



Full wwPDB X-ray Structure Validation Report ⓘ

Apr 5, 2022 – 01:27 pm BST

PDB ID : 7Q4I
Title : Crystal structure of DmC1GalT1 in complex with UDP-Mn²⁺ and the APD-TGalNAc-RP
Authors : Gonzalez-Ramirez, A.M.; Coelho, H.; Companon, I.; Grosso, A.S.; Yang, Z.; Narimatsu, Y.; Clausen, H.; Marcelo, F.; Corzana, F.; Hurtado-Guerrero, R.
Deposited on : 2021-10-31
Resolution : 2.40 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 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 : 2.27
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.27

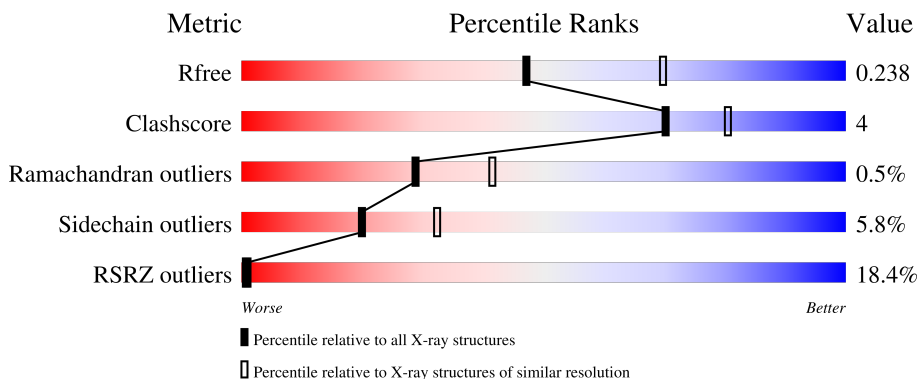
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 $\leq 5\%$. 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	316	 17% (poor fit), 79% (0-3 outliers), 9% (1 outlier), 12% (2 outliers), 3% (3+ outliers), 0% (not modelled)
1	B	316	 15% (poor fit), 77% (0-3 outliers), 10% (1 outlier), 11% (2 outliers), 1% (3+ outliers), 0% (not modelled)
2	F	6	 33% (poor fit), 33% (0-3 outliers), 50% (1 outlier), 17% (2 outliers), 0% (3+ outliers), 0% (not modelled)
2	G	6	 33% (poor fit), 67% (0-3 outliers), 17% (1 outlier), 17% (2 outliers), 0% (3+ outliers), 0% (not modelled)

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 4830 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.

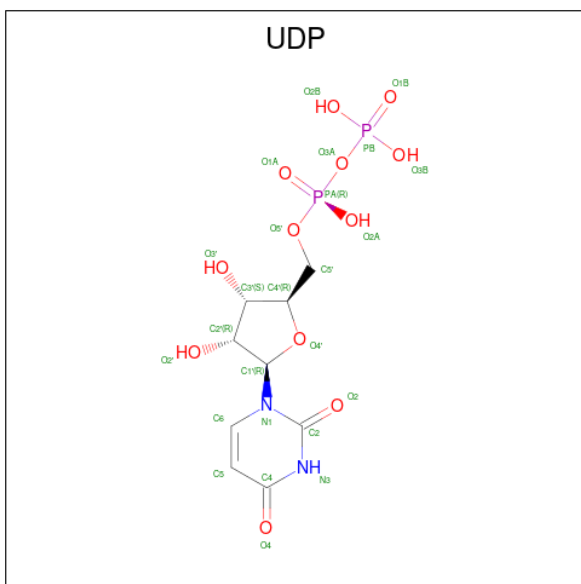
- Molecule 1 is a protein called Glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	279	Total 2284	C 1463	N 390	O 417	S 14	0	1	0
1	B	281	Total 2294	C 1470	N 391	O 419	S 14	0	0	0

- Molecule 2 is a protein called Mucin-1.

