Full wwPDB X-ray Structure Validation Report ⓘ

May 7, 2025 – 12:21 AM JST

PDB ID	:	8X2C / pdb_00008x2c
Title	:	Crystal structure of H5 hemagglutinin from swan-infecting H5N8 influenza virus
Authors	:	Jin, X.Y.; Han, P.; Song, H.; Qi, J.X.
Deposited on	:	2023-11-09
Resolution	:	2.59 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 ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

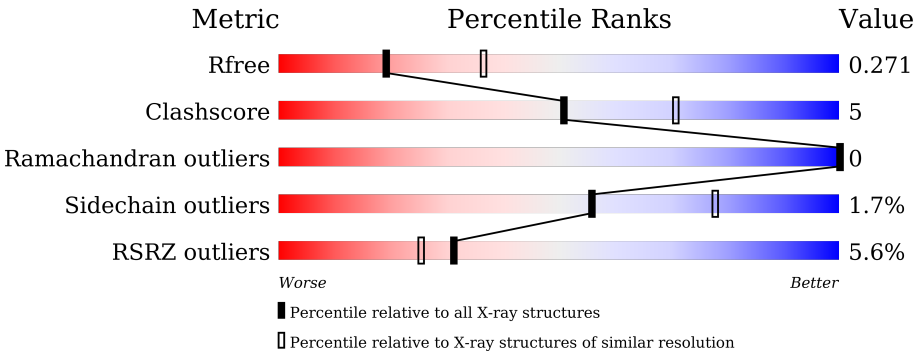
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

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 2.59 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}	164625	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (2.60-2.60)

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	506	3% 85% 12% .
1	B	506	3% 84% 13% .
1	C	506	11% 81% 15% .
2	D	2	100%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 12022 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
1	A	492	Total	C	N	O	S	0	0	0
			3930	2475	682	751	22			
1	B	492	Total	C	N	O	S	0	0	0
			3930	2475	682	751	22			
1	C	489	Total	C	N	O	S	0	0	0
			3906	2461	676	748	21			

There are 9 discrepancies between the modelled and reference sequences:

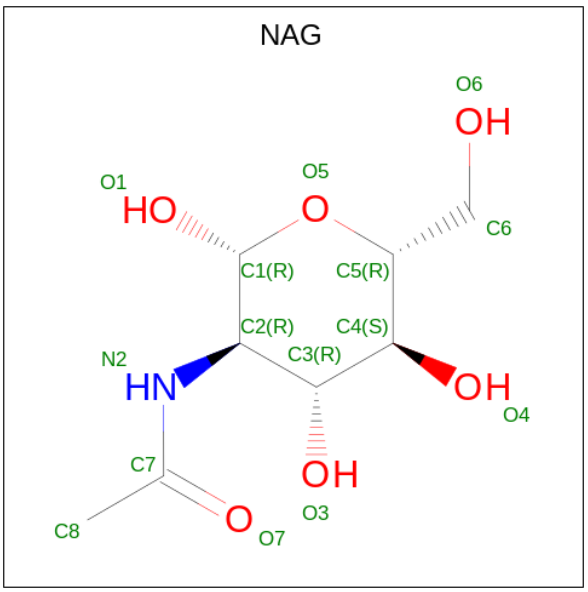
Chain	Residue	Modelled	Actual	Comment	Reference
A	86	VAL	ALA	engineered mutation	UNP A0A8E4ZAK5
A	188	ILE	THR	engineered mutation	UNP A0A8E4ZAK5
A	273	ASN	HIS	engineered mutation	UNP A0A8E4ZAK5
B	86	VAL	ALA	engineered mutation	UNP A0A8E4ZAK5
B	188	ILE	THR	engineered mutation	UNP A0A8E4ZAK5
B	273	ASN	HIS	engineered mutation	UNP A0A8E4ZAK5
C	86	VAL	ALA	engineered mutation	UNP A0A8E4ZAK5
C	188	ILE	THR	engineered mutation	UNP A0A8E4ZAK5
C	273	ASN	HIS	engineered mutation	UNP A0A8E4ZAK5

- Molecule 2 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
2	D	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (CCD ID: NAG) (formula:

$C_8H_{15}NO_6$) (labeled as "Ligand of Interest" by depositor).



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

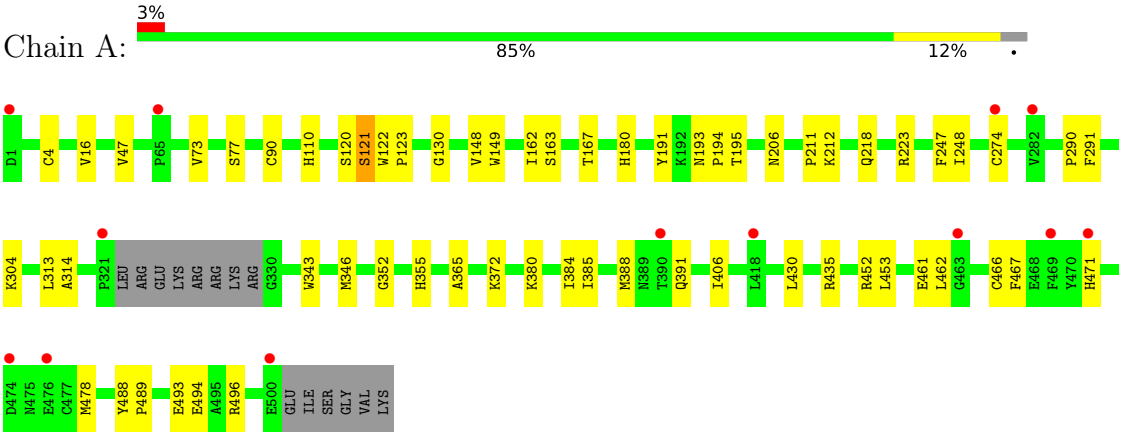
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	42	Total O 42 42	0	0
4	B	47	Total O 47 47	0	0
4	C	27	Total O 27 27	0	0

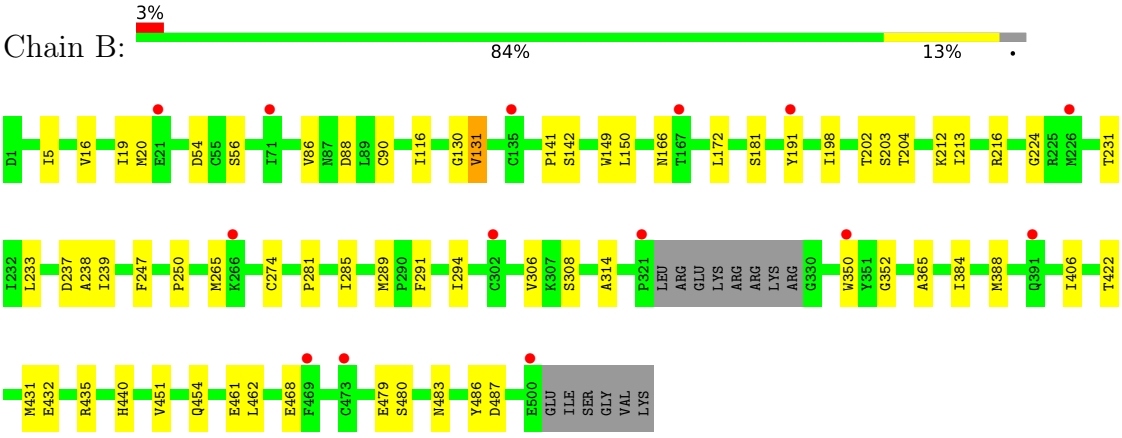
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

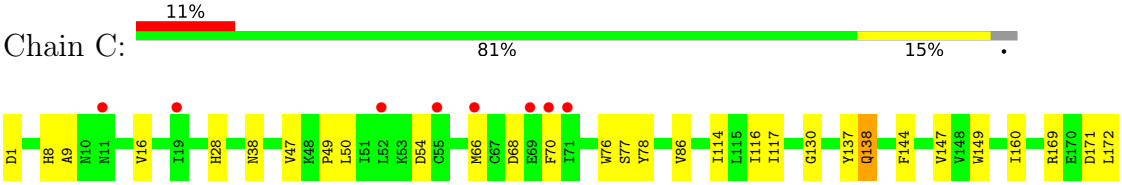
• Molecule 1: Hemagglutinin

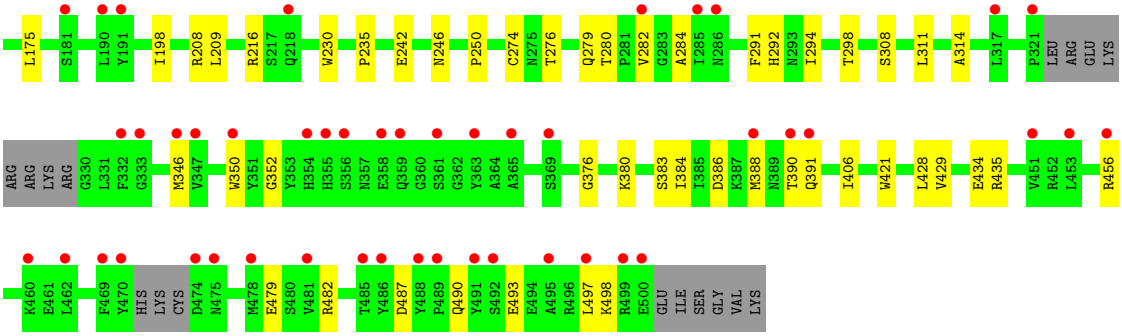


• Molecule 1: Hemagglutinin



• Molecule 1: Hemagglutinin





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

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	81.41Å 107.03Å 134.76Å 90.00° 101.05° 90.00°	Depositor
Resolution (Å)	28.27 – 2.59 28.27 – 2.59	Depositor EDS
% Data completeness (in resolution range)	98.8 (28.27-2.59) 87.7 (28.27-2.59)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	-0.01 (at 2.58Å)	Xtriage
Refinement program	PHENIX 1.19rc3_4028	Depositor
R, R_{free}	0.231 , 0.270 0.231 , 0.271	Depositor DCC
R_{free} test set	67913 reflections (2.85%)	wwPDB-VP
Wilson B-factor (Å^2)	41.2	Xtriage
Anisotropy	0.479	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 51.9	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	0.93	EDS
Total number of atoms	12022	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.46% 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: 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	0.09	0/4020	0.26	0/5444
1	B	0.10	0/4020	0.26	0/5444
1	C	0.15	0/3994	0.33	0/5409
All	All	0.12	0/12034	0.29	0/16297

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	3930	0	3791	38	0
1	B	3930	0	3790	39	0
1	C	3906	0	3772	52	0
2	D	28	0	24	2	0
3	A	42	0	39	0	0
3	B	42	0	39	0	0
3	C	28	0	26	0	0
4	A	42	0	0	0	0
4	B	47	0	0	0	0
4	C	27	0	0	0	0
All	All	12022	0	11481	118	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 5.

