78	107.20
3	A	702	NAG	O7-C7-C8	-3.27	115.98	122.06
3	A	701	NAG	O3-C3-C2	-2.69	103.90	109.47
3	A	701	NAG	O6-C6-C5	-2.49	102.75	111.29
3	A	703	NAG	O3-C3-C4	-2.43	104.74	110.35

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	701	NAG	O7-C7-C8	-2.40	117.60	122.06
3	A	703	NAG	O5-C1-C2	2.33	114.97	111.29
3	A	703	NAG	C1-C2-N2	-2.29	106.57	110.49

There are no chirality outliers.

All (21) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	705	GOL	C1-C2-C3-O3
4	A	706	GOL	O1-C1-C2-C3
3	A	702	NAG	C8-C7-N2-C2
3	A	702	NAG	O7-C7-N2-C2
3	A	703	NAG	C4-C5-C6-O6
6	A	709	1PE	OH4-C13-C23-OH3
6	A	709	1PE	OH5-C14-C24-OH4
4	A	705	GOL	O2-C2-C3-O3
6	A	709	1PE	OH6-C15-C25-OH5
3	A	703	NAG	O5-C5-C6-O6
4	A	705	GOL	O1-C1-C2-C3
4	A	706	GOL	C1-C2-C3-O3
4	A	706	GOL	O2-C2-C3-O3
6	A	709	1PE	OH2-C12-C22-OH3
6	A	709	1PE	OH7-C16-C26-OH6
4	A	705	GOL	O1-C1-C2-O2
6	A	709	1PE	C25-C15-OH6-C26
4	A	706	GOL	O1-C1-C2-O2
6	A	709	1PE	C13-C23-OH3-C22
6	A	709	1PE	C15-C25-OH5-C14
6	A	709	1PE	C23-C13-OH4-C24

There are no ring outliers.

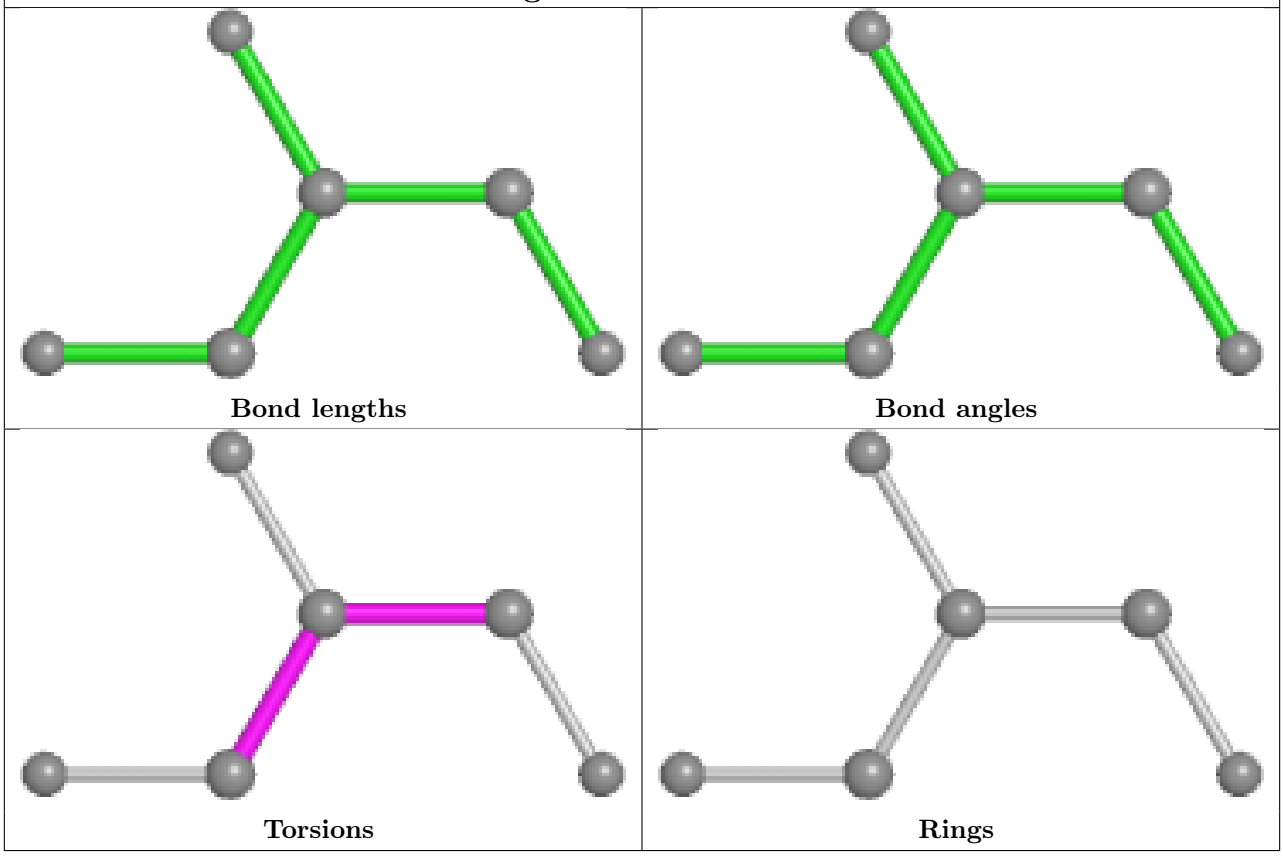
4 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	706	GOL	2	0
3	A	702	NAG	1	0
6	A	709	1PE	7	0
4	A	705	GOL	1	0

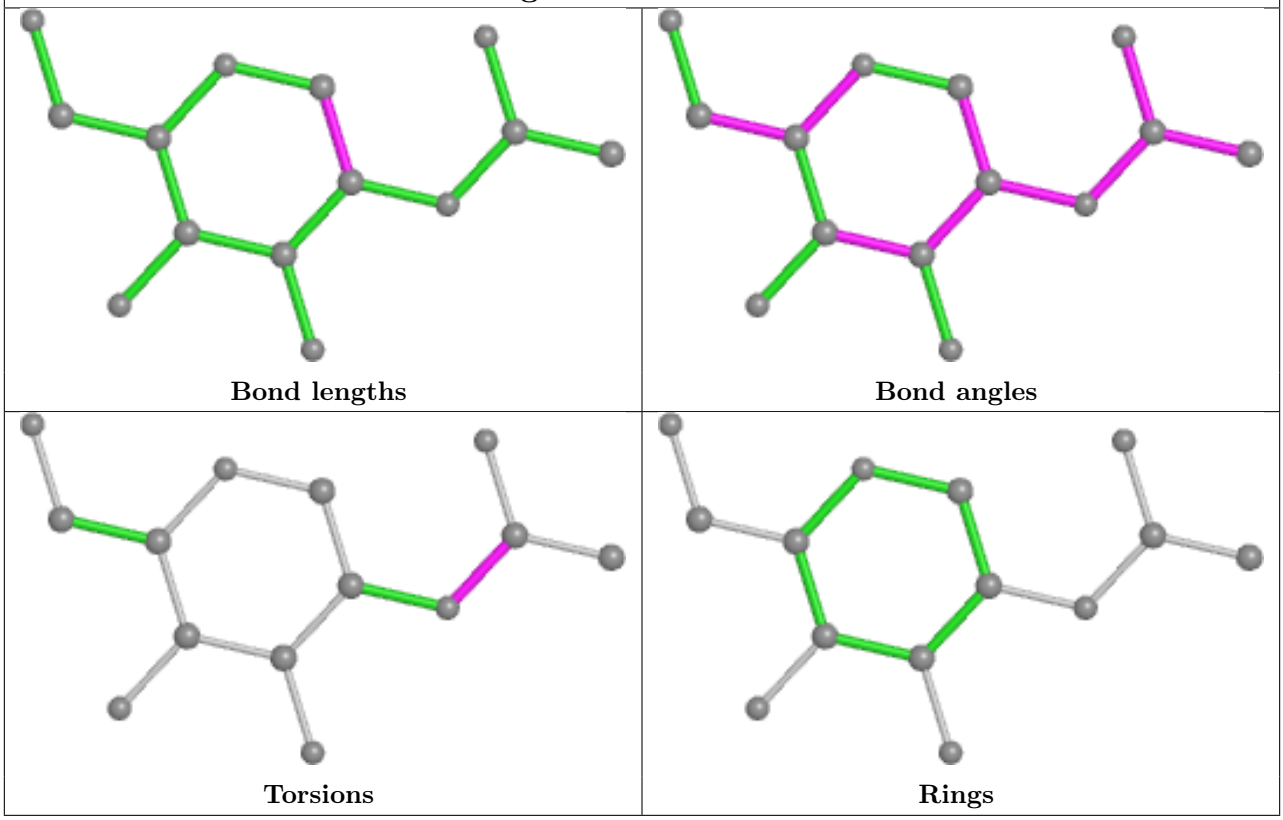
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In

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.

