Full wwPDB X-ray Structure Validation Report ⓘ

Jan 16, 2024 – 03:33 am GMT

PDB ID	:	6S22
Title	:	Crystal structure of the TgGalNAc-T3 in complex with UDP, manganese and FGF23c
Authors	:	de las Rivas, M.; Daniel, E.J.P.; Narimatsu, Y.; Companon, I.; Kato, K.; Hermosilla, P.; Thureau, A.; Ceballos-Laita, L.; Coelho, H.; Bernado, P.; Marcelo, F.; Hansen, L.; Lostao, A.; Corzana, F.; Clausen, H.; Gerken, T.A.; Hurtado-Guerrero, R.
Deposited on	:	2019-06-20
Resolution	:	1.96 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:

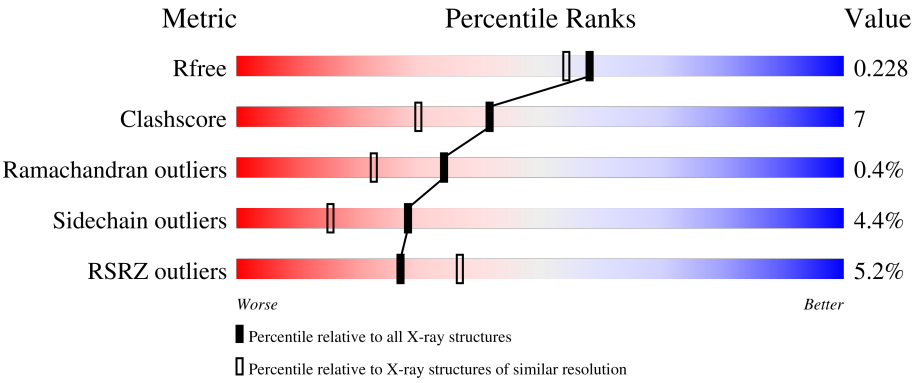
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)

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 1.96 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	631	4% 72% 10% • 16%
2	F	12	42% 58% 8% 33%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	EDO	A	711	-	-	X	-
9	NGA	F	201	X	-	-	-

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 4791 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.

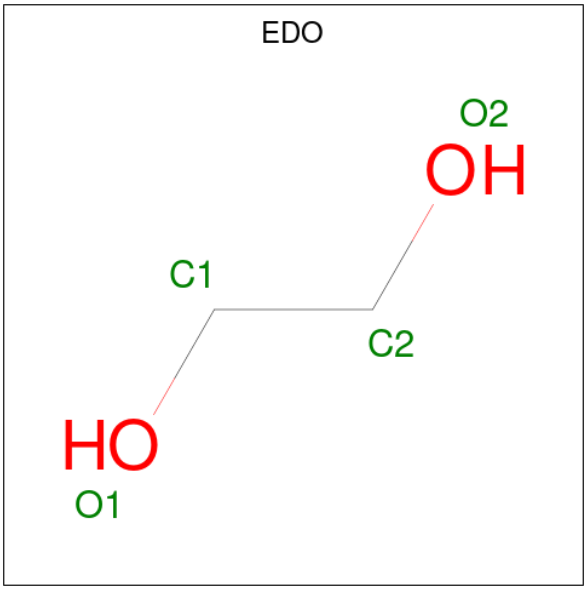
- Molecule 1 is a protein called Polypeptide N-acetylgalactosaminyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	529	Total	C	N	O	S	0	0	0
			4269	2716	741	789	23			

- Molecule 2 is a protein called Fibroblast growth factor 23.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	F	12	Total	C	N	O	0	0	0
			98	58	24	16			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



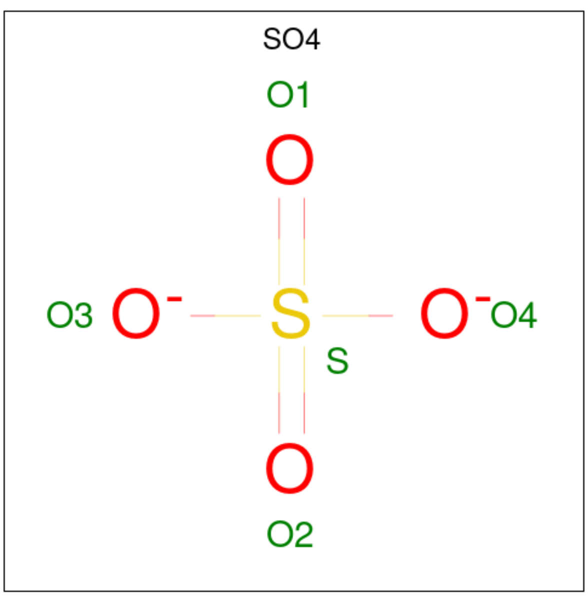
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		
3	A	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

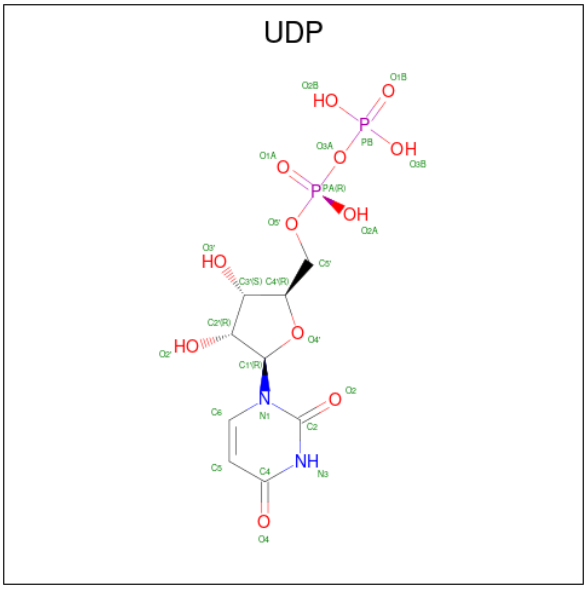
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 4 2 2	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 5 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: $C_9H_{14}N_2O_{12}P_2$).