All (118) 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:435:ARG:NH1	1:C:434:GLU:OE2	2.12	0.82
1:B:19:ILE:HD11	1:B:431:MET:HA	1.67	0.77
1:C:479:GLU:HA	1:C:482:ARG:HE	1.57	0.69
1:C:346:MET:HE1	1:C:352:GLY:HA3	1.76	0.68
1:B:216:ARG:HH21	1:B:224:GLY:HA2	1.59	0.66
1:B:19:ILE:HD12	1:B:19:ILE:H	1.64	0.63
1:A:435:ARG:HE	1:B:435:ARG:HH12	1.48	0.61
1:A:291:PHE:HZ	1:A:388:MET:HG3	1.65	0.61
1:B:212:LYS:HD3	1:C:208:ARG:HB2	1.82	0.61
1:A:435:ARG:HE	1:B:435:ARG:NH1	2.01	0.59
1:A:471:HIS:H	1:A:471:HIS:CD2	2.22	0.57
1:B:352:GLY:HA3	1:B:365:ALA:HA	1.86	0.57
1:A:352:GLY:HA3	1:A:365:ALA:HA	1.87	0.56
1:A:218:GLN:HG3	1:A:223:ARG:HG3	1.88	0.56
1:A:355:HIS:HB2	1:A:478:MET:HE3	1.88	0.56
1:B:131:VAL:HG13	1:B:141:PRO:HB2	1.88	0.56
1:C:493:GLU:H	1:C:493:GLU:CD	2.14	0.55
1:C:280:THR:HG22	1:C:298:THR:HG22	1.87	0.55
1:A:380:LYS:O	1:A:384:ILE:HG13	2.08	0.54
1:B:285:ILE:HD11	1:B:294:ILE:HG13	1.90	0.54
1:A:452:ARG:HH21	1:A:453:LEU:HD21	1.73	0.53
1:B:265:MET:HE3	1:B:281:PRO:HA	1.90	0.53
1:C:47:VAL:HB	1:C:77:SER:HB3	1.90	0.53
1:C:50:LEU:HD12	1:C:76:TRP:CE3	2.44	0.53
1:C:346:MET:HE1	1:C:352:GLY:CA	2.39	0.52
1:B:130:GLY:HA3	1:B:149:TRP:HB3	1.92	0.52
1:B:16:VAL:HG21	1:B:314:ALA:HB2	1.90	0.52
1:A:313:LEU:HD11	1:A:384:ILE:HD12	1.92	0.52
1:B:462:LEU:HD21	1:B:468:GLU:HB2	1.90	0.52
1:B:203:SER:OG	1:B:237:ASP:OD1	2.25	0.51
1:B:131:VAL:HG22	1:B:142:SER:HA	1.92	0.51
1:B:350:TRP:CZ3	1:B:440:HIS:HE1	2.28	0.51
1:A:180:HIS:CD2	1:A:212:LYS:H	2.29	0.50
1:B:454:GLN:HG2	1:B:486:TYR:HB3	1.93	0.50
1:B:5:ILE:HD11	1:B:451:VAL:HG21	1.93	0.49
1:A:16:VAL:HG21	1:A:314:ALA:HB2	1.94	0.49
1:A:304:LYS:HD2	1:A:391:GLN:HB2	1.94	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:202:THR:HG22	1:B:239:ILE:HA	1.95	0.49
1:A:247:PHE:C	1:A:248:ILE:HD13	2.37	0.49
1:C:130:GLY:HA3	1:C:149:TRP:HB3	1.95	0.49
1:C:386:ASP:C	1:C:388:MET:H	2.19	0.49
1:C:291:PHE:HZ	1:C:388:MET:HG3	1.77	0.49
2:D:1:NAG:H61	2:D:2:NAG:O5	2.13	0.49
1:C:311:LEU:HD22	1:C:429:VAL:HG21	1.94	0.49
1:A:121:SER:O	1:A:123:PRO:HD3	2.13	0.48
1:B:150:LEU:O	1:B:191:TYR:HE1	1.97	0.48
1:B:56:SER:HB3	1:B:88:ASP:HB2	1.96	0.48
1:B:480:SER:HA	1:B:483:ASN:HB2	1.95	0.48
1:C:114:ILE:HD11	1:C:172:LEU:HD21	1.95	0.48
1:C:171:ASP:OD2	1:C:235:PRO:HD3	2.14	0.48
1:A:406:ILE:HG21	1:C:406:ILE:HD11	1.95	0.47
1:C:54:ASP:HB2	1:C:86:VAL:HG22	1.96	0.47
1:B:265:MET:HG2	1:B:281:PRO:HG3	1.97	0.47
1:A:122:TRP:HH2	1:A:162:ILE:HG21	1.79	0.47
1:C:54:ASP:OD2	1:C:54:ASP:N	2.45	0.47
1:C:292:HIS:HD2	1:C:294:ILE:H	1.63	0.47
1:B:20:MET:HE3	1:C:380:LYS:HE3	1.97	0.47
1:A:343:TRP:HE3	1:A:346:MET:HE2	1.80	0.47
1:C:384:ILE:HD12	1:C:428:LEU:HD21	1.97	0.46
1:B:198:ILE:HD11	1:B:247:PHE:HA	1.98	0.46
1:C:38:ASN:ND2	1:C:284:ALA:HB3	2.30	0.46
1:B:384:ILE:HD11	1:B:432:GLU:HG3	1.96	0.46
1:C:198:ILE:HB	1:C:209:LEU:HB2	1.96	0.46
1:C:390:THR:O	1:C:390:THR:HG23	2.15	0.46
1:C:175:LEU:HD23	1:C:230:TRP:HB3	1.97	0.46
1:A:496:ARG:HG2	1:A:496:ARG:HH11	1.81	0.46
1:A:493:GLU:CD	1:A:493:GLU:H	2.23	0.46
1:C:117:ILE:HD12	1:C:250:PRO:HG2	1.98	0.45
1:C:487:ASP:HB3	1:C:490:GLN:HB2	1.98	0.45
1:A:206:ASN:OD1	1:C:216:ARG:HD3	2.17	0.45
1:C:391:GLN:HG2	1:C:421:TRP:CD2	2.52	0.45
1:A:461:GLU:HG2	1:A:467:PHE:CE1	2.52	0.45
1:B:291:PHE:HZ	1:B:388:MET:HG3	1.81	0.45
1:C:291:PHE:CZ	1:C:388:MET:HG3	2.52	0.45
1:C:311:LEU:HD23	1:C:311:LEU:HA	1.82	0.45
1:A:191:TYR:O	1:A:193:ASN:N	2.49	0.44
1:C:8:HIS:ND1	1:C:9:ALA:N	2.65	0.44
1:A:435:ARG:HH21	1:C:435:ARG:HH11	1.66	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:54:ASP:HB3	1:B:86:VAL:HG22	2.00	0.44
1:C:28:HIS:CD2	1:C:350:TRP:HE1	2.34	0.44
1:C:380:LYS:O	1:C:384:ILE:HG12	2.17	0.44
1:A:130:GLY:HA3	1:A:149:TRP:HB3	2.00	0.44
1:B:289:MET:HE3	1:B:289:MET:HB3	1.73	0.44
1:C:160:ILE:O	1:C:242:GLU:HA	2.18	0.43
1:B:20:MET:HE2	1:C:376:GLY:C	2.42	0.43
1:C:198:ILE:HD11	1:C:246:ASN:O	2.18	0.43
1:A:435:ARG:HH21	1:C:435:ARG:NH1	2.16	0.43
1:B:238:ALA:N	2:D:1:NAG:H82	2.34	0.43
1:A:4:CYS:HA	1:A:466:CYS:HA	2.01	0.43
1:C:68:ASP:HB3	1:C:70:PHE:CE1	2.54	0.43
1:A:372:LYS:HE3	1:A:372:LYS:HB2	1.85	0.42
1:B:172:LEU:HD12	1:B:172:LEU:HA	1.89	0.42
1:C:498:LYS:HA	1:C:498:LYS:HD3	1.74	0.42
1:A:167:THR:O	1:A:167:THR:OG1	2.34	0.42
1:C:406:ILE:HD13	1:C:406:ILE:HA	1.83	0.42
1:B:406:ILE:HD11	1:C:406:ILE:HG21	2.01	0.42
1:C:169:ARG:HA	1:C:235:PRO:HB3	2.01	0.42
1:B:461:GLU:O	1:C:456:ARG:NH1	2.53	0.42
1:B:479:GLU:O	1:B:483:ASN:N	2.45	0.42
1:A:162:ILE:HG13	1:A:163:SER:N	2.35	0.41
1:C:16:VAL:HG21	1:C:314:ALA:HB2	2.02	0.41
1:B:116:ILE:HD11	1:B:250:PRO:HB2	2.01	0.41
1:B:166:ASN:HB2	1:B:233:LEU:HD23	2.01	0.41
1:A:488:TYR:N	1:A:489:PRO:HD2	2.35	0.41
1:C:66:MET:HB2	1:C:66:MET:HE3	1.78	0.41
1:C:276:THR:HG21	1:C:279:GLN:HG2	2.02	0.41
1:B:306:VAL:HG23	1:B:308:SER:H	1.84	0.41
1:A:180:HIS:HD2	1:A:211:PRO:HA	1.85	0.41
1:C:116:ILE:HD11	1:C:250:PRO:HB2	2.02	0.41
1:A:47:VAL:HB	1:A:77:SER:HB3	2.02	0.41
1:A:290:PRO:HG3	1:A:385:ILE:HD12	2.02	0.41
1:A:193:ASN:HA	1:A:194:PRO:HD3	1.87	0.40
1:B:181:SER:OG	1:B:213:ILE:HG12	2.21	0.40
1:C:137:TYR:C	1:C:138:GLN:HG3	2.45	0.40
1:A:73:VAL:O	1:A:110:HIS:HE1	2.05	0.40
1:A:430:LEU:HD12	1:A:430:LEU:HA	1.98	0.40
1:C:49:PRO:HG3	1:C:78:TYR:CZ	2.56	0.40
1:C:144:PHE:HB2	1:C:147:VAL:HG12	2.03	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	488/506 (96%)	468 (96%)	20 (4%)	0	100	100
1	B	488/506 (96%)	470 (96%)	18 (4%)	0	100	100
1	C	483/506 (96%)	465 (96%)	18 (4%)	0	100	100
All	All	1459/1518 (96%)	1403 (96%)	56 (4%)	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	433/446 (97%)	425 (98%)	8 (2%)	54	77
1	B	432/446 (97%)	425 (98%)	7 (2%)	58	79
1	C	430/446 (96%)	423 (98%)	7 (2%)	58	79
All	All	1295/1338 (97%)	1273 (98%)	22 (2%)	56	78