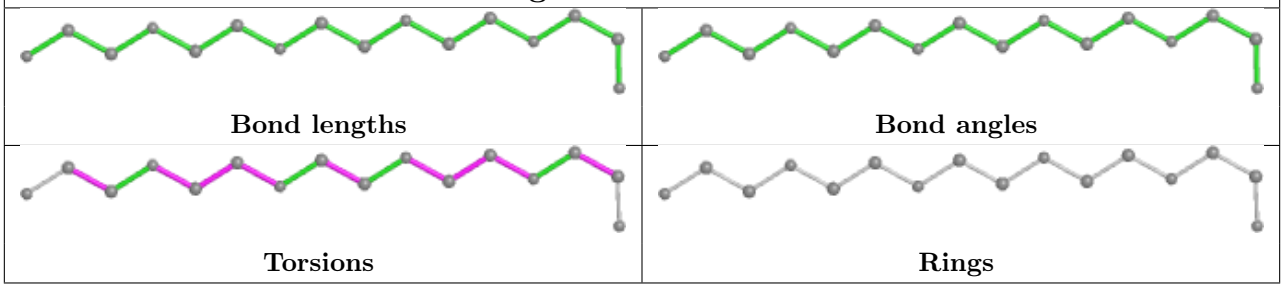
Ligand GOL A 706



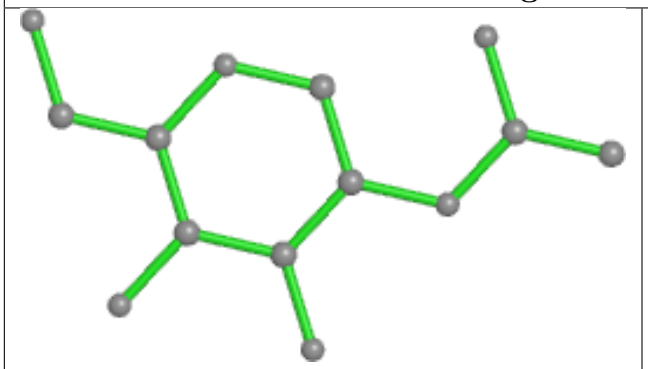
Ligand NAG A 702



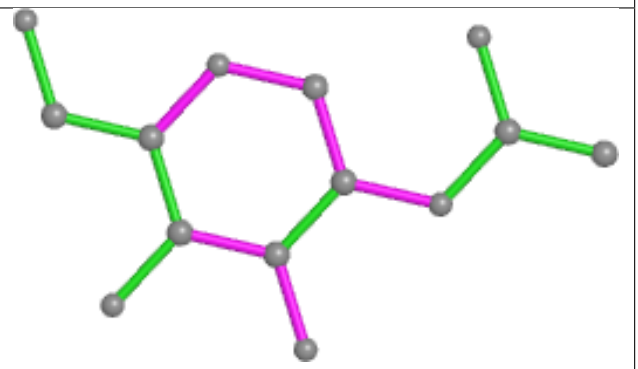
Ligand 1PE A 709



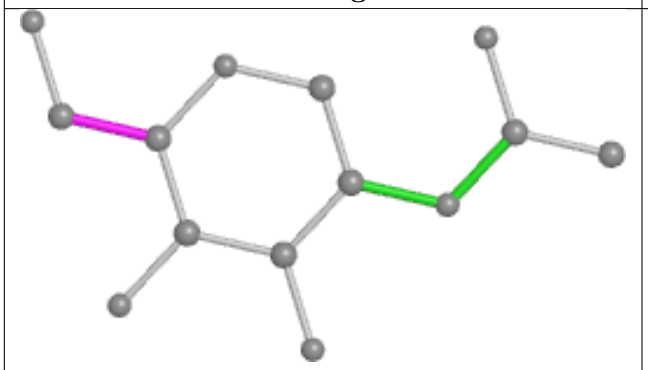
Ligand NAG A 703



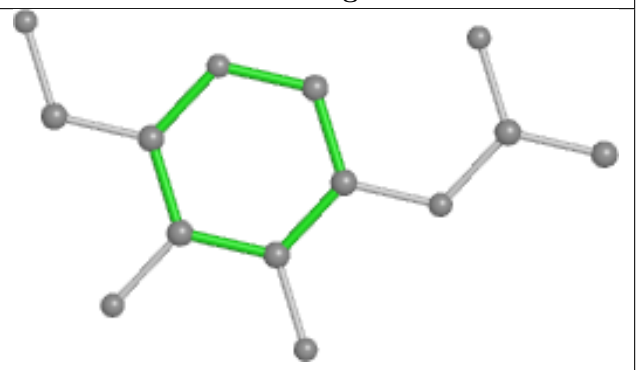
Bond lengths



Bond angles

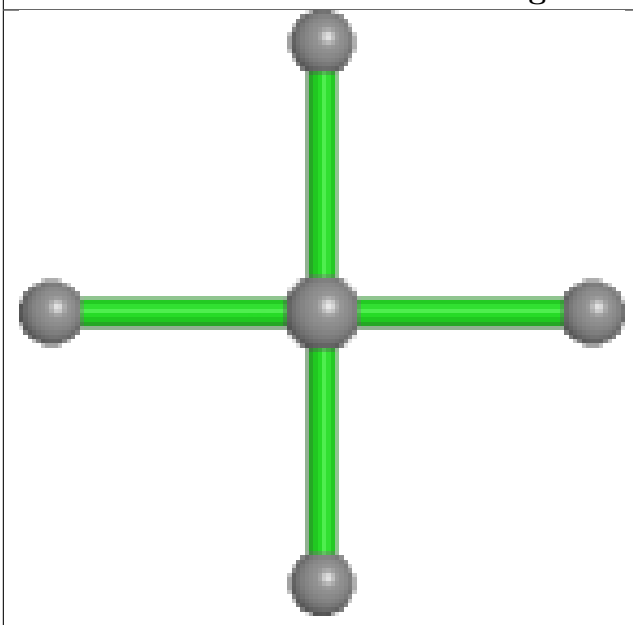


Torsions

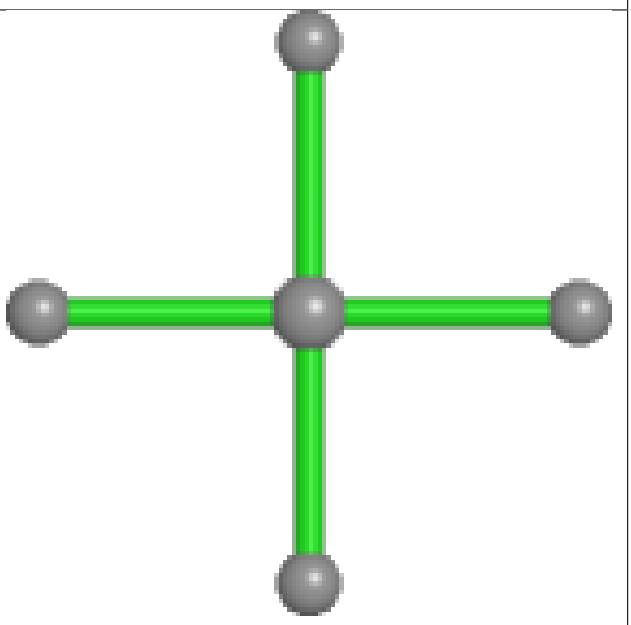


Rings

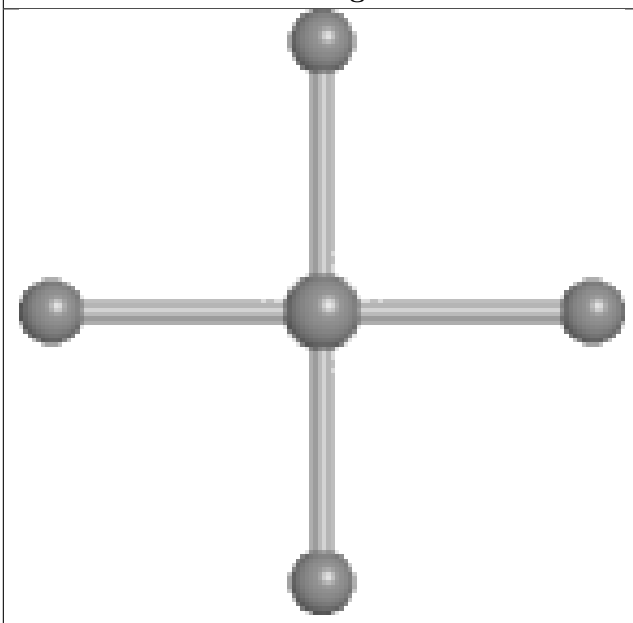
Ligand SO4 A 707



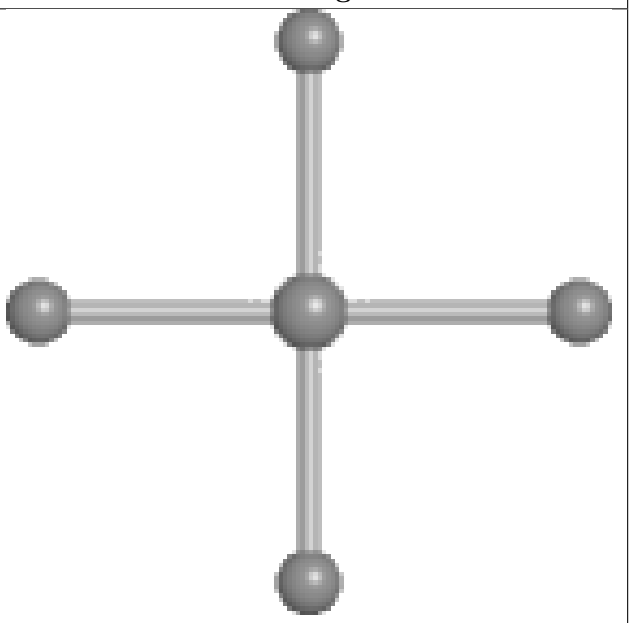
Bond lengths



Bond angles