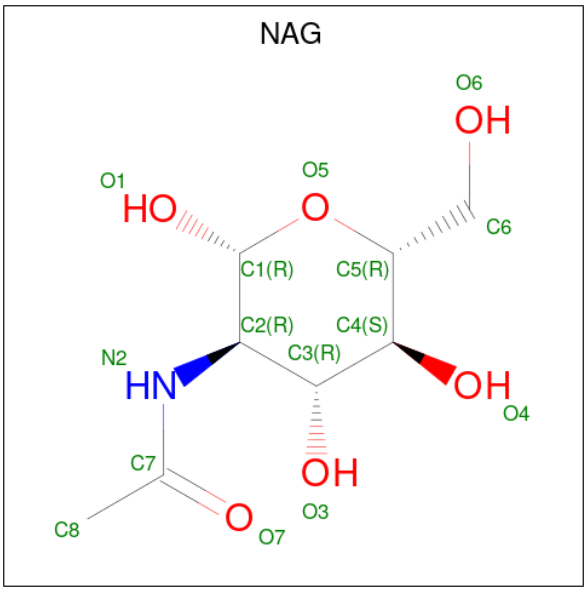


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O P 25 9 2 12 2	0	0

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

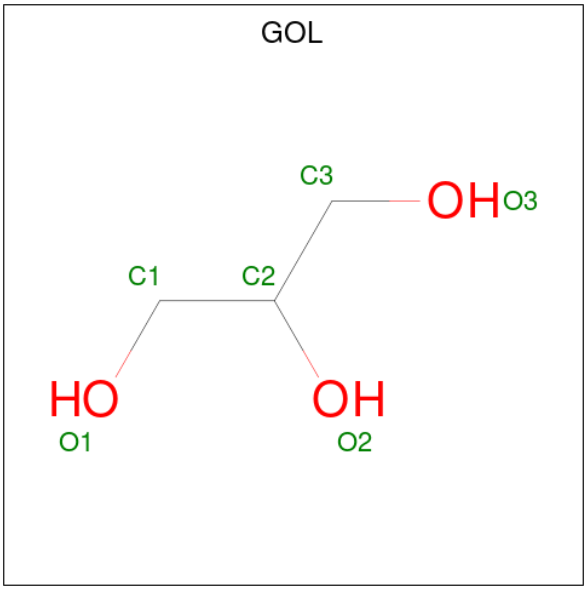
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Mn 1 1	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



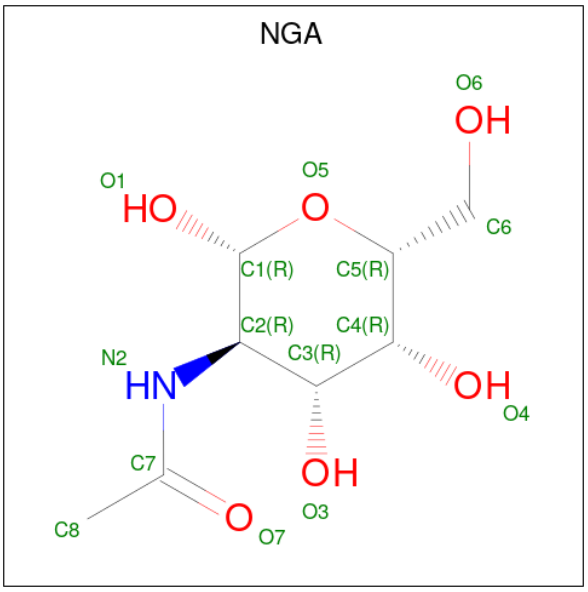
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C N O 14 8 1 5	0	0

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 9 is 2-acetamido-2-deoxy-beta-D-galactopyranose (three-letter code: NGA) (formula: $C_8H_{15}NO_6$) (labeled as "Ligand of Interest" by depositor).