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	90	CYS
1	A	120	SER
1	A	121	SER
1	A	148	VAL
1	A	195	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	274	CYS
1	A	462	LEU
1	A	494	GLU
1	B	90	CYS
1	B	131	VAL
1	B	204	THR
1	B	231	THR
1	B	274	CYS
1	B	422	THR
1	B	487	ASP
1	C	1	ASP
1	C	138	GLN
1	C	274	CYS
1	C	282	VAL
1	C	308	SER
1	C	383	SER
1	C	497	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	110	HIS
1	A	168	ASN
1	A	180	HIS
1	A	193	ASN
1	A	309	ASN
1	A	354	HIS
1	A	389	ASN
1	B	45	ASN
1	B	273	ASN
1	B	371	GLN
1	C	103	HIS
1	C	240	HIS
1	C	279	GLN
1	C	354	HIS

5.3.3 RNA

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)

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	NAG	D	1	1,2	14,14,15	2.28	4 (28%)	17,19,21	1.59	4 (23%)
2	NAG	D	2	2	14,14,15	2.59	5 (35%)	17,19,21	5.82	7 (41%)

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	NAG	D	1	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	2	2	-	5/6/23/26	0/1/1/1

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	NAG	O5-C1	5.94	1.53	1.43
2	D	2	NAG	O5-C1	5.20	1.52	1.43
2	D	2	NAG	C2-N2	4.23	1.53	1.46
2	D	2	NAG	C7-N2	3.93	1.47	1.34
2	D	1	NAG	C7-N2	3.82	1.47	1.34
2	D	2	NAG	C1-C2	3.59	1.57	1.52
2	D	2	NAG	O7-C7	-3.56	1.15	1.23
2	D	1	NAG	O7-C7	-2.92	1.16	1.23
2	D	1	NAG	C2-N2	2.11	1.49	1.46

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	2	NAG	C1-C2-N2	-13.71	87.07	110.49
2	D	2	NAG	C2-N2-C7	12.02	140.02	122.90
2	D	2	NAG	C8-C7-N2	11.83	136.13	116.10
2	D	2	NAG	O7-C7-N2	-8.59	106.16	121.95
2	D	1	NAG	C4-C3-C2	3.78	116.56	111.02
2	D	2	NAG	O5-C1-C2	3.26	116.43	111.29
2	D	1	NAG	O4-C4-C3	-2.74	104.02	110.35
2	D	2	NAG	O7-C7-C8	-2.67	117.11	122.06
2	D	1	NAG	C1-O5-C5	-2.54	108.75	112.19
2	D	2	NAG	C1-O5-C5	-2.41	108.93	112.19
2	D	1	NAG	C8-C7-N2	2.37	120.11	116.10

There are no chirality outliers.

All (5) torsion outliers are listed below:

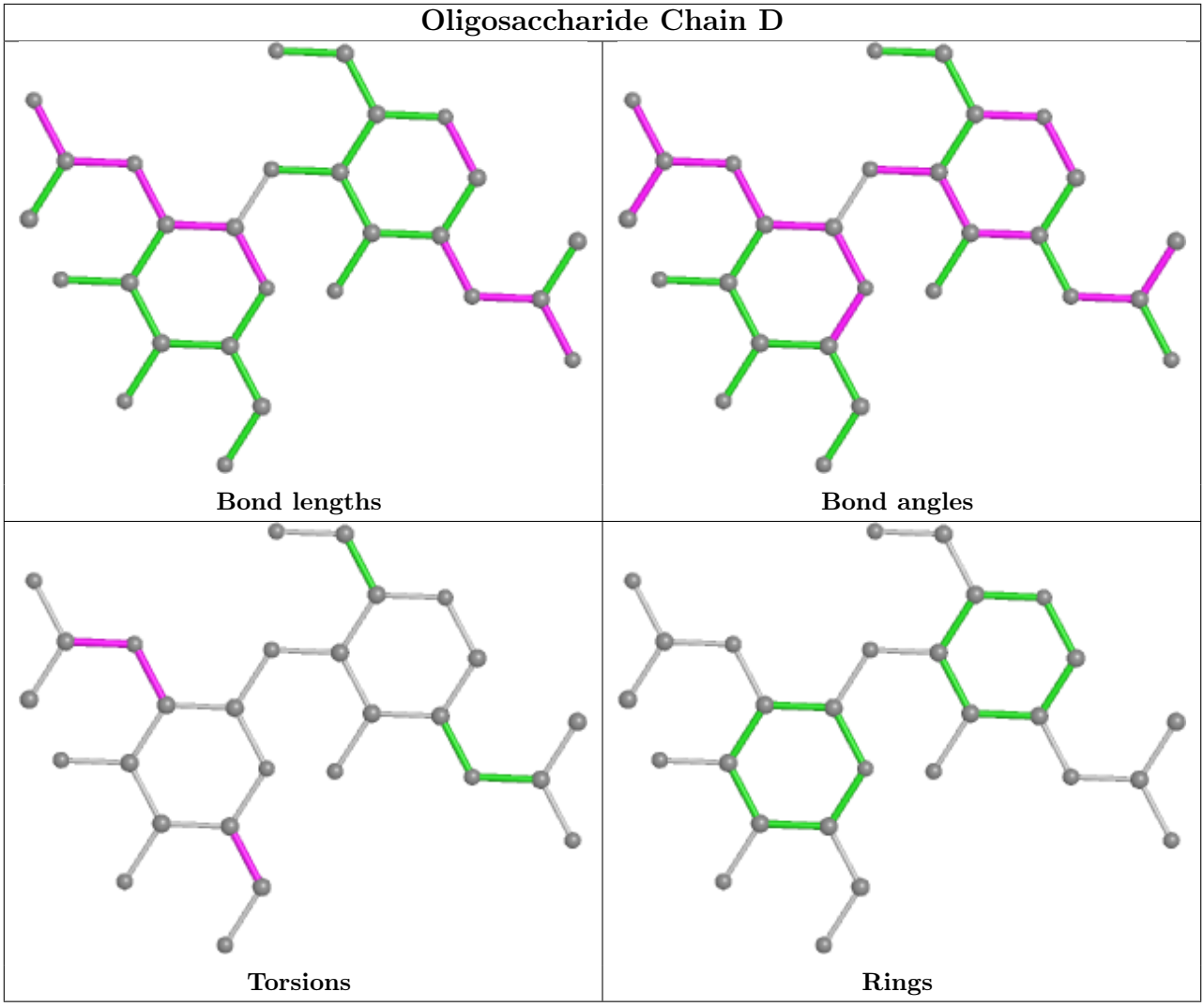
Mol	Chain	Res	Type	Atoms
2	D	2	NAG	C3-C2-N2-C7
2	D	2	NAG	C8-C7-N2-C2
2	D	2	NAG	O5-C5-C6-O6
2	D	2	NAG	O7-C7-N2-C2
2	D	2	NAG	C4-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	1	NAG	2	0
2	D	2	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 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	$\#\|Z\| > 2$
3	NAG	C	602	1	14,14,15	2.01	4 (28%)	17,19,21	1.10	2 (11%)
3	NAG	A	601	1	14,14,15	1.98	4 (28%)	17,19,21	1.15	1 (5%)
3	NAG	B	603	1	14,14,15	2.02	4 (28%)	17,19,21	1.02	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	NAG	A	603	1	14,14,15	2.02	4 (28%)	17,19,21	1.17	2 (11%)
3	NAG	B	602	1	14,14,15	1.99	4 (28%)	17,19,21	1.14	1 (5%)
3	NAG	A	602	1	14,14,15	1.99	4 (28%)	17,19,21	1.15	2 (11%)
3	NAG	C	601	1	14,14,15	2.00	4 (28%)	17,19,21	1.14	1 (5%)
3	NAG	B	601	1	14,14,15	2.02	4 (28%)	17,19,21	1.05	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
3	NAG	C	602	1	-	0/6/23/26	0/1/1/1
3	NAG	A	601	1	-	0/6/23/26	0/1/1/1
3	NAG	B	603	1	-	0/6/23/26	0/1/1/1
3	NAG	A	603	1	-	2/6/23/26	0/1/1/1
3	NAG	B	602	1	-	0/6/23/26	0/1/1/1
3	NAG	A	602	1	-	2/6/23/26	0/1/1/1
3	NAG	C	601	1	-	0/6/23/26	0/1/1/1
3	NAG	B	601	1	-	1/6/23/26	0/1/1/1