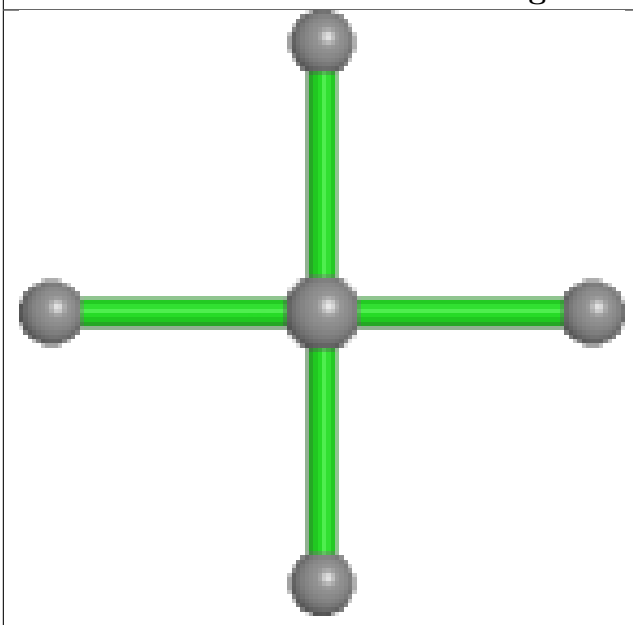


Torsions

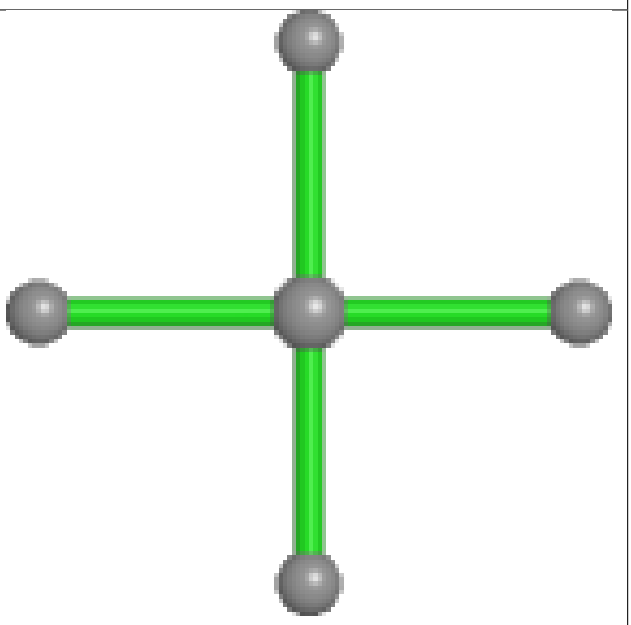


Rings

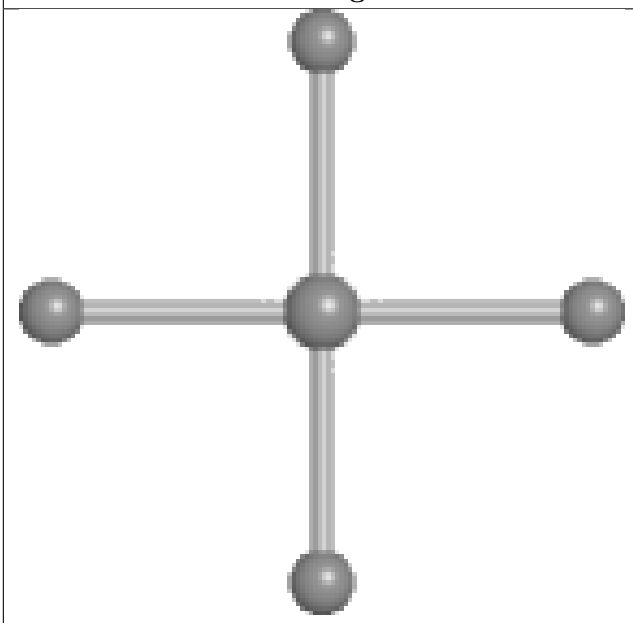
Ligand SO4 A 708



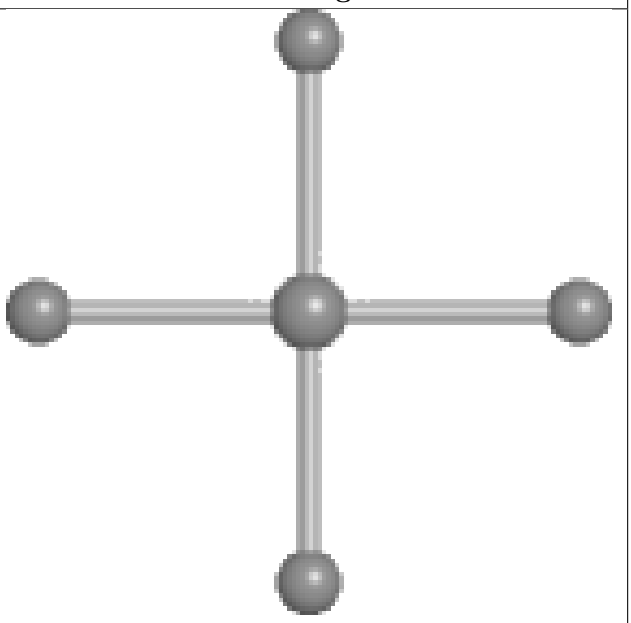
Bond lengths



Bond angles

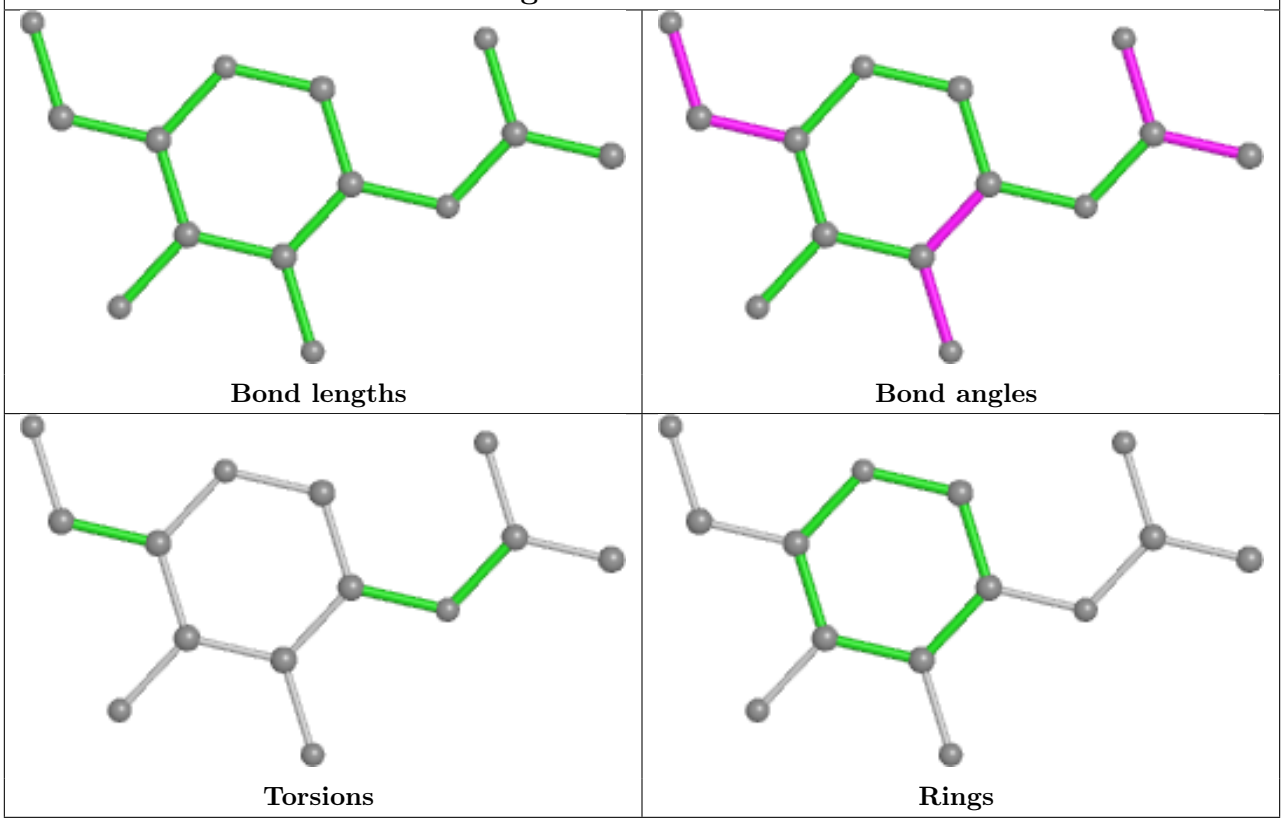


Torsions

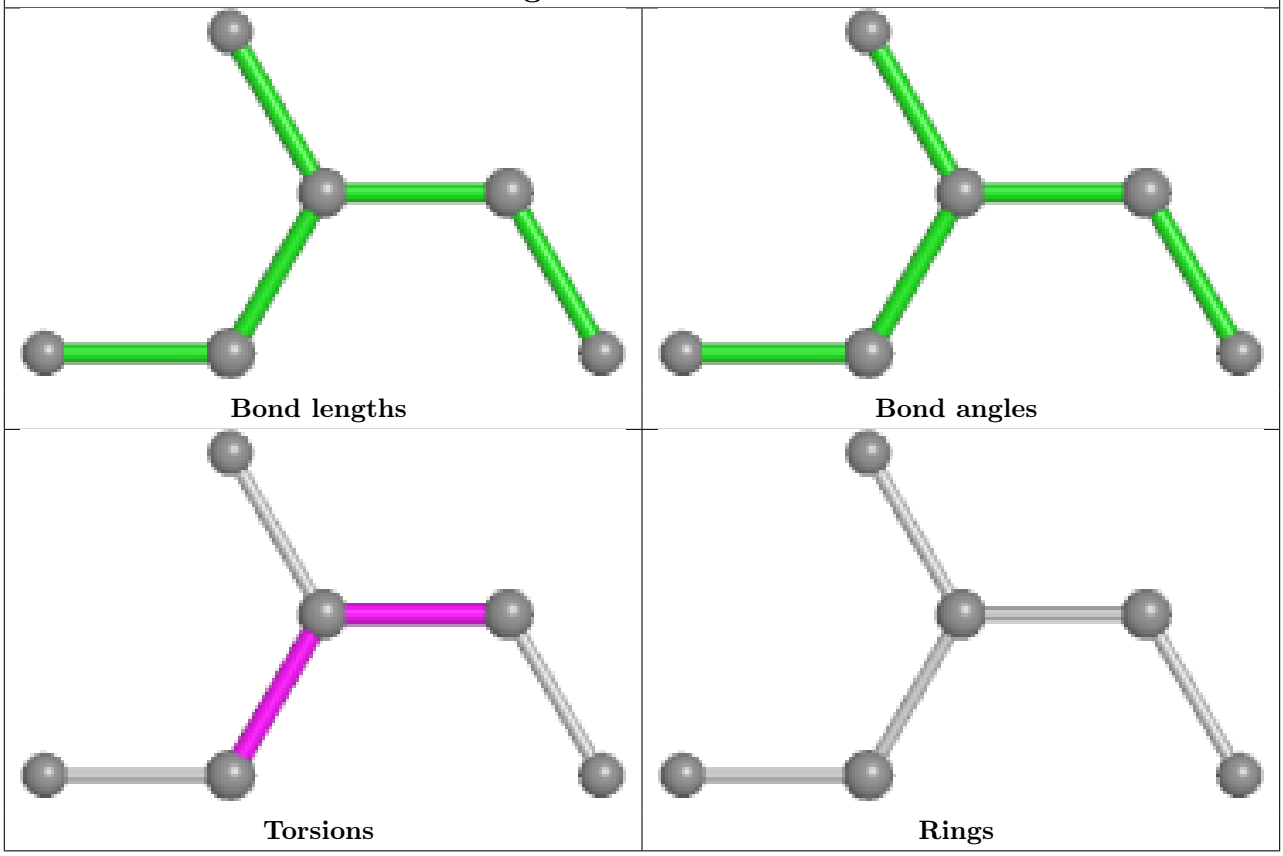


Rings

Ligand NAG A 701



Ligand GOL A 705



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, 95th 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	606/609 (99%)	-0.35	3 (0%) 91 93	16, 24, 39, 86	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	-2	HIS	3.6