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

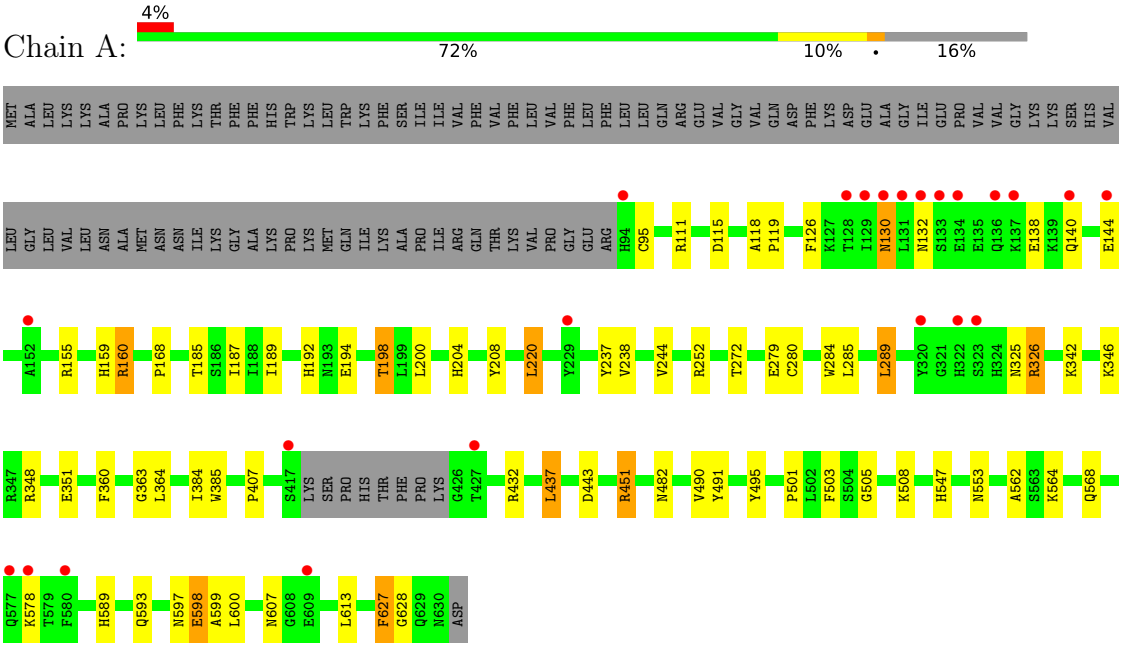
- Molecule 10 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	298	Total O 298 298	0	0
10	F	5	Total O 5 5	0	0

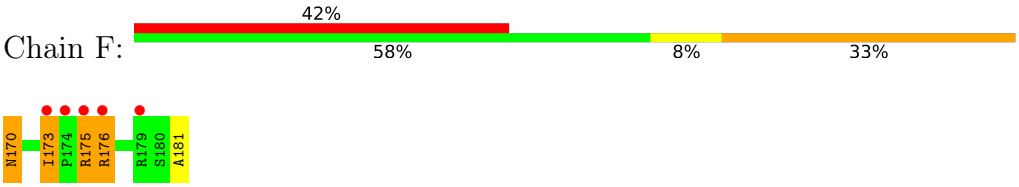
3 Residue-property plots (i)

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: Polypeptide N-acetylgalactosaminyltransferase



• Molecule 2: Fibroblast growth factor 23



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	50.96Å 104.82Å 143.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.01 – 1.96 50.96 – 1.96	Depositor EDS
% Data completeness (in resolution range)	100.0 (51.01-1.96) 100.0 (50.96-1.96)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.02 (at 1.97Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.179 , 0.216 0.188 , 0.228	Depositor DCC
R_{free} test set	2215 reflections (3.95%)	wwPDB-VP
Wilson B-factor (Å^2)	29.6	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.1	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	0.96	EDS
Total number of atoms	4791	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% 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: MN, NGA, EDO, NAG, SO4, UDP, GOL

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	0.74	0/4385	0.90	6/5939 (0.1%)
2	F	0.76	0/100	1.01	0/135
All	All	0.74	0/4485	0.90	6/6074 (0.1%)

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	160	ARG	NE-CZ-NH1	9.96	125.28	120.30
1	A	160	ARG	NE-CZ-NH2	-9.82	115.39	120.30
1	A	451	ARG	N-CA-CB	-5.37	100.93	110.60
1	A	432	ARG	NE-CZ-NH2	-5.37	117.61	120.30
1	A	326	ARG	CG-CD-NE	5.32	122.97	111.80
1	A	326	ARG	CB-CA-C	-5.23	99.95	110.40

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	4269	0	4129	51	0
2	F	98	0	99	7	0
3	A	56	0	83	11	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	5	0	0	0	0
5	A	25	0	11	0	0
6	A	1	0	0	0	0
7	A	14	0	13	0	0
8	A	6	0	8	1	0
9	F	14	0	13	0	0
10	A	298	0	0	12	0
10	F	5	0	0	0	0
All	All	4791	0	4356	62	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 7.