All (32) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	603	NAG	O5-C1	4.77	1.51	1.43
3	A	603	NAG	O5-C1	4.77	1.51	1.43
3	C	602	NAG	O5-C1	4.75	1.51	1.43
3	B	601	NAG	O5-C1	4.72	1.51	1.43
3	C	601	NAG	O5-C1	4.65	1.51	1.43
3	B	602	NAG	O5-C1	4.62	1.51	1.43
3	A	601	NAG	O5-C1	4.59	1.51	1.43
3	A	602	NAG	O5-C1	4.51	1.50	1.43
3	B	601	NAG	C7-N2	3.63	1.46	1.34
3	A	602	NAG	C7-N2	3.62	1.46	1.34
3	A	603	NAG	C7-N2	3.62	1.46	1.34
3	C	601	NAG	C7-N2	3.61	1.46	1.34
3	B	602	NAG	C7-N2	3.61	1.46	1.34
3	A	601	NAG	C7-N2	3.61	1.46	1.34
3	C	602	NAG	C7-N2	3.61	1.46	1.34
3	B	603	NAG	C7-N2	3.53	1.46	1.34
3	B	601	NAG	C2-N2	2.44	1.50	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	601	NAG	C2-N2	2.41	1.50	1.46
3	A	601	NAG	C2-N2	2.39	1.50	1.46
3	A	603	NAG	C2-N2	2.37	1.50	1.46
3	B	602	NAG	C2-N2	2.37	1.50	1.46
3	A	602	NAG	C2-N2	2.35	1.50	1.46
3	B	603	NAG	C2-N2	2.32	1.50	1.46
3	C	602	NAG	C2-N2	2.30	1.50	1.46
3	A	602	NAG	O5-C5	2.18	1.47	1.43
3	B	603	NAG	O5-C5	2.10	1.47	1.43
3	B	601	NAG	O5-C5	2.07	1.47	1.43
3	C	601	NAG	O5-C5	2.07	1.47	1.43
3	A	601	NAG	O5-C5	2.03	1.47	1.43
3	A	603	NAG	O5-C5	2.03	1.47	1.43
3	B	602	NAG	O5-C5	2.02	1.47	1.43
3	C	602	NAG	O5-C5	2.00	1.47	1.43

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	603	NAG	C8-C7-N2	2.32	120.03	116.10
3	A	602	NAG	C8-C7-N2	2.31	120.00	116.10
3	C	601	NAG	C8-C7-N2	2.27	119.95	116.10
3	B	601	NAG	C8-C7-N2	2.27	119.94	116.10
3	C	602	NAG	C8-C7-N2	2.24	119.90	116.10
3	B	602	NAG	C8-C7-N2	2.22	119.85	116.10
3	A	601	NAG	C8-C7-N2	2.21	119.85	116.10
3	A	603	NAG	C1-O5-C5	2.10	115.04	112.19
3	A	602	NAG	C2-N2-C7	-2.08	119.94	122.90
3	C	602	NAG	C2-N2-C7	-2.06	119.97	122.90

There are no chirality outliers.

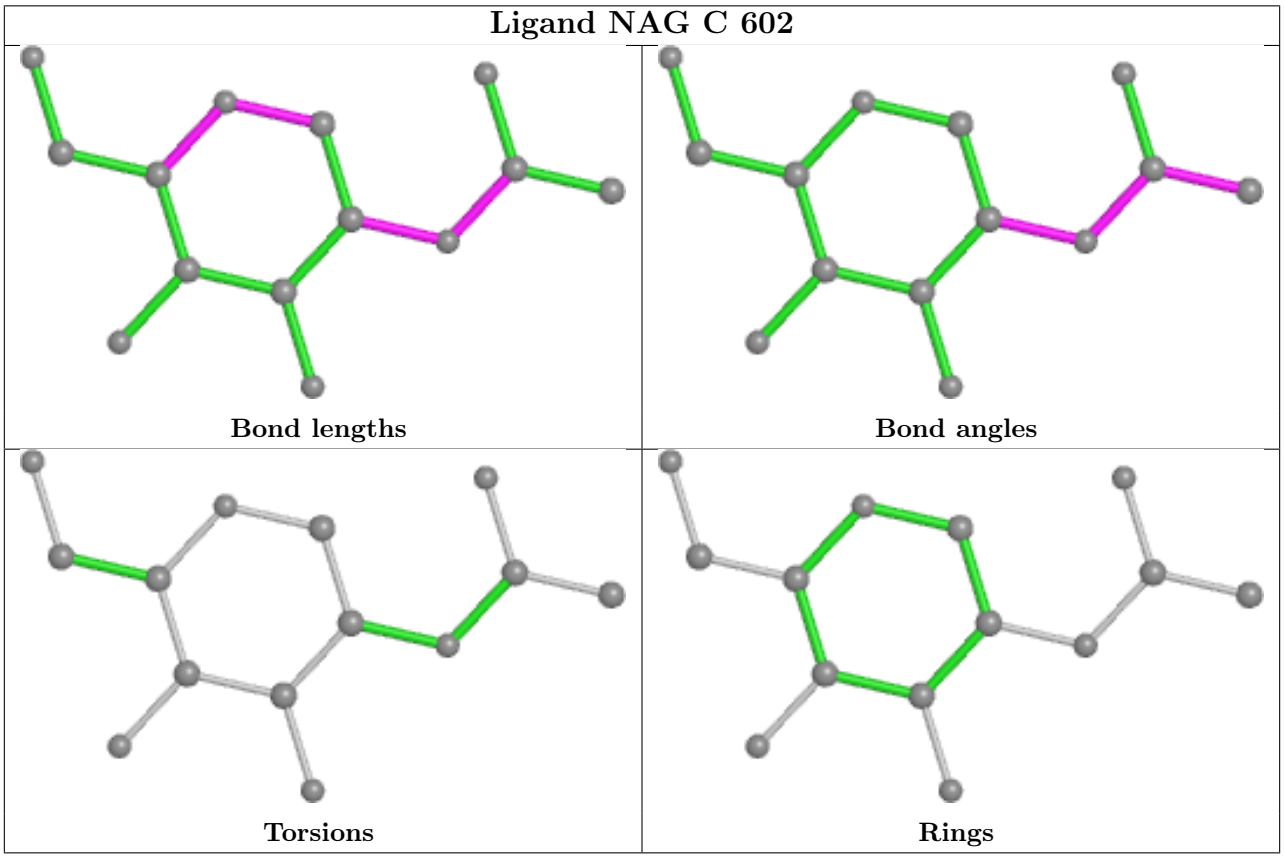
All (5) torsion outliers are listed below:

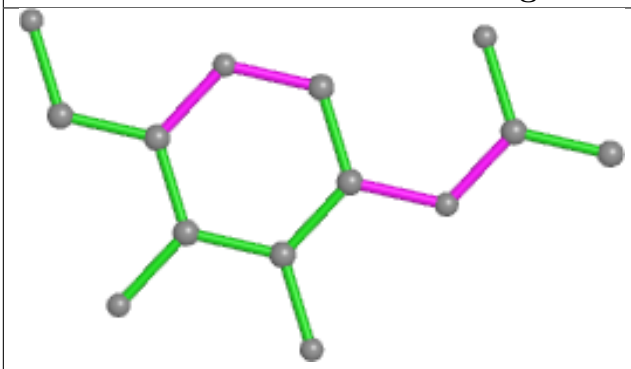
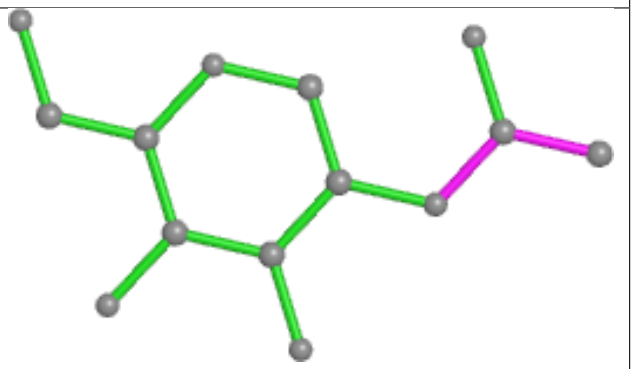
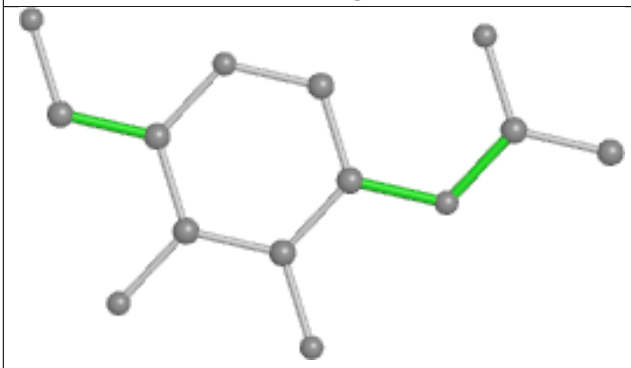
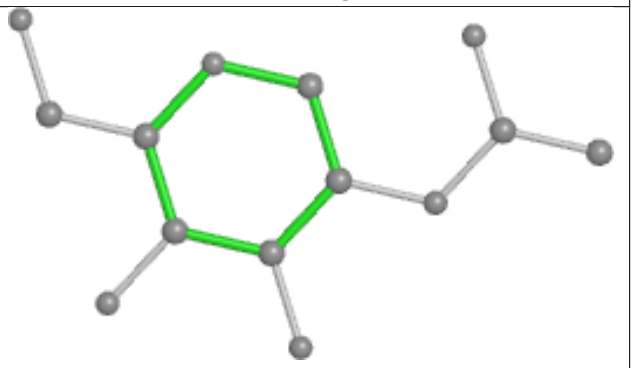
Mol	Chain	Res	Type	Atoms
3	A	602	NAG	C4-C5-C6-O6
3	A	602	NAG	O5-C5-C6-O6
3	A	603	NAG	O5-C5-C6-O6
3	A	603	NAG	C4-C5-C6-O6
3	B	601	NAG	O5-C5-C6-O6