1	A	586	HIS	3.3
1	A	-1	ALA	2.7

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

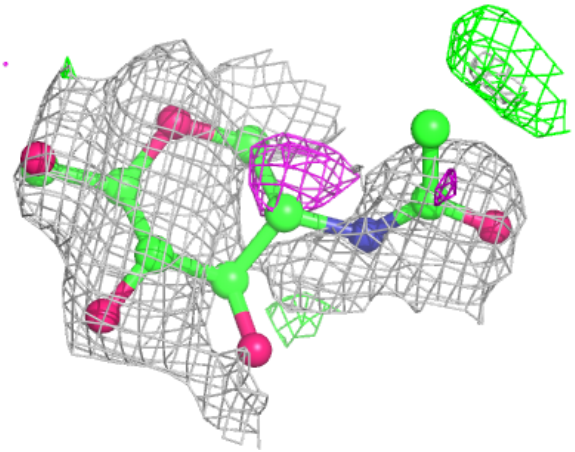
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

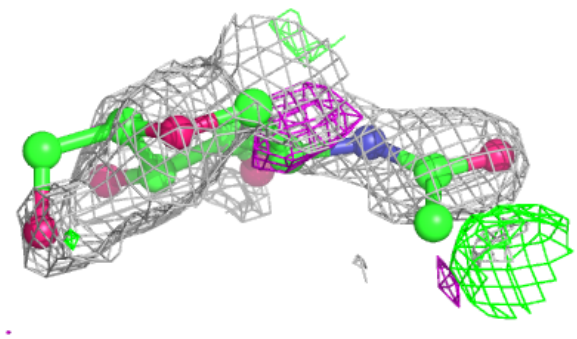
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	A	702	14/15	0.66	0.27	63,77,86,87	0
3	NAG	A	703	14/15	0.77	0.38	56,77,96,109	0
3	NAG	A	701	14/15	0.91	0.21	44,52,63,65	0

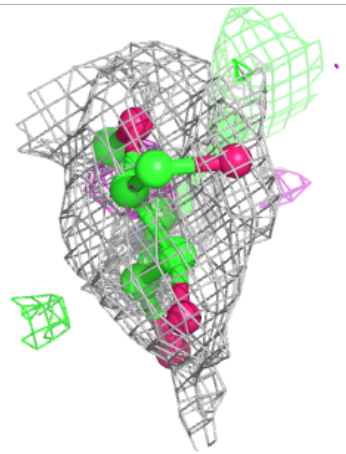
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around NAG A 702:

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

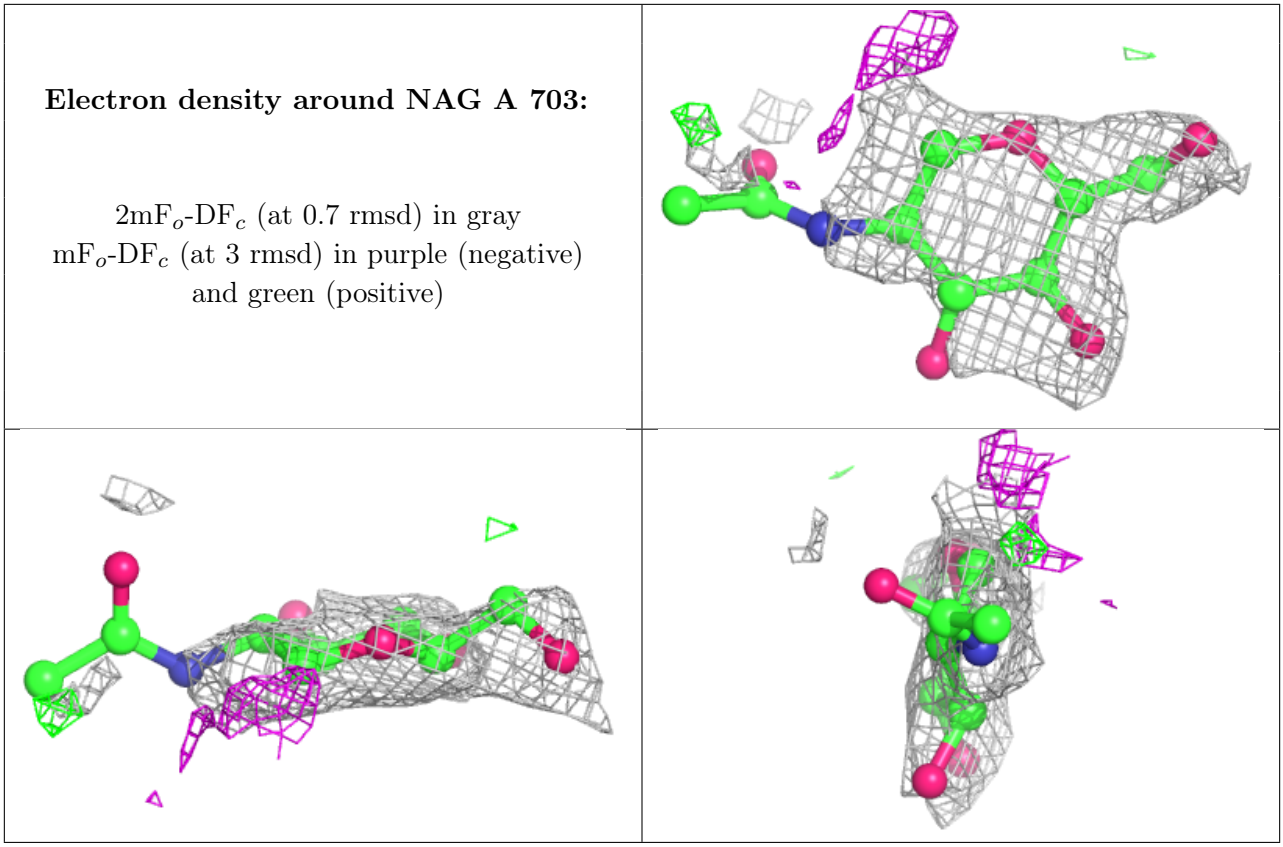






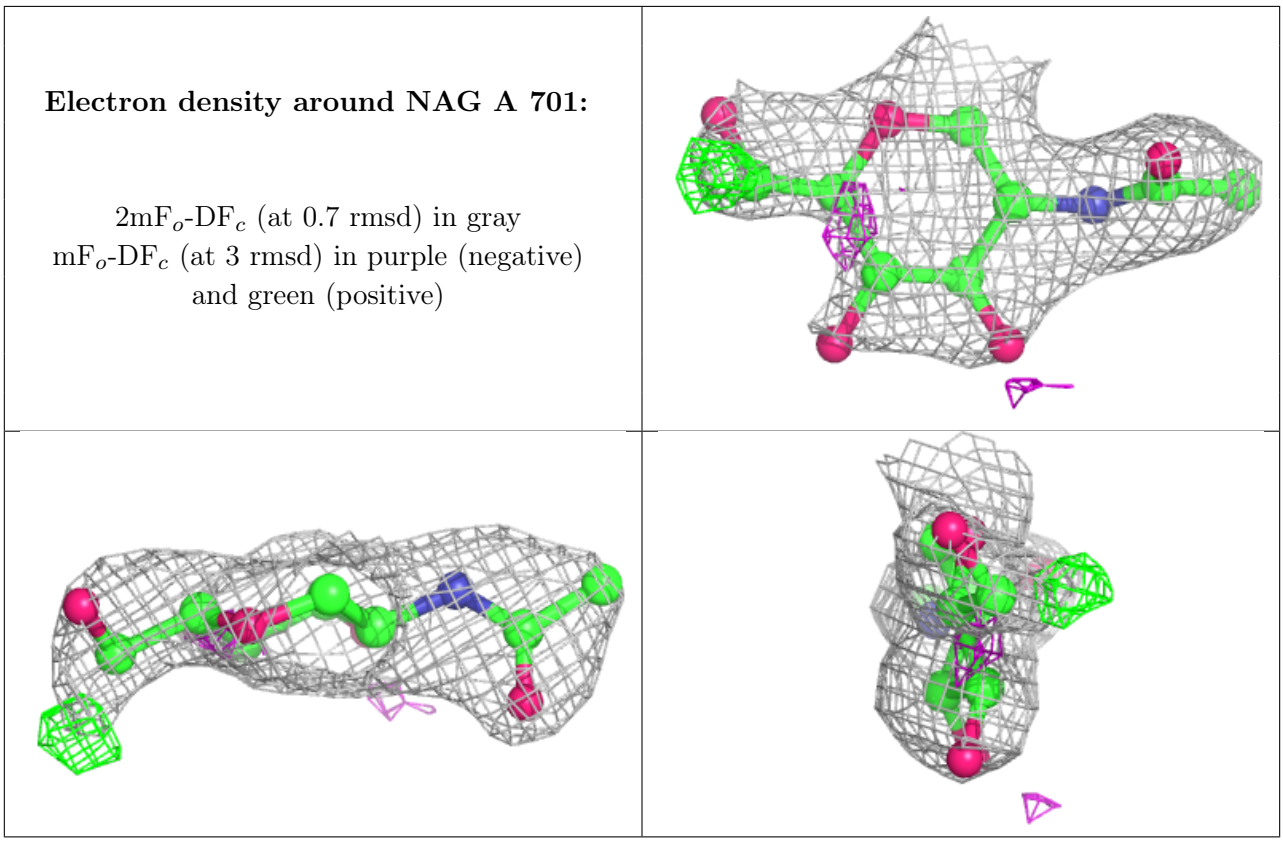
Electron density around NAG A 703:

$2mF_o$-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 A 701:

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



6.3 Carbohydrates (i)

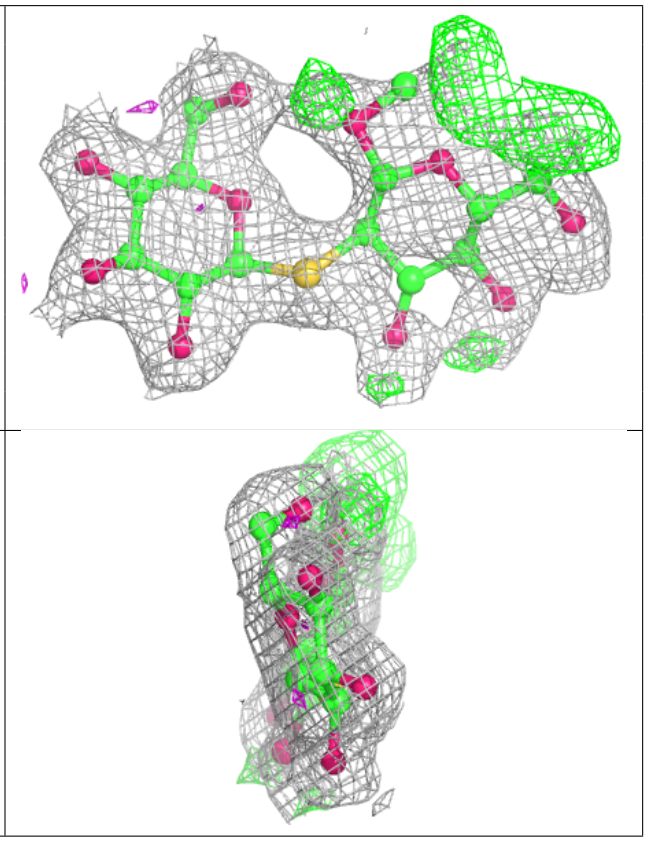
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