All (62) 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:562:ALA:HB1	1:A:613:LEU:HD11	1.67	0.76
8:A:719:GOL:H12	10:A:892:HOH:O	1.86	0.75
1:A:160:ARG:HD3	1:A:279:GLU:OE1	1.87	0.75
1:A:598:GLU:CA	10:A:806:HOH:O	2.39	0.70
1:A:194:GLU:OE2	1:A:198:THR:HG21	1.94	0.68
1:A:597:ASN:HD22	1:A:600:LEU:H	1.40	0.68
2:F:170:ASN:C	2:F:170:ASN:HD22	1.97	0.68
1:A:200:LEU:HD23	1:A:237:TYR:CE2	2.30	0.66
1:A:285:LEU:HG	1:A:289:LEU:HD22	1.76	0.66
2:F:173:ILE:O	2:F:175:ARG:NH2	2.30	0.63
1:A:119:PRO:HB3	1:A:126:PHE:CE1	2.36	0.61
1:A:598:GLU:HA	10:A:806:HOH:O	2.00	0.59
1:A:598:GLU:CB	10:A:806:HOH:O	2.50	0.59
3:A:708:EDO:H11	3:A:711:EDO:C2	2.33	0.58
1:A:598:GLU:HB3	10:A:806:HOH:O	2.04	0.58
1:A:159:HIS:HD2	1:A:208:TYR:OH	1.87	0.57
1:A:194:GLU:OE2	1:A:198:THR:CG2	2.51	0.57
2:F:170:ASN:C	2:F:170:ASN:ND2	2.57	0.57
3:A:708:EDO:H11	3:A:711:EDO:O1	2.06	0.55
1:A:115:ASP:HB3	1:A:118:ALA:HB2	1.89	0.55
1:A:589:HIS:HD2	3:A:707:EDO:O1	1.91	0.53
1:A:192:HIS:CE1	1:A:252:ARG:HD2	2.44	0.53
3:A:709:EDO:H11	3:A:711:EDO:O1	2.09	0.53
1:A:443:ASP:OD2	1:A:482:ASN:HB2	2.10	0.51
1:A:597:ASN:HD21	1:A:599:ALA:HB3	1.75	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:111:ARG:NH1	10:A:812:HOH:O	2.45	0.50
1:A:363:GLY:C	1:A:364:LEU:HD12	2.32	0.49
1:A:495:TYR:HE2	1:A:553:ASN:HB2	1.78	0.49
1:A:140:GLN:O	1:A:144:GLU:HG3	2.13	0.49
1:A:597:ASN:ND2	1:A:600:LEU:H	2.09	0.48
1:A:159:HIS:CD2	1:A:208:TYR:OH	2.67	0.48
1:A:138:GLU:OE1	1:A:155:ARG:NH2	2.47	0.48
1:A:593:GLN:HE22	3:A:711:EDO:C1	2.26	0.47
2:F:170:ASN:ND2	2:F:170:ASN:O	2.47	0.47
1:A:326:ARG:HD2	10:A:935:HOH:O	2.14	0.47
1:A:547:HIS:HD2	10:A:1061:HOH:O	1.97	0.47
1:A:598:GLU:HG3	10:A:826:HOH:O	2.15	0.47
1:A:325:ASN:OD1	2:F:181:ALA:HB1	2.15	0.47
3:A:705:EDO:H11	10:A:1052:HOH:O	2.14	0.47
1:A:168:PRO:HG2	1:A:351:GLU:HG3	1.97	0.46
1:A:140:GLN:O	1:A:144:GLU:CG	2.63	0.46
1:A:280:CYS:HB3	1:A:284:TRP:CD1	2.50	0.46
3:A:708:EDO:H11	3:A:711:EDO:O2	2.15	0.46
1:A:508:LYS:HD2	3:A:710:EDO:H22	1.96	0.46
1:A:187:ILE:HD13	1:A:272:THR:HB	1.99	0.45
1:A:111:ARG:NH2	10:A:822:HOH:O	2.49	0.45
2:F:175:ARG:O	2:F:176:ARG:HB3	2.16	0.45
1:A:238:VAL:HG13	1:A:244:VAL:HB	1.98	0.45
1:A:451:ARG:HD3	1:A:501:PRO:HG3	1.99	0.45
1:A:508:LYS:CD	3:A:710:EDO:H22	2.47	0.44
1:A:589:HIS:CD2	3:A:707:EDO:O1	2.69	0.44
1:A:189:ILE:HB	1:A:220:LEU:HD12	1.99	0.44
1:A:130:ASN:HD22	1:A:130:ASN:HA	1.59	0.43
1:A:348:ARG:NH2	1:A:407:PRO:HG2	2.32	0.43
1:A:200:LEU:CD2	1:A:237:TYR:CE2	3.00	0.43
1:A:505:GLY:HA3	1:A:628:GLY:O	2.19	0.42
1:A:437:LEU:HD12	1:A:437:LEU:C	2.39	0.42
1:A:598:GLU:CB	10:A:826:HOH:O	2.66	0.42
1:A:200:LEU:O	1:A:204:HIS:HD2	2.03	0.41
1:A:503:PHE:HE1	1:A:627:PHE:CD1	2.38	0.40
1:A:385:TRP:HZ2	2:F:176:ARG:O	2.05	0.40
1:A:490:VAL:HA	3:A:705:EDO:H21	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	525/631 (83%)	509 (97%)	15 (3%)	1 (0%)	47 38
2	F	10/12 (83%)	6 (60%)	3 (30%)	1 (10%)	0 0
All	All	535/643 (83%)	515 (96%)	18 (3%)	2 (0%)	34 22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	176	ARG
1	A	384	ILE

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	466/556 (84%)	448 (96%)	18 (4%)	32 19
2	F	11/11 (100%)	8 (73%)	3 (27%)	0 0
All	All	477/567 (84%)	456 (96%)	21 (4%)	28 15