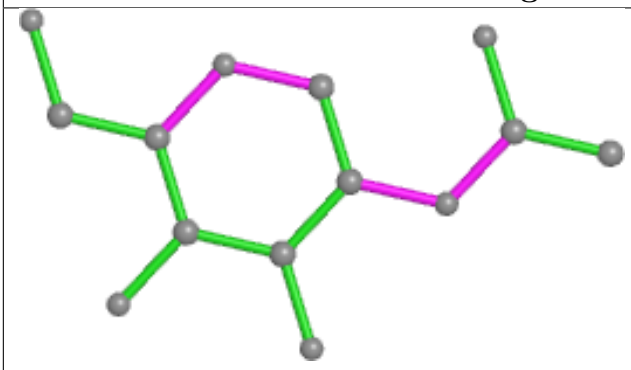
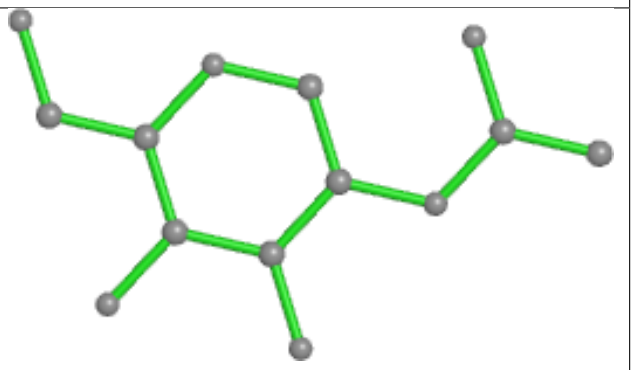
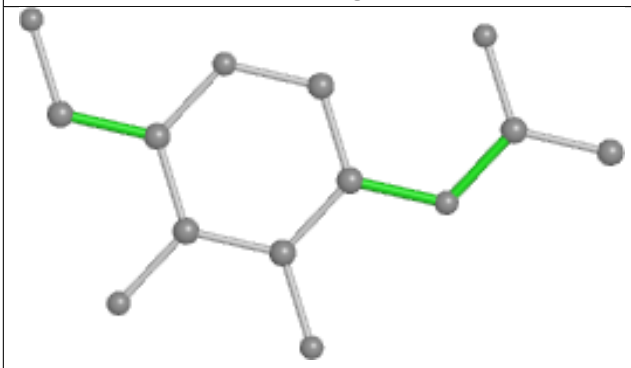
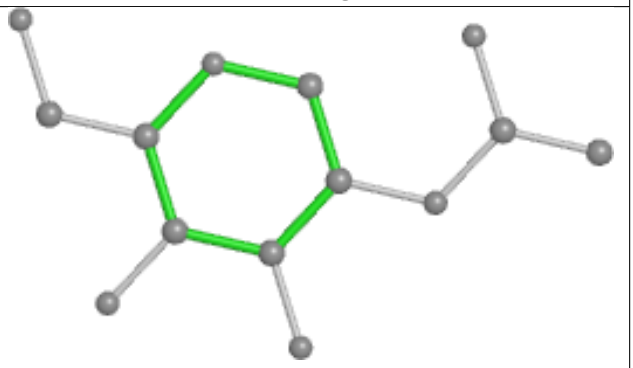
There are no ring outliers.

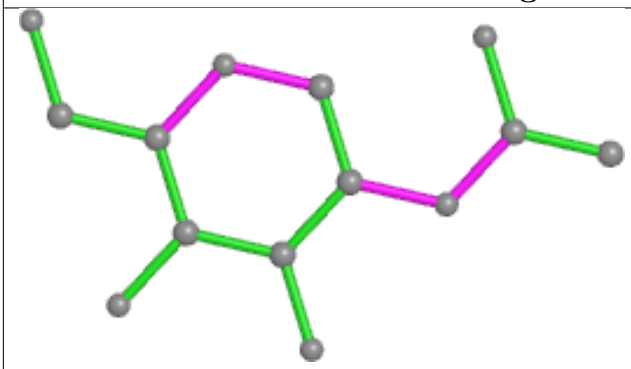
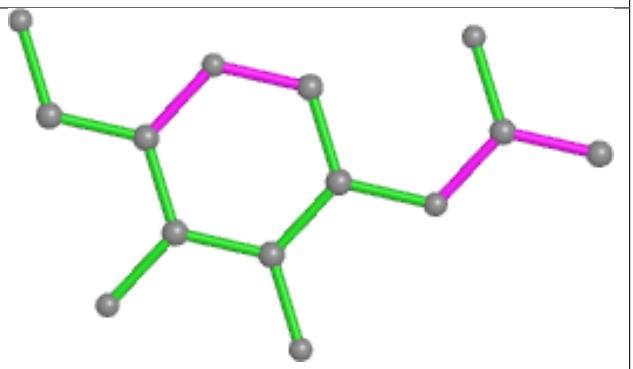
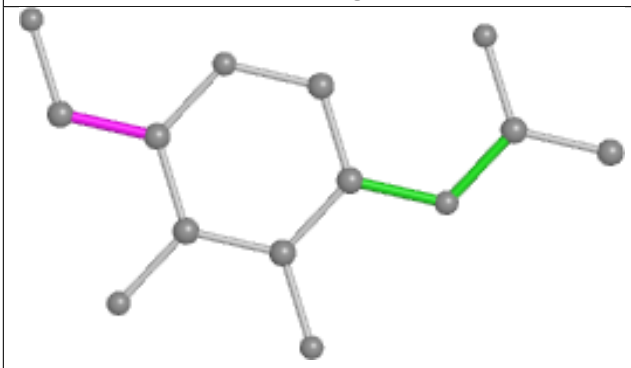
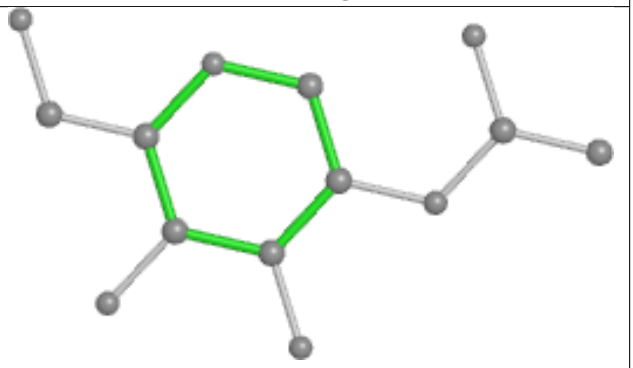
No monomer is involved in short contacts.

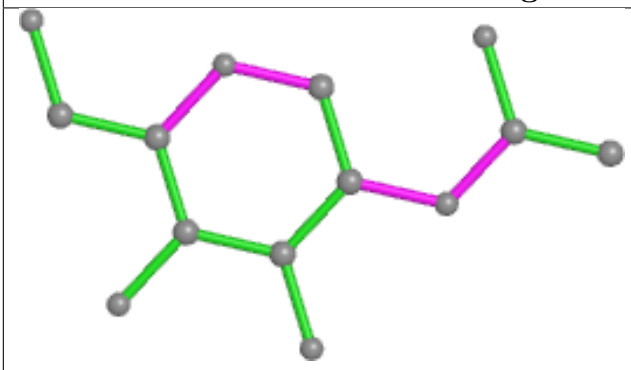
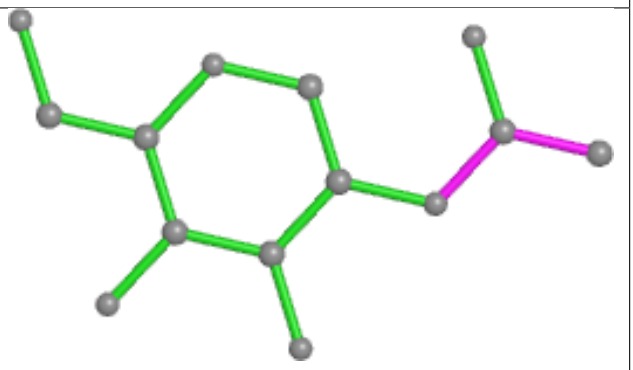
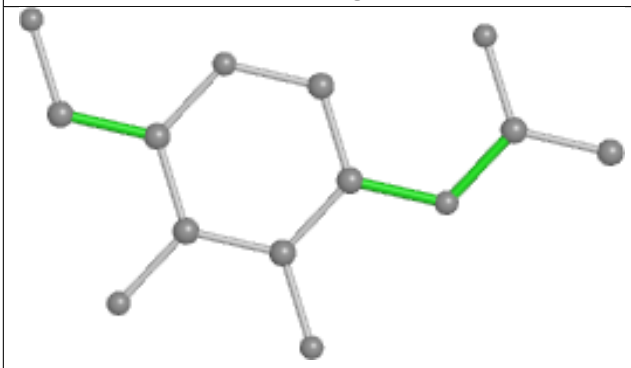
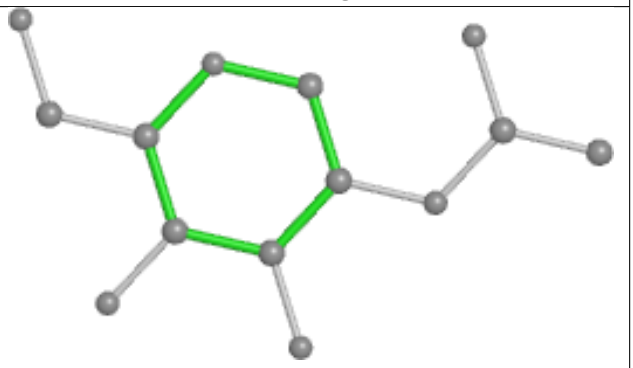
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.



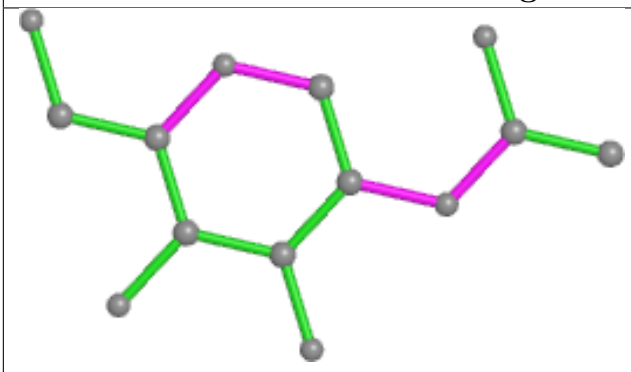
Ligand NAG A 601	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand NAG B 603	
 Bond lengths	 Bond angles
 Torsions	 Rings

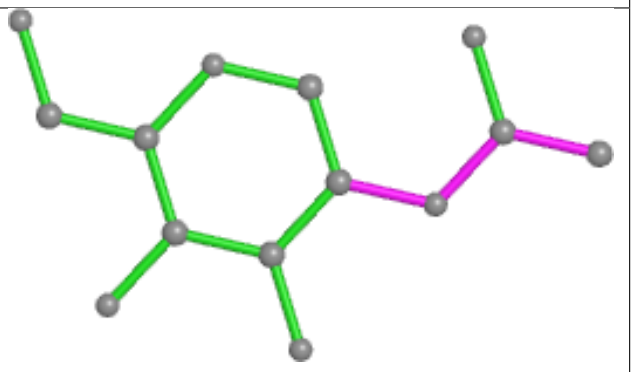
Ligand NAG A 603	
 Bond lengths	 Bond angles
 Torsions	 Rings