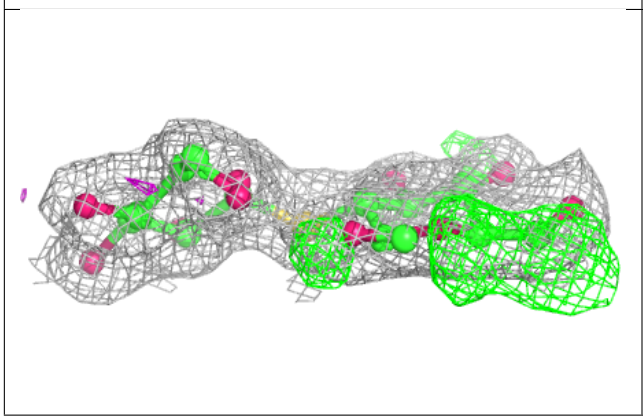
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	U2A	B	1	13/13	0.72	0.18	37,48,55,55	12
2	BGC	B	2	11/12	0.93	0.12	24,28,32,38	0

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 B:

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, 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	NAG	A	702	14/15	0.66	0.27	63,77,86,87	0

Continued on next page...

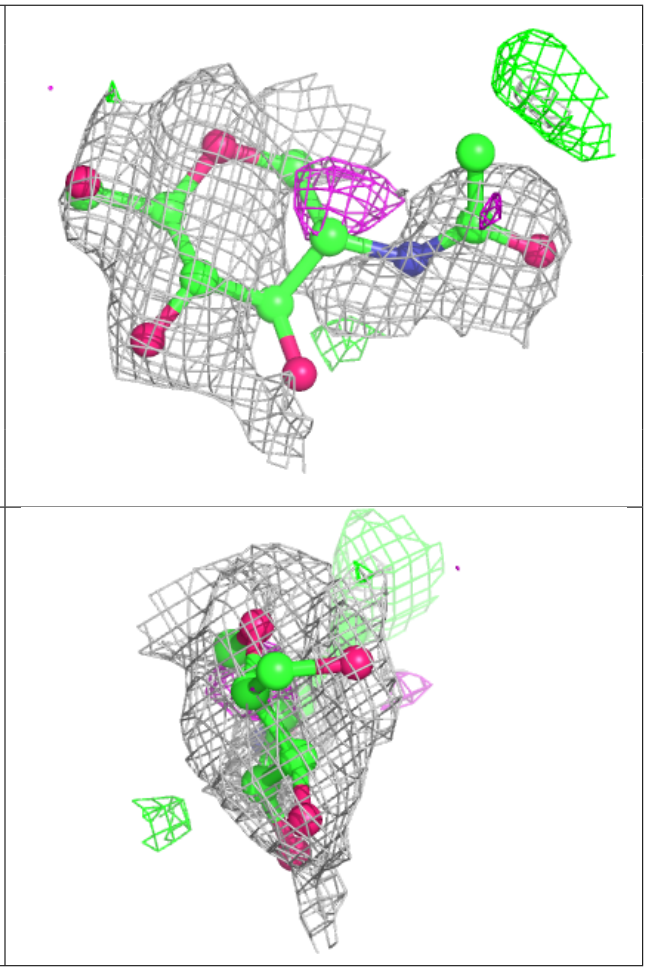
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	1PE	A	709	16/16	0.72	0.24	56,67,73,77	0
3	NAG	A	703	14/15	0.77	0.38	56,77,96,109	0
4	GOL	A	706	6/6	0.85	0.18	35,48,52,54	0
4	GOL	A	705	6/6	0.88	0.12	49,55,57,63	0
5	SO4	A	708	5/5	0.89	0.15	59,60,69,85	0
3	NAG	A	701	14/15	0.91	0.21	44,52,63,65	0
5	SO4	A	707	5/5	0.94	0.17	65,65,77,84	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.

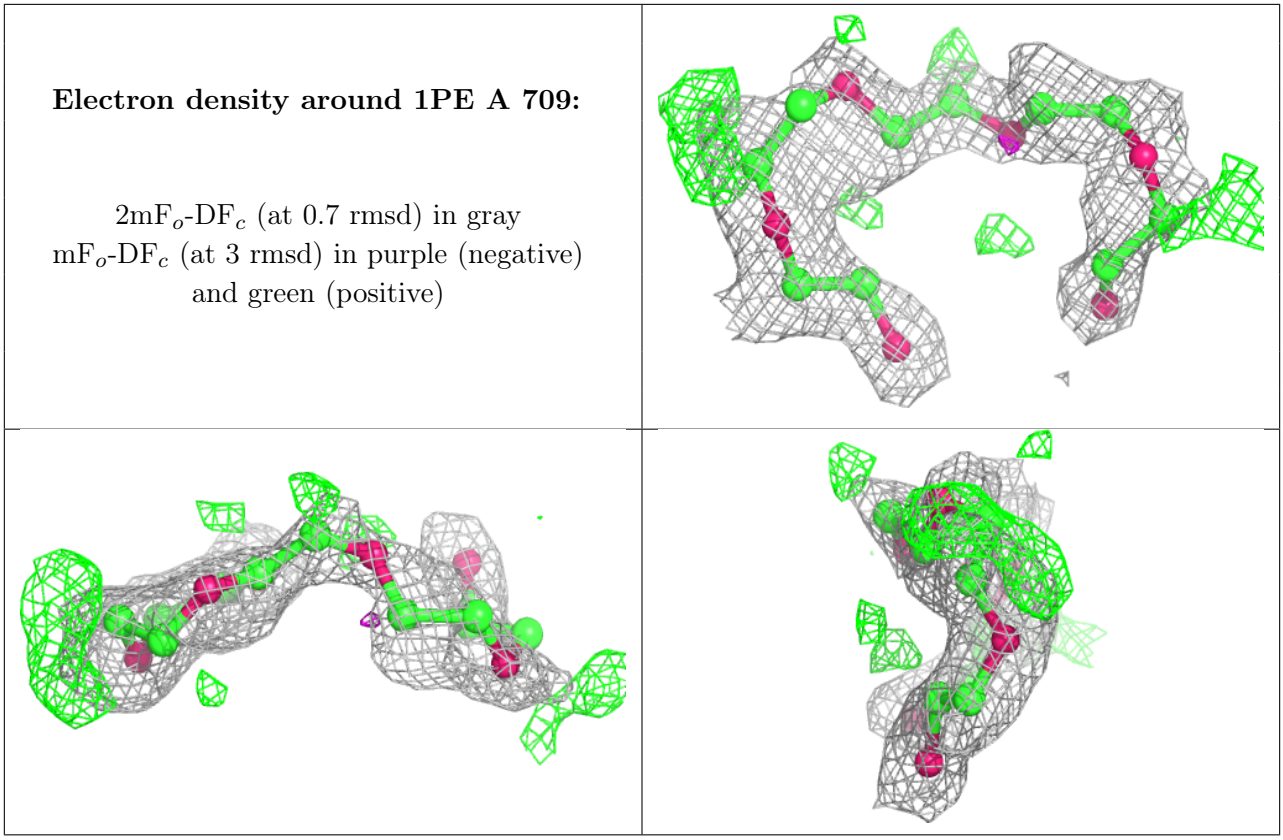
Electron density around NAG A 702:

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



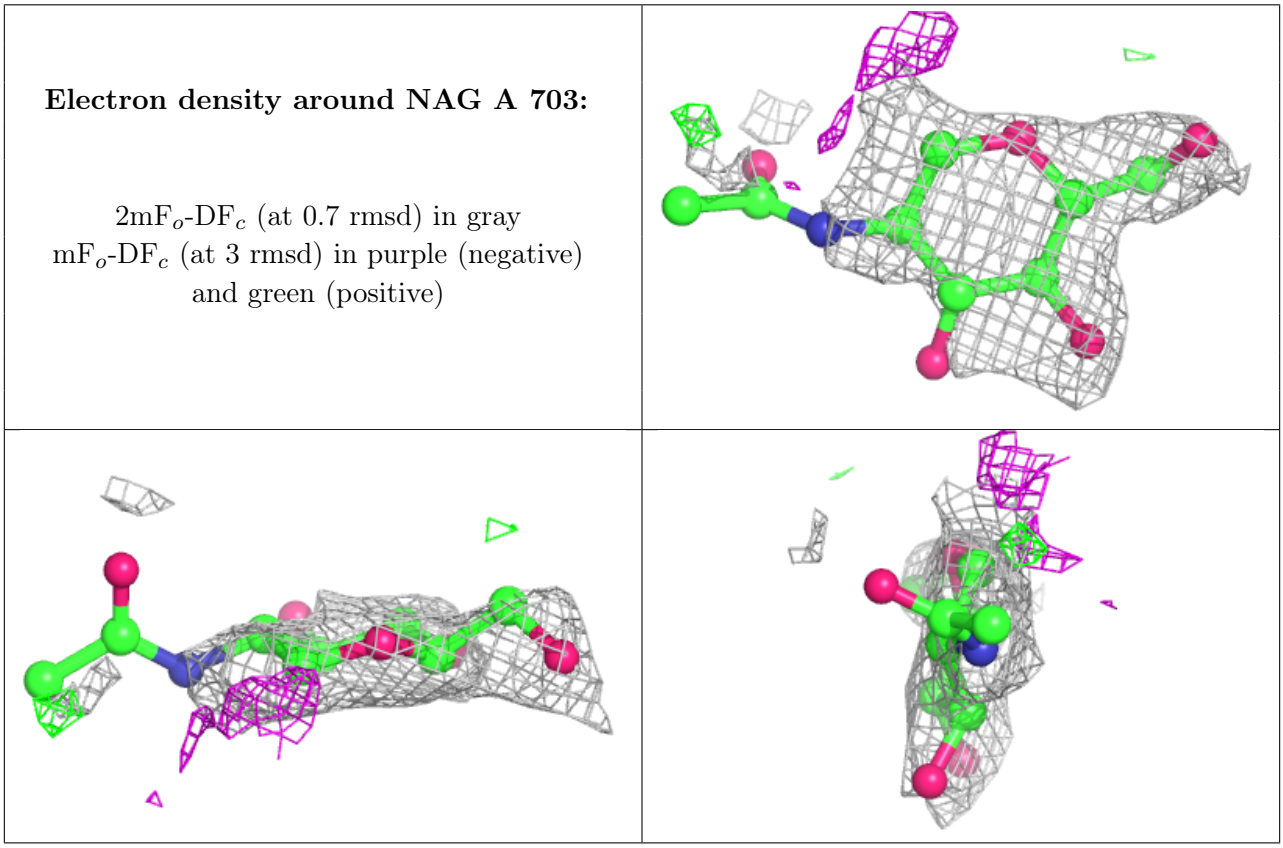
Electron density around 1PE A 709:

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



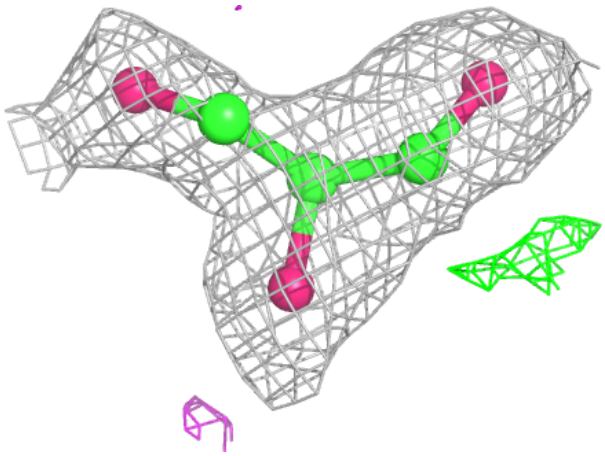
Electron density around NAG A 703:

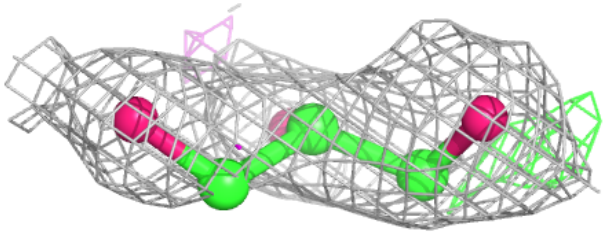
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

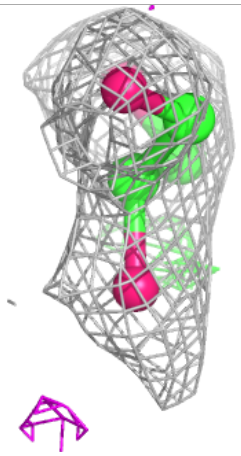


Electron density around GOL A 706:

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

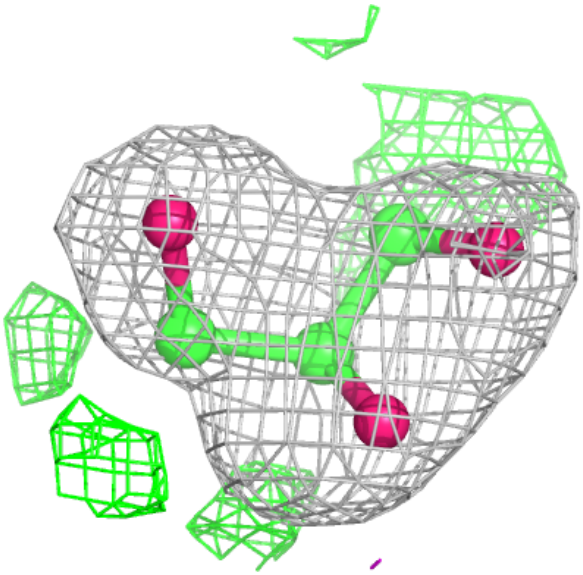


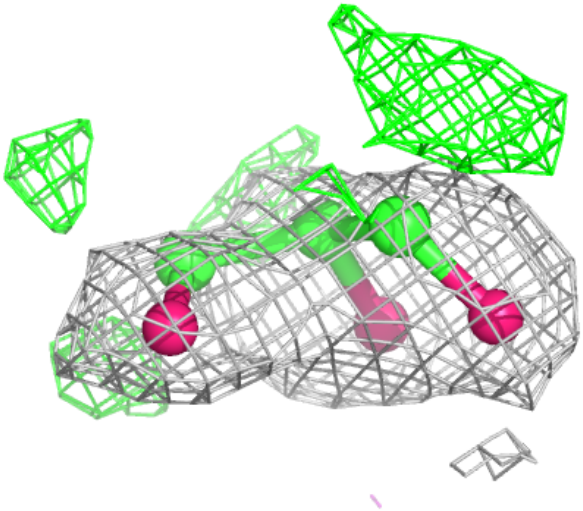


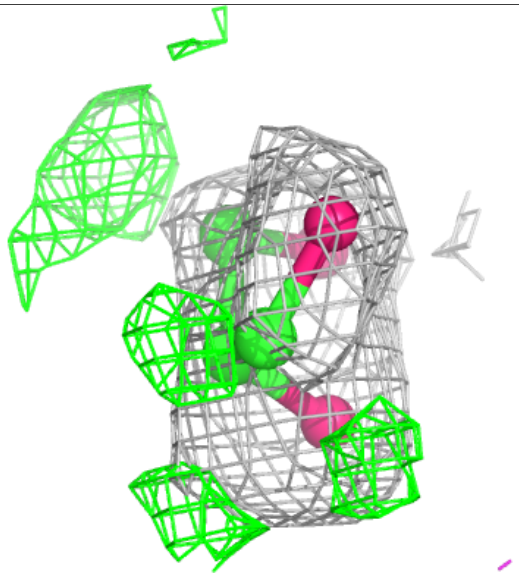


Electron density around GOL A 705:

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

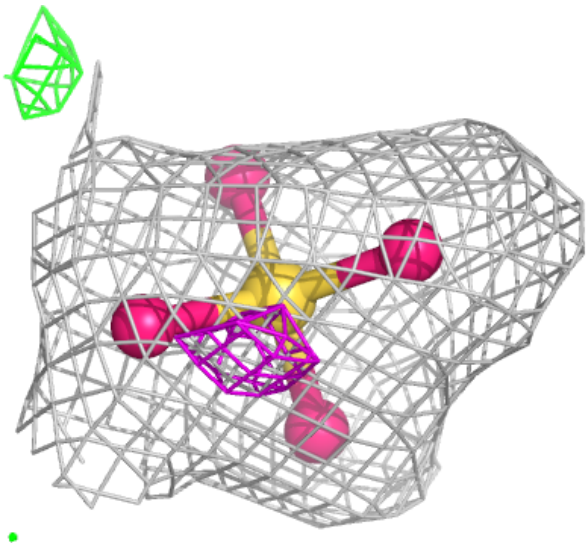


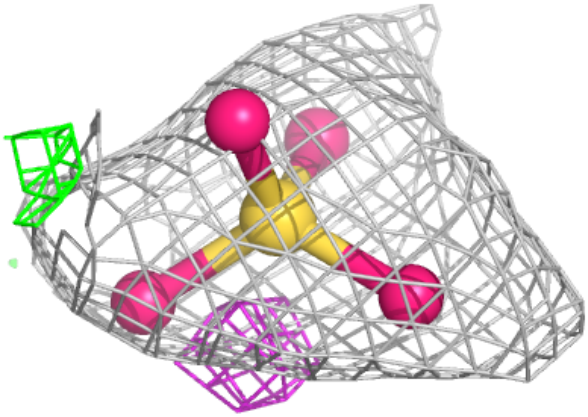


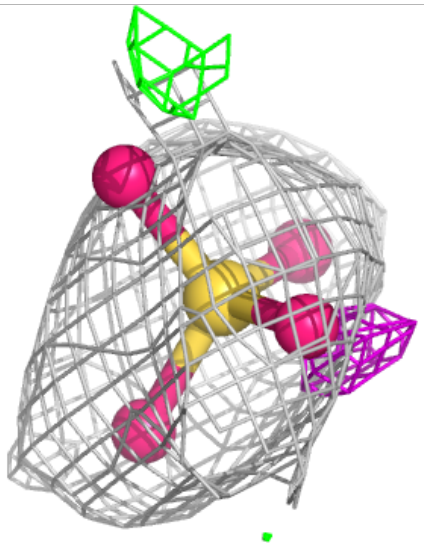


Electron density around SO4 A 708:

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

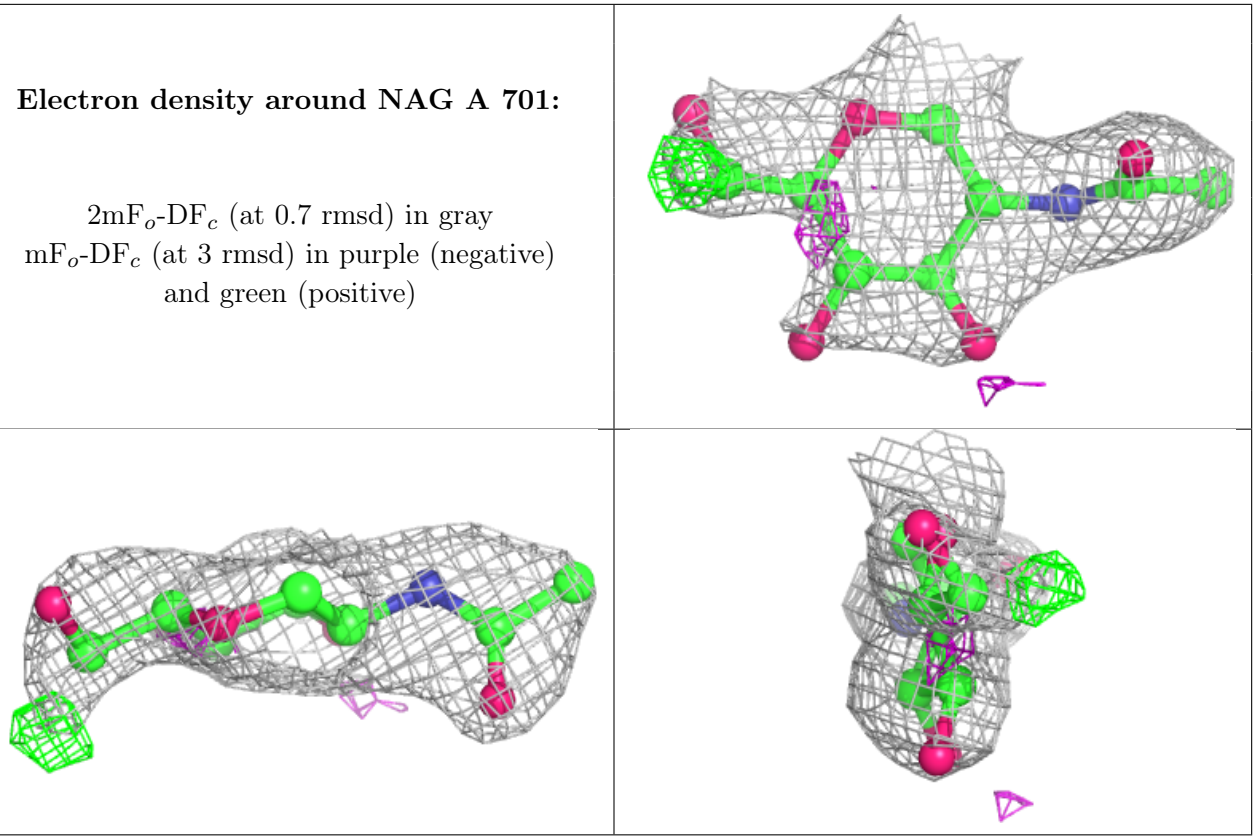






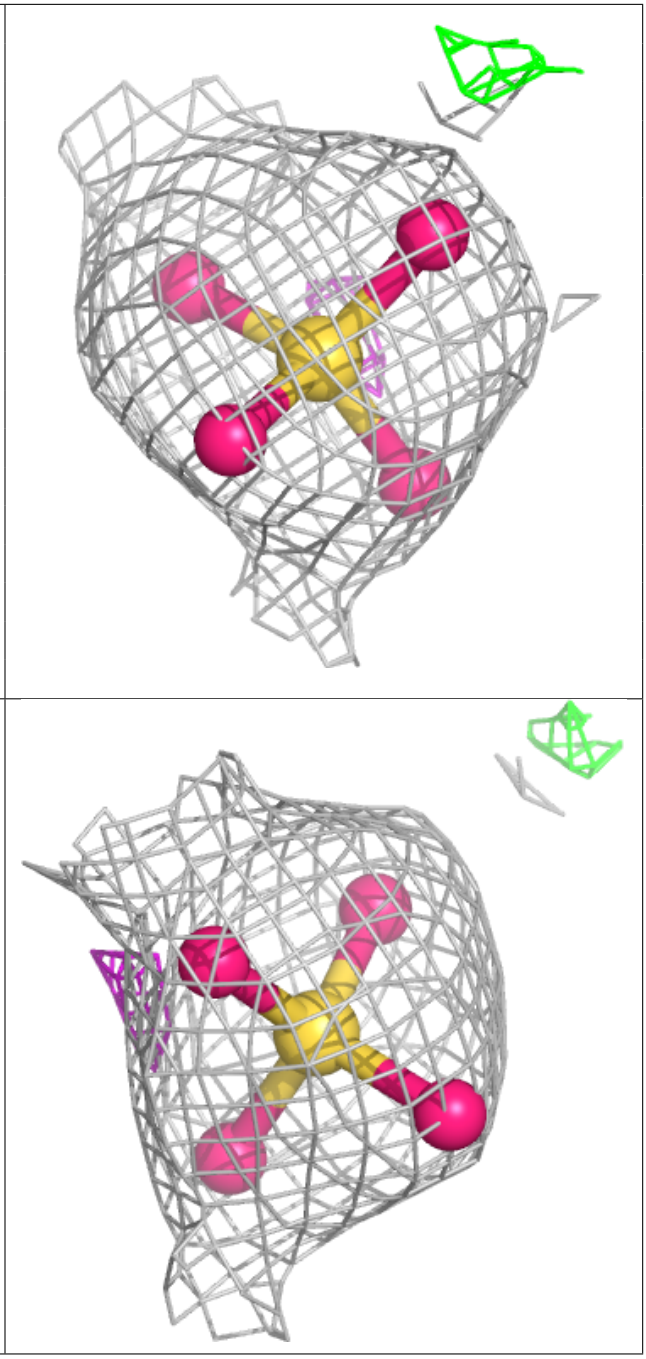
Electron density around NAG A 701:

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



Electron density around SO4 A 707:

$2mF_o$-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 (i)

There are no such residues in this entry.