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	95	CYS
1	A	130	ASN
1	A	132	ASN
1	A	185	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	198	THR
1	A	220	LEU
1	A	289	LEU
1	A	342	LYS
1	A	346	LYS
1	A	360	PHE
1	A	437	LEU
1	A	491	TYR
1	A	564	LYS
1	A	568	GLN
1	A	578	LYS
1	A	598	GLU
1	A	607	ASN
1	A	627	PHE
2	F	170	ASN
2	F	173	ILE
2	F	175	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	114	GLN
1	A	130	ASN
1	A	136	GLN
1	A	159	HIS
1	A	174	GLN
1	A	523	HIS
1	A	589	HIS
1	A	593	GLN
1	A	597	ASN
1	A	607	ASN
2	F	170	ASN
2	F	177	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)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 20 ligands modelled in this entry, 1 is monoatomic - leaving 19 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
7	NAG	A	718	1	14,14,15	0.38	0	17,19,21	1.44	2 (11%)
3	EDO	A	707	-	3,3,3	0.27	0	2,2,2	0.35	0
8	GOL	A	719	-	5,5,5	0.42	0	5,5,5	0.91	0
3	EDO	A	702	-	3,3,3	0.12	0	2,2,2	0.33	0
3	EDO	A	704	-	3,3,3	0.38	0	2,2,2	0.21	0
3	EDO	A	711	-	3,3,3	0.38	0	2,2,2	0.70	0
3	EDO	A	706	-	3,3,3	0.06	0	2,2,2	0.35	0
3	EDO	A	708	-	3,3,3	0.38	0	2,2,2	0.59	0
4	SO4	A	715	-	4,4,4	0.24	0	6,6,6	0.08	0
3	EDO	A	705	-	3,3,3	0.29	0	2,2,2	0.24	0
9	NGA	F	201	2	14,14,15	1.12	1 (7%)	17,19,21	1.00	0
3	EDO	A	703	-	3,3,3	0.09	0	2,2,2	0.18	0
3	EDO	A	709	-	3,3,3	0.71	0	2,2,2	0.47	0
5	UDP	A	716	6	24,26,26	1.18	4 (16%)	37,40,40	1.85	8 (21%)
3	EDO	A	712	-	3,3,3	0.43	0	2,2,2	0.53	0
3	EDO	A	714	-	3,3,3	0.21	0	2,2,2	0.08	0
3	EDO	A	701	-	3,3,3	0.10	0	2,2,2	0.70	0
3	EDO	A	713	-	3,3,3	0.22	0	2,2,2	0.38	0
3	EDO	A	710	-	3,3,3	0.44	0	2,2,2	0.20	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
3	EDO	A	712	-	-	1/1/1/1	-
3	EDO	A	714	-	-	1/1/1/1	-
7	NAG	A	718	1	-	0/6/23/26	0/1/1/1
3	EDO	A	711	-	-	0/1/1/1	-
3	EDO	A	707	-	-	0/1/1/1	-
3	EDO	A	701	-	-	1/1/1/1	-
3	EDO	A	706	-	-	1/1/1/1	-
8	GOL	A	719	-	-	4/4/4/4	-
3	EDO	A	709	-	-	0/1/1/1	-
3	EDO	A	702	-	-	1/1/1/1	-
3	EDO	A	708	-	-	1/1/1/1	-
3	EDO	A	713	-	-	0/1/1/1	-
3	EDO	A	704	-	-	0/1/1/1	-
3	EDO	A	705	-	-	1/1/1/1	-
5	UDP	A	716	6	-	0/16/32/32	0/2/2/2
3	EDO	A	710	-	-	1/1/1/1	-
9	NGA	F	201	2	1/1/5/7	0/6/23/26	0/1/1/1
3	EDO	A	703	-	-	1/1/1/1	-

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	F	201	NGA	C1-C2	3.36	1.57	1.52
5	A	716	UDP	C4-N3	-3.02	1.33	1.38
5	A	716	UDP	C6-C5	2.62	1.41	1.35
5	A	716	UDP	C2-N3	-2.47	1.33	1.38
5	A	716	UDP	C2-N1	2.40	1.42	1.38

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	716	UDP	N3-C2-N1	6.14	123.04	114.89
5	A	716	UDP	C4-N3-C2	-4.22	121.01	126.58
5	A	716	UDP	O2-C2-N3	-3.49	115.00	121.50
7	A	718	NAG	O5-C5-C6	3.44	112.60	107.20
5	A	716	UDP	C6-N1-C2	-3.15	116.96	120.99
7	A	718	NAG	C1-O5-C5	-2.76	108.45	112.19
5	A	716	UDP	C5-C4-N3	2.63	118.77	114.84
5	A	716	UDP	O2B-PB-O1B	2.42	120.14	110.68
5	A	716	UDP	PA-O3A-PB	-2.28	125.01	132.83

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	716	UDP	O5'-PA-O1A	2.02	116.97	109.07

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
9	F	201	NGA	C1

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	A	719	GOL	O1-C1-C2-C3
8	A	719	GOL	C1-C2-C3-O3
8	A	719	GOL	O1-C1-C2-O2
8	A	719	GOL	O2-C2-C3-O3
3	A	702	EDO	O1-C1-C2-O2
3	A	708	EDO	O1-C1-C2-O2
3	A	701	EDO	O1-C1-C2-O2
3	A	706	EDO	O1-C1-C2-O2
3	A	714	EDO	O1-C1-C2-O2
3	A	703	EDO	O1-C1-C2-O2
3	A	705	EDO	O1-C1-C2-O2
3	A	710	EDO	O1-C1-C2-O2
3	A	712	EDO	O1-C1-C2-O2

There are no ring outliers.