Ligand NAG B 602	
 Bond lengths	 Bond angles
 Torsions	 Rings

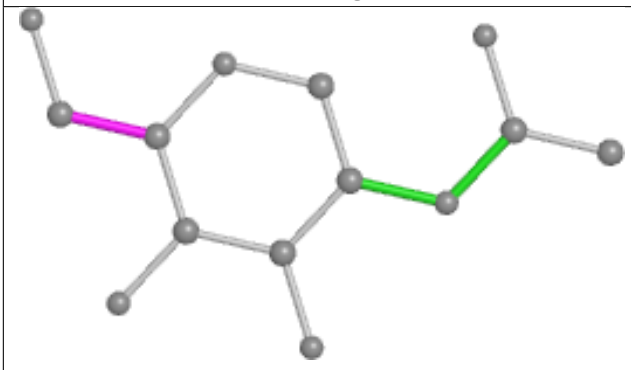
Ligand NAG A 602



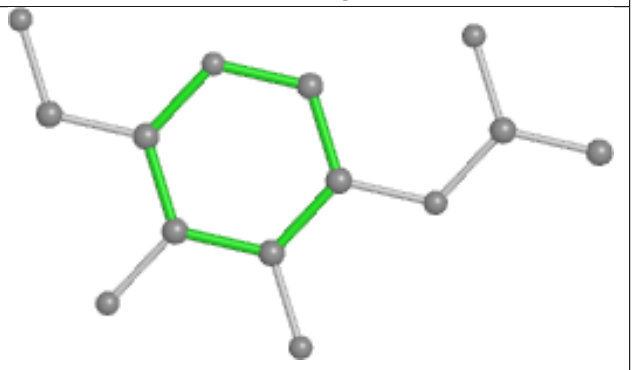
Bond lengths



Bond angles

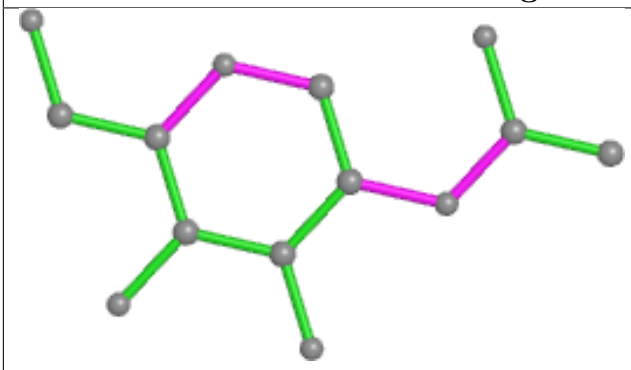


Torsions

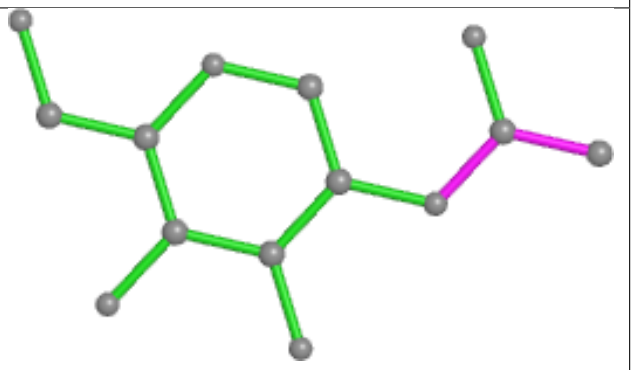


Rings

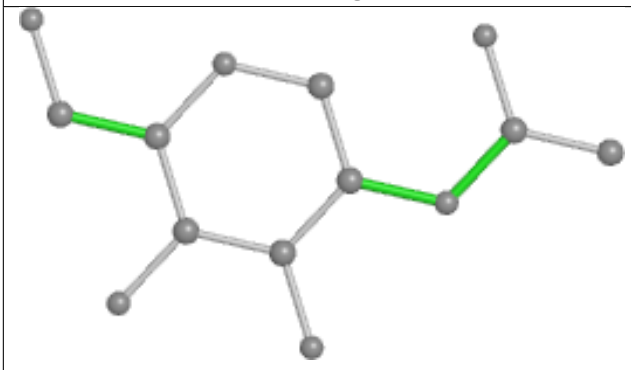
Ligand NAG C 601



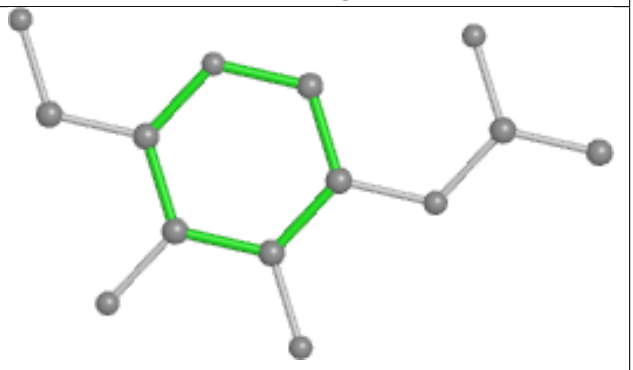
Bond lengths



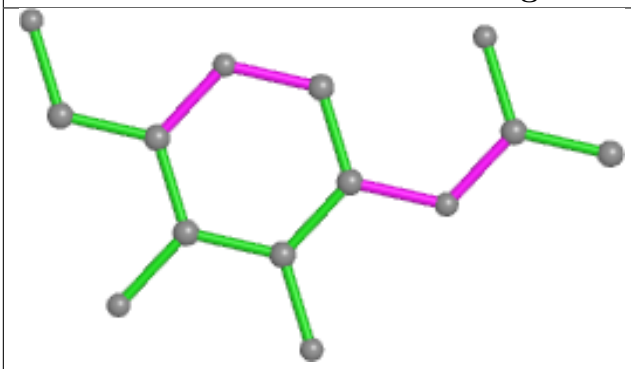
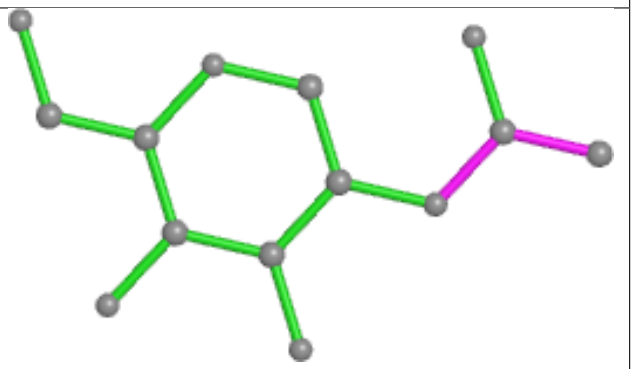
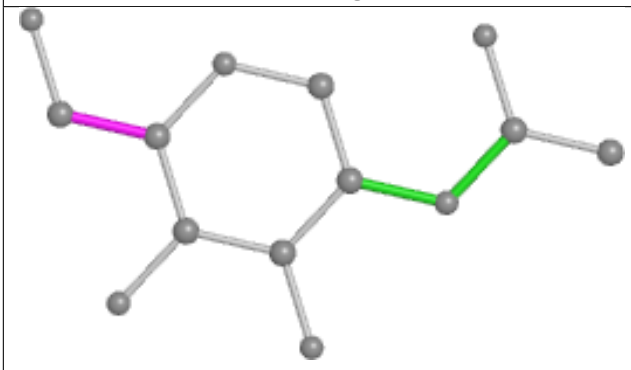
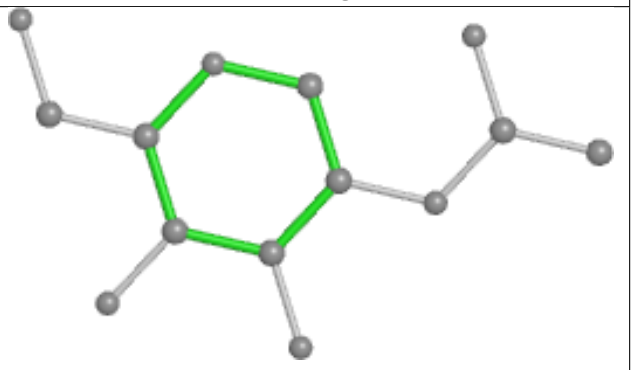
Bond angles



Torsions



Rings

Ligand NAG B 601	
 Bond lengths	 Bond angles
 Torsions	 Rings

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

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	492/506 (97%)	0.43	13 (2%) 57 51	37, 58, 90, 127	0
1	B	492/506 (97%)	0.49	14 (2%) 55 49	38, 64, 94, 120	0
1	C	489/506 (96%)	0.91	55 (11%) 11 9	39, 73, 132, 150	0
All	All	1473/1518 (97%)	0.61	82 (5%) 31 25	37, 64, 106, 150	0

All (82) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	469	PHE	9.1
1	C	497	LEU	6.0
1	C	470	TYR	4.4
1	C	71	ILE	4.2
1	C	500	GLU	3.8
1	A	274	CYS	3.4
1	C	66	MET	3.4
1	C	391	GLN	3.3
1	C	69	GLU	3.3
1	B	226	MET	3.2
1	B	469	PHE	3.2
1	B	391	GLN	3.2
1	B	135	CYS	3.1
1	C	489	PRO	3.1
1	A	476	GLU	3.1
1	A	1	ASP	3.1
1	C	321	PRO	3.1
1	A	500	GLU	3.1
1	C	181	SER	2.9
1	A	321	PRO	2.9
1	C	491	TYR	2.9
1	A	282	VAL	2.9
1	B	350	TRP	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	347	VAL	2.8
1	C	453	LEU	2.8
1	C	363	TYR	2.7
1	B	473	CYS	2.7
1	C	495	ALA	2.6
1	C	55	CYS	2.6
1	C	390	THR	2.6
1	A	418	LEU	2.5
1	C	365	ALA	2.5
1	C	474	ASP	2.5
1	C	486	TYR	2.5
1	B	191	TYR	2.5
1	C	358	GLU	2.5
1	C	218	GLN	2.5
1	C	462	LEU	2.4
1	C	499	ARG	2.4
1	C	11	ASN	2.4
1	B	500	GLU	2.4
1	C	356	SER	2.4
1	C	350	TRP	2.4
1	C	70	PHE	2.4
1	C	481	VAL	2.4
1	A	469	PHE	2.3
1	C	488	TYR	2.3
1	C	285	ILE	2.3
1	C	191	TYR	2.3
1	C	359	GLN	2.3
1	C	317	LEU	2.3
1	C	361	SER	2.3
1	C	485	THR	2.3
1	C	19	ILE	2.3
1	C	478	MET	2.2
1	C	456	ARG	2.2
1	B	321	PRO	2.2
1	C	190	LEU	2.2
1	C	333	GLY	2.2
1	B	71	ILE	2.2
1	C	354	HIS	2.2
1	A	390	THR	2.2
1	B	21	GLU	2.2
1	C	388	MET	2.1
1	B	302	CYS	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	460	LYS	2.1
1	C	346	MET	2.1
1	B	167	THR	2.1
1	A	65	PRO	2.1
1	C	332	PHE	2.1
1	C	369	SER	2.1
1	B	266	LYS	2.1
1	A	474	ASP	2.1
1	C	52	LEU	2.1
1	C	492	SER	2.1
1	C	355	HIS	2.1
1	C	282	VAL	2.0
1	A	471	HIS	2.0
1	C	286	ASN	2.0
1	C	451	VAL	2.0
1	A	463	GLY	2.0
1	C	475	ASN	2.0

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

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

6.3 Carbohydrates (i)

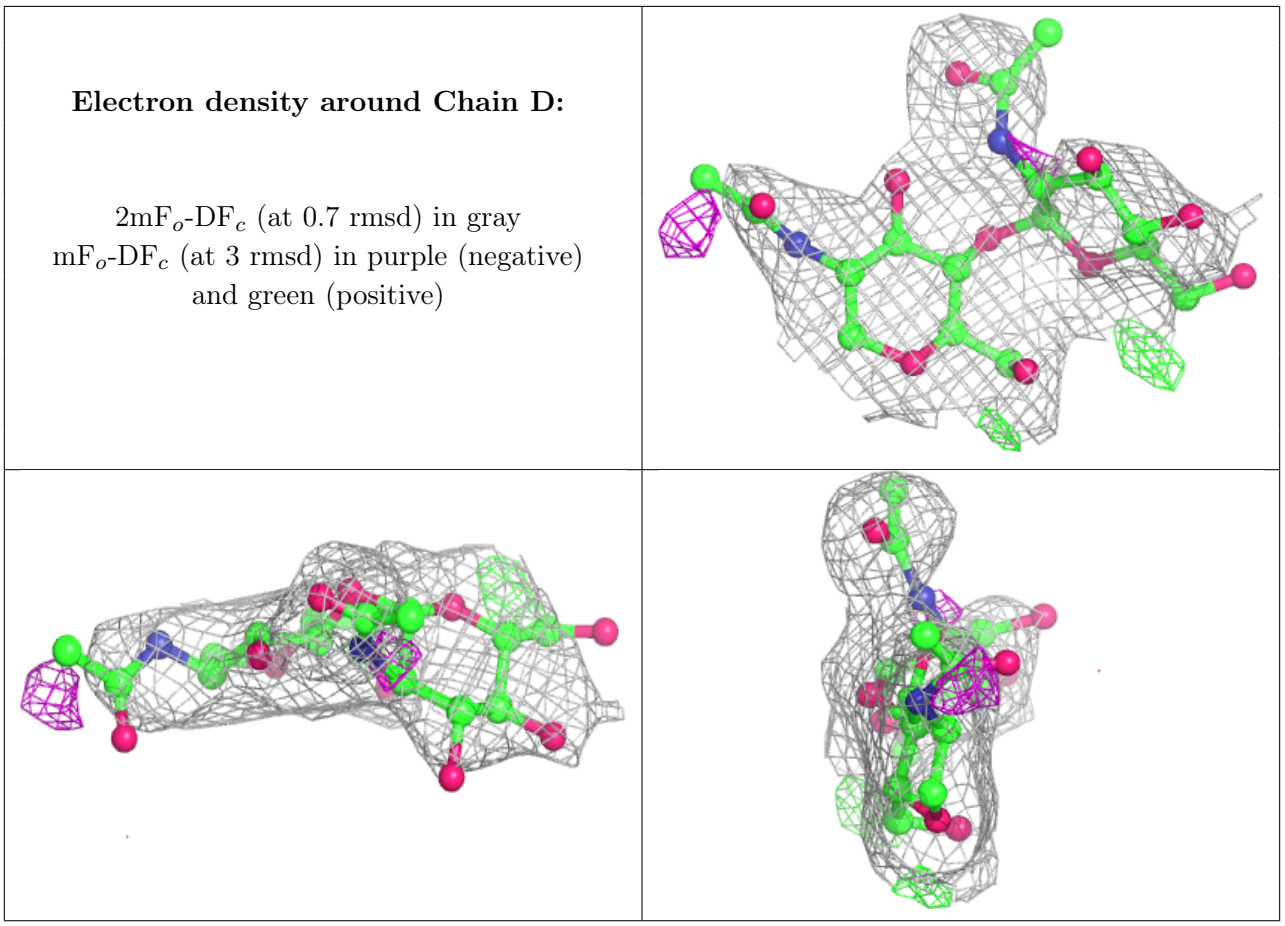
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	NAG	D	2	14/15	0.63	0.15	79,96,100,106	0
2	NAG	D	1	14/15	0.88	0.13	57,69,88,89	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 D:

$2mF_o$-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.

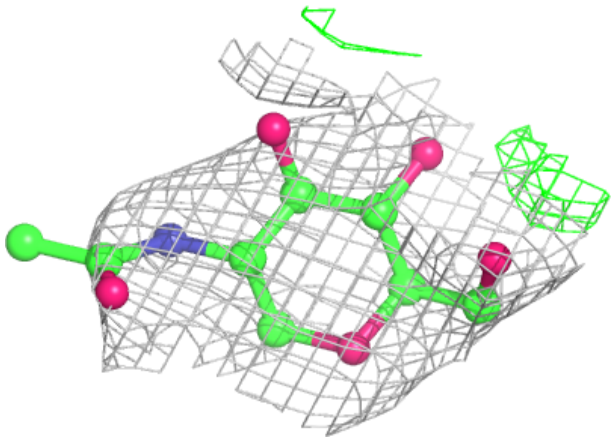
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	NAG	A	602	14/15	0.72	0.14	73,83,91,92	0
3	NAG	C	601	14/15	0.73	0.15	80,92,105,105	0
3	NAG	B	602	14/15	0.75	0.14	76,82,89,92	0
3	NAG	A	601	14/15	0.77	0.12	64,76,85,86	0
3	NAG	B	603	14/15	0.77	0.14	88,98,112,118	0
3	NAG	B	601	14/15	0.77	0.12	57,71,76,81	0
3	NAG	C	602	14/15	0.83	0.13	64,79,88,89	0
3	NAG	A	603	14/15	0.86	0.11	51,63,66,67	0

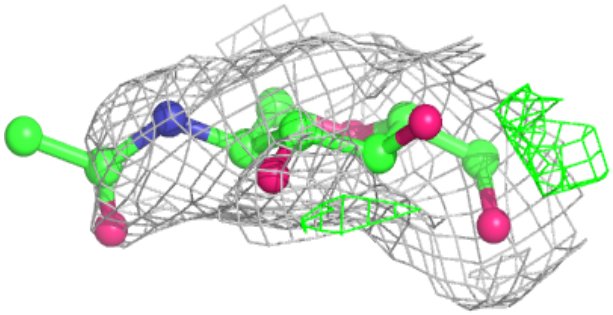
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different

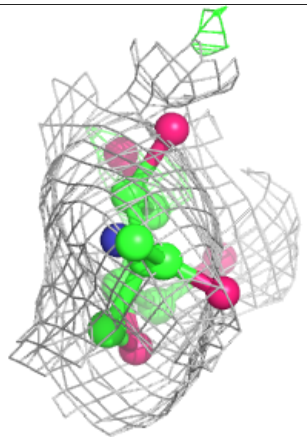
orientation to approximate a three-dimensional view.

Electron density around NAG A 602:

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

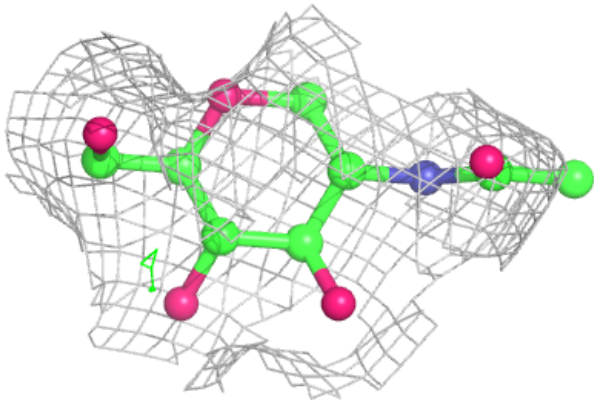


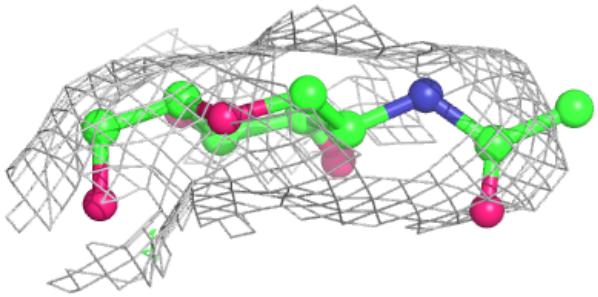


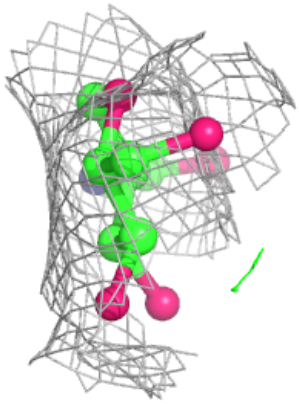


Electron density around NAG C 601:

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

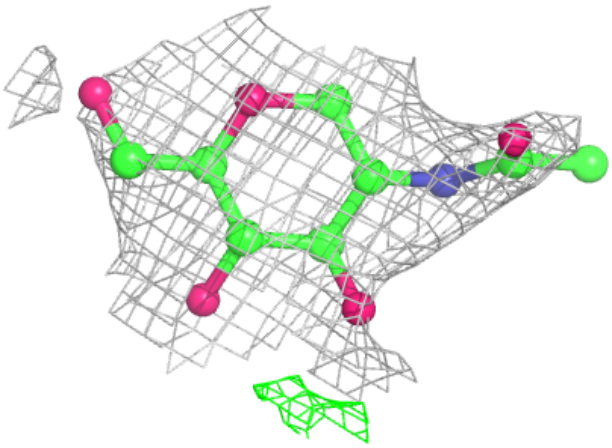


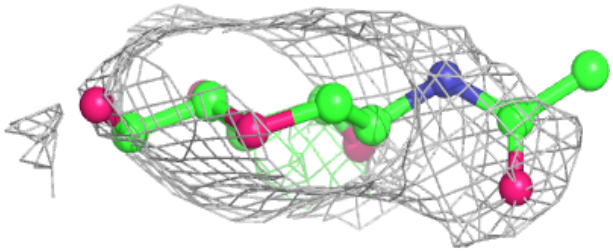


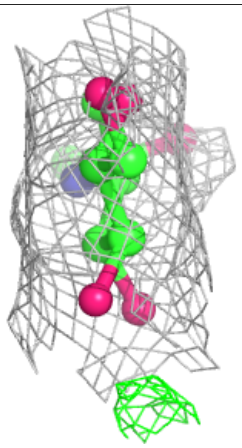


Electron density around NAG B 602:

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

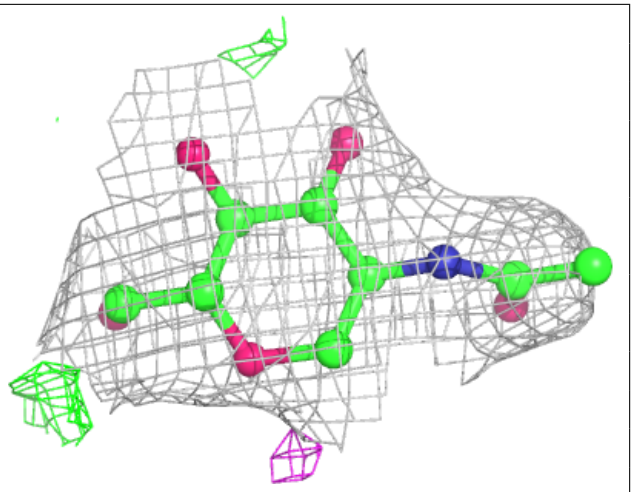


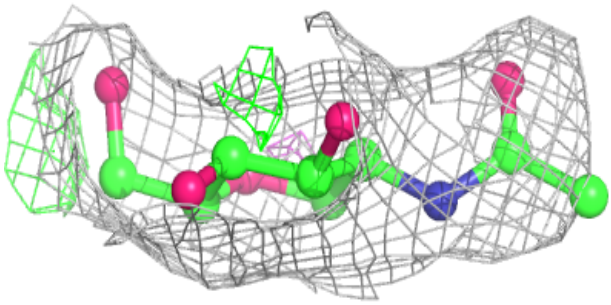


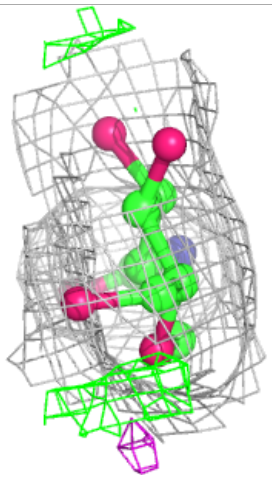


Electron density around NAG A 601:

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

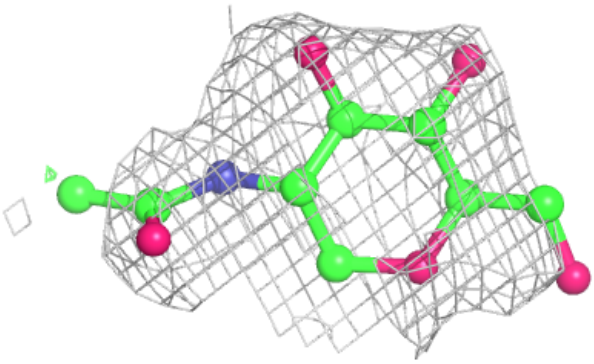


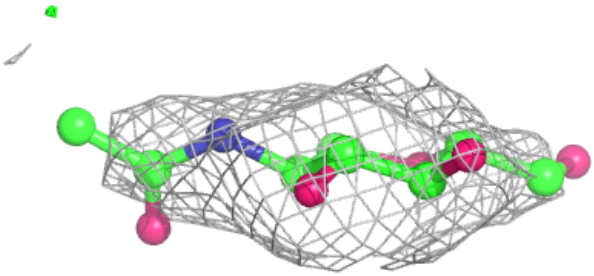


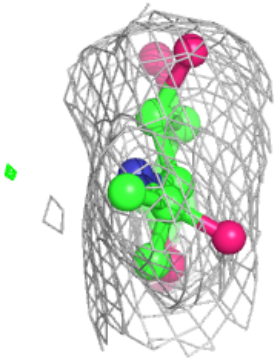


Electron density around NAG B 603:

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

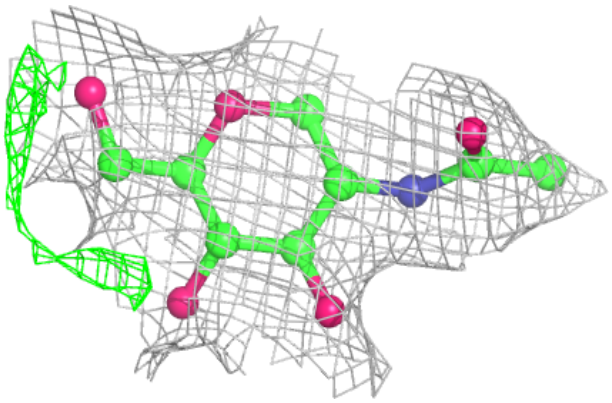


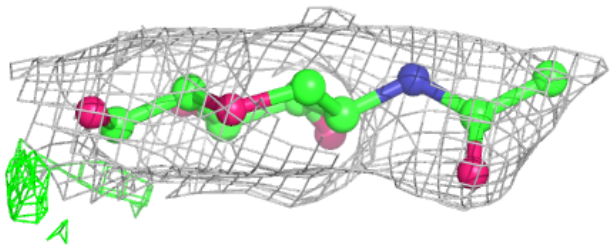


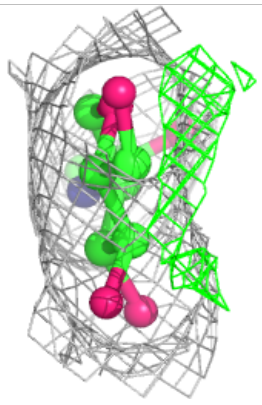


Electron density around NAG B 601:

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

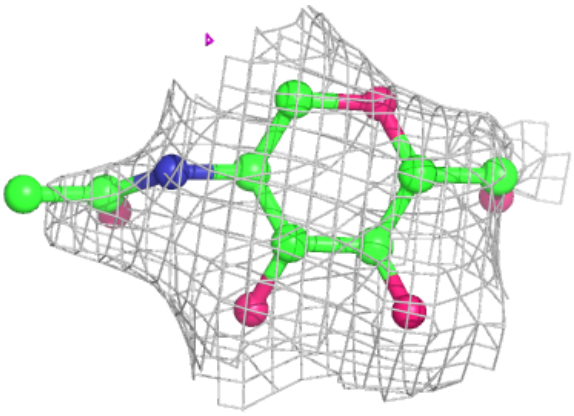


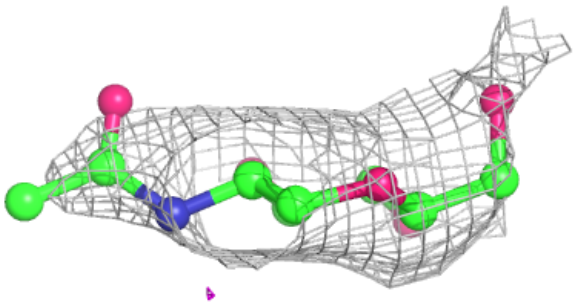


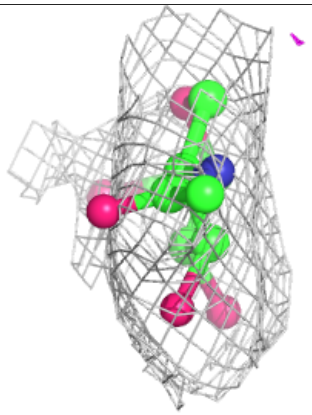


Electron density around NAG C 602:

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

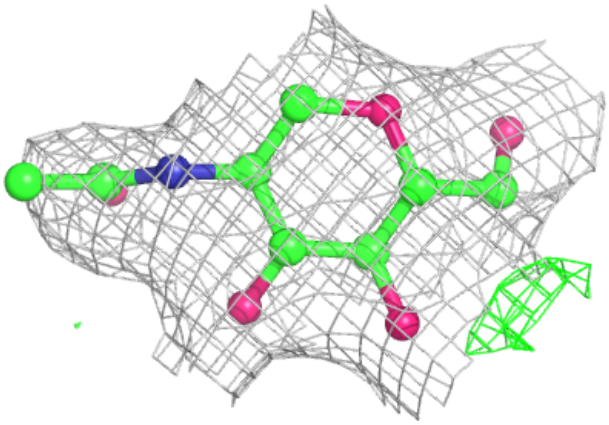


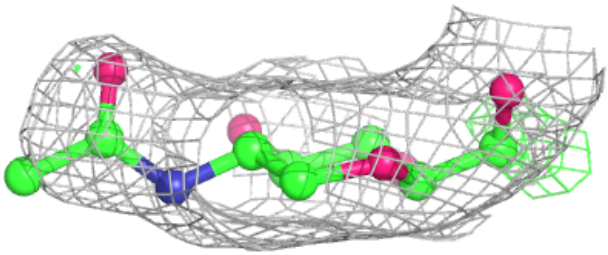


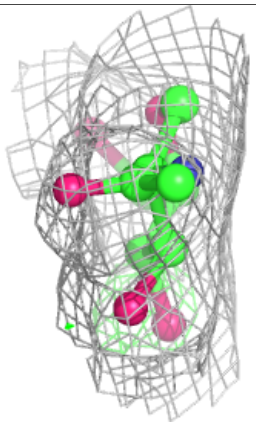


Electron density around NAG A 603:

$2mF_o$-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 (i)

There are no such residues in this entry.