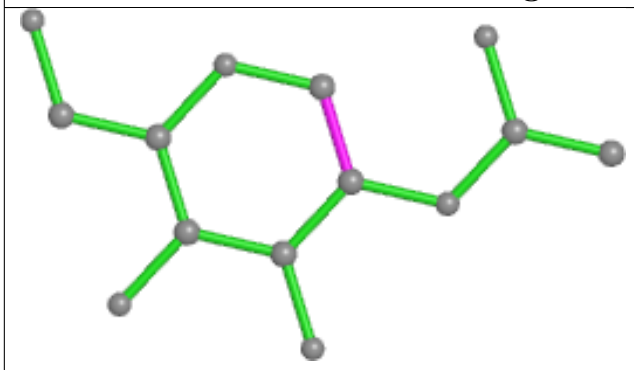
7 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	707	EDO	2	0
8	A	719	GOL	1	0
3	A	711	EDO	5	0
3	A	708	EDO	3	0
3	A	705	EDO	2	0
3	A	709	EDO	1	0
3	A	710	EDO	2	0

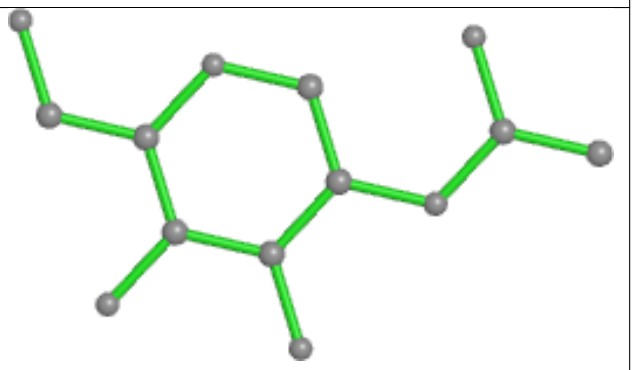
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

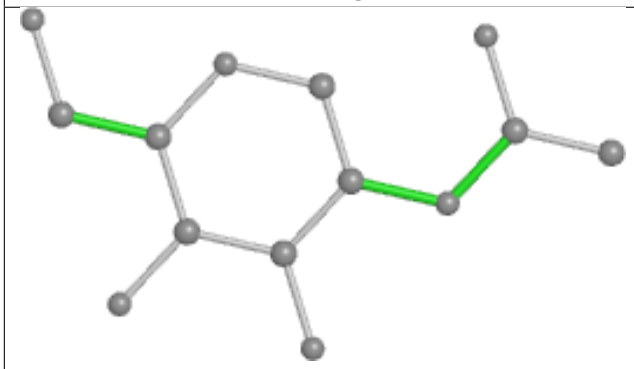
Ligand NGA F 201



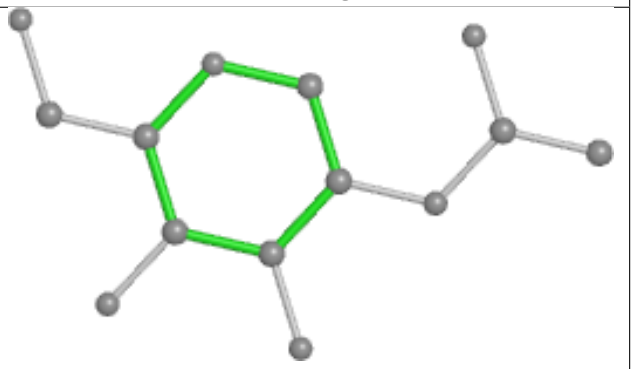
Bond lengths



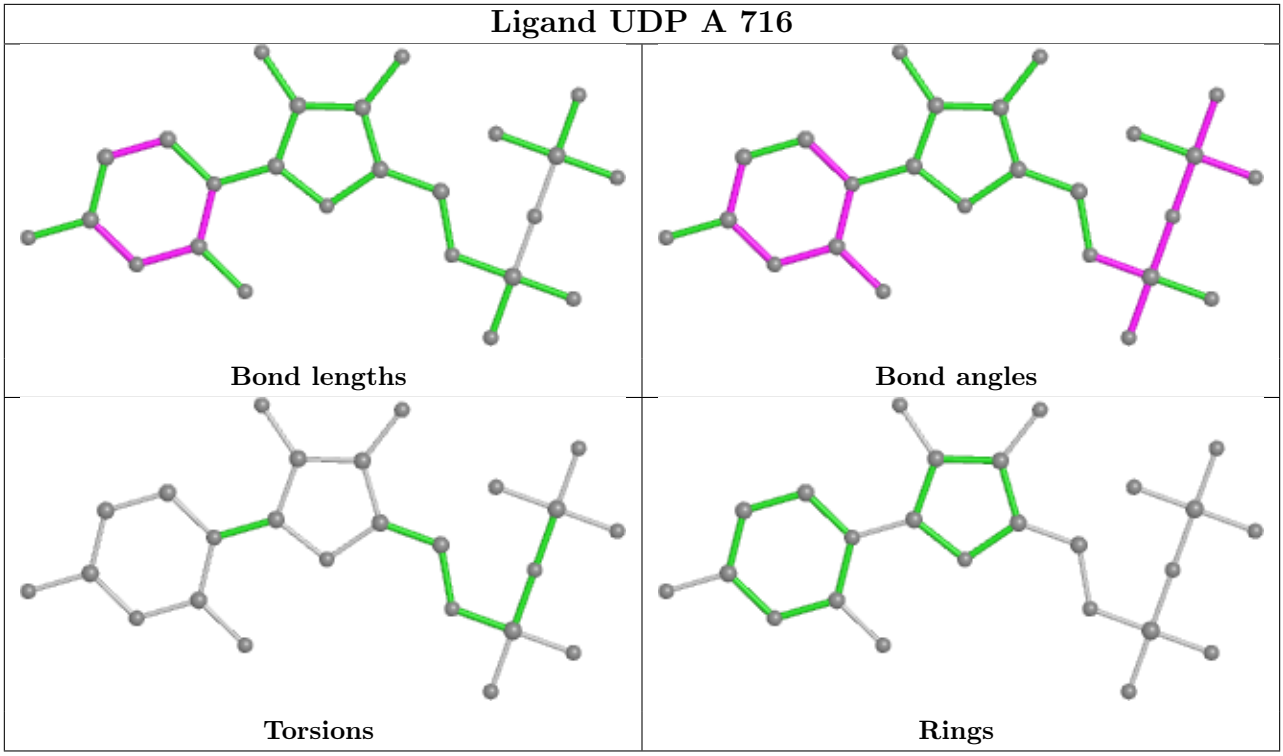
Bond angles



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	529/631 (83%)	0.02	23 (4%) 35 45	19, 32, 67, 117	0
2	F	12/12 (100%)	2.16	5 (41%) 0 0	52, 80, 115, 118	0
All	All	541/643 (84%)	0.06	28 (5%) 27 37	19, 32, 75, 118	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	174	PRO	6.9
1	A	94	HIS	5.7
1	A	577	GLN	5.1
1	A	580	PHE	4.9
2	F	176	ARG	4.7
1	A	130	ASN	4.2
1	A	131	LEU	4.2
1	A	322	HIS	4.0
1	A	129	ILE	3.7
1	A	134	GLU	3.5
2	F	175	ARG	3.4
2	F	173	ILE	3.4
1	A	128	THR	3.3
1	A	140	GLN	3.1
1	A	320	TYR	2.8
1	A	132	ASN	2.7
1	A	427	THR	2.6
1	A	133	SER	2.5
1	A	609	GLU	2.4
1	A	152	ALA	2.4
1	A	144	GLU	2.4
2	F	179	ARG	2.4
1	A	229	TYR	2.4
1	A	323	SER	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	578	LYS	2.3
1	A	136	GLN	2.1
1	A	137	LYS	2.0
1	A	417	SER	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)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{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	EDO	A	714	4/4	0.75	0.20	56,58,64,72	0
3	EDO	A	713	4/4	0.82	0.20	61,62,64,67	0
8	GOL	A	719	6/6	0.84	0.22	28,40,47,50	0
3	EDO	A	712	4/4	0.86	0.25	60,62,63,66	0
3	EDO	A	711	4/4	0.87	0.13	49,65,65,71	0
3	EDO	A	710	4/4	0.88	0.16	53,54,57,59	0
3	EDO	A	703	4/4	0.88	0.12	56,57,60,61	0
3	EDO	A	707	4/4	0.88	0.14	47,56,57,59	0
4	SO4	A	715	5/5	0.89	0.32	71,79,87,94	0
3	EDO	A	705	4/4	0.91	0.24	48,53,53,66	0
3	EDO	A	709	4/4	0.92	0.17	34,38,38,44	0
3	EDO	A	702	4/4	0.93	0.16	42,45,47,50	0
3	EDO	A	706	4/4	0.93	0.13	39,48,48,52	0
3	EDO	A	701	4/4	0.93	0.25	49,49,50,52	0
3	EDO	A	708	4/4	0.93	0.33	41,43,45,51	0
7	NAG	A	718	14/15	0.94	0.17	38,51,57,57	0
3	EDO	A	704	4/4	0.95	0.12	39,39,42,48	0
5	UDP	A	716	25/25	0.96	0.12	40,52,87,91	0
9	NGA	F	201	14/15	0.96	0.07	28,31,40,45	0

Continued on next page...

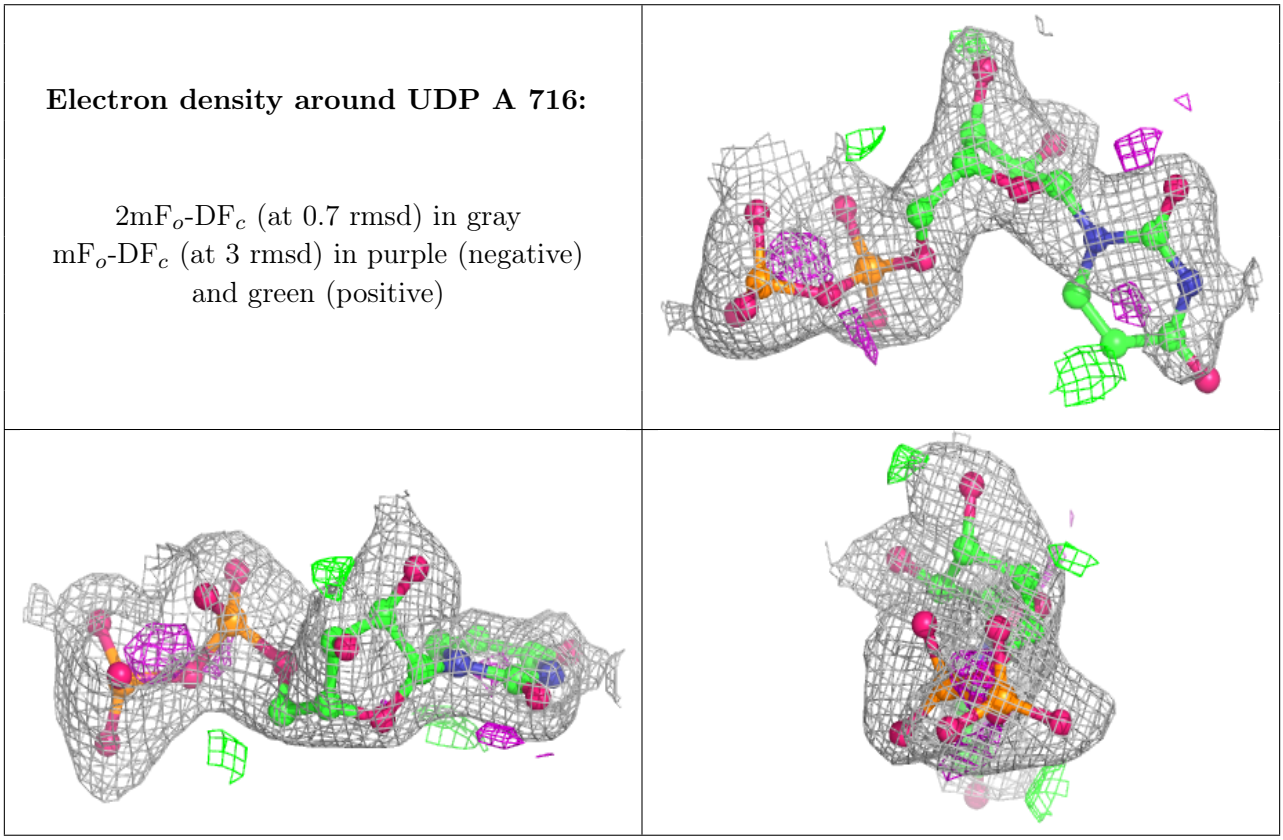
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	MN	A	717	1/1	0.99	0.08	34,34,34,34	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.

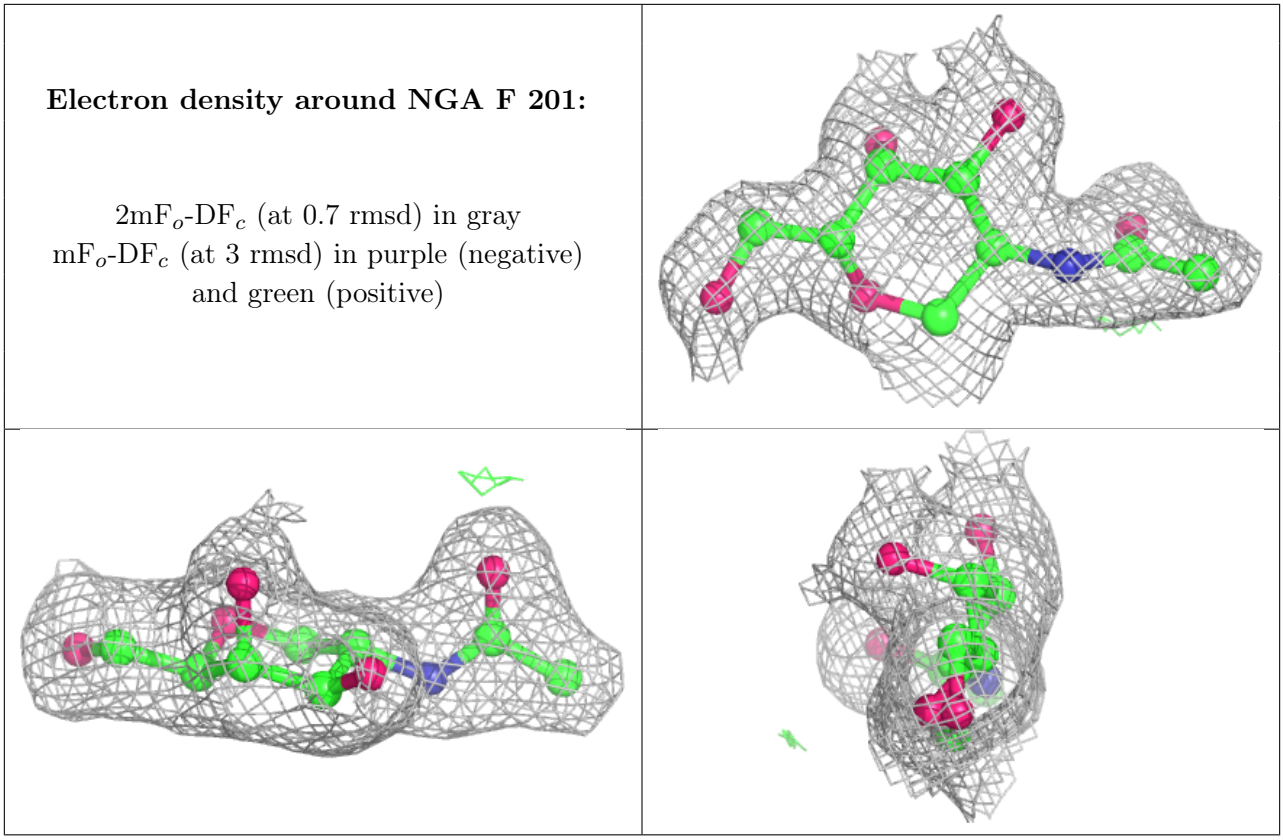
Electron density around UDP A 716:

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



Electron density around NGA F 201:

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



6.5 Other polymers (i)

There are no such residues in this entry.