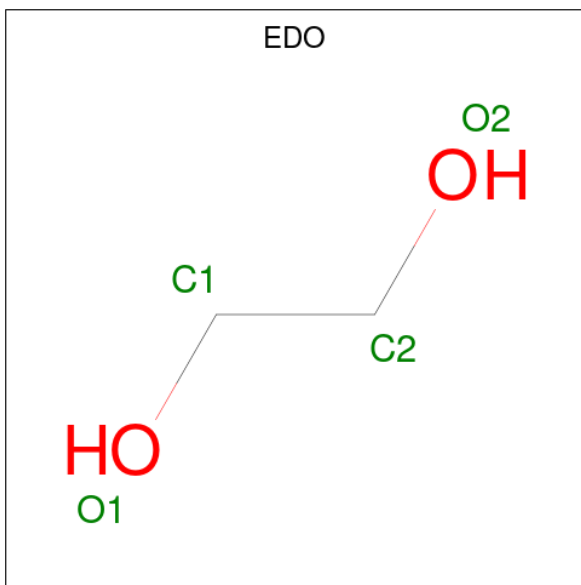
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	G	6	Total 43	C 25	N 9	O 9	0	0	1
2	F	6	Total 43	C 25	N 9	O 9	0	0	1

- Molecule 3 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: C₉H₁₄N₂O₁₂P₂).



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

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).

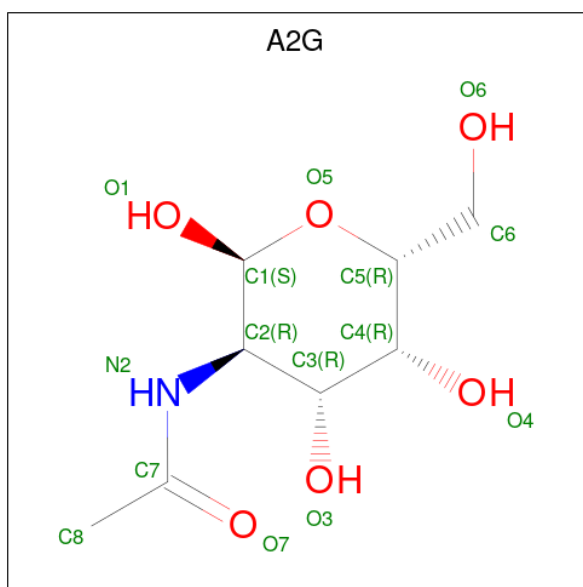


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	C O	0	0
			4	2 2		
4	B	1	Total	C O	0	0
			4	2 2		
4	F	1	Total	C O	0	0
			4	2 2		

- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mn	0	0
			1	1		
5	B	1	Total	Mn	0	0
			1	1		

- Molecule 6 is 2-acetamido-2-deoxy-alpha-D-galactopyranose (three-letter code: A2G) (formula: C₈H₁₅NO₆) (labeled as "Ligand of Interest" by depositor).



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

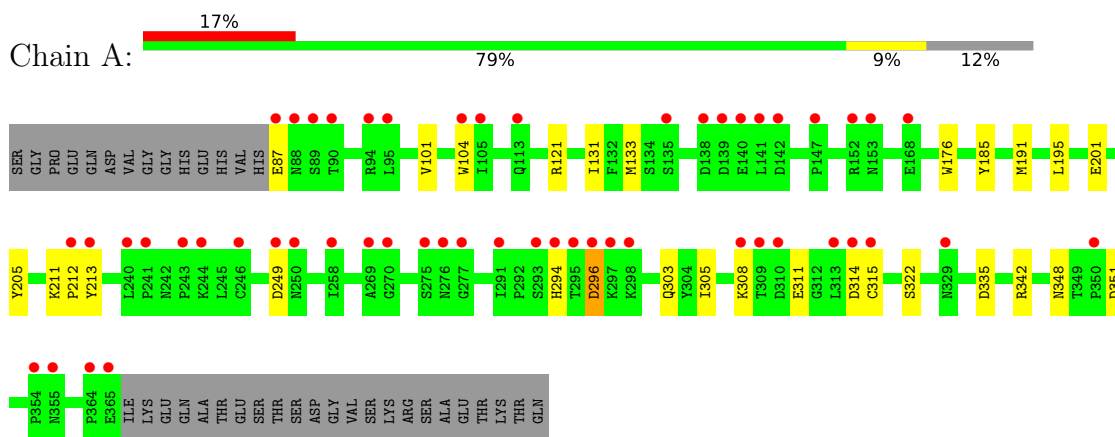
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	34	Total	O	0	0
			34	34		
7	B	37	Total	O	0	0
			37	37		
7	F	3	Total	O	0	0
			3	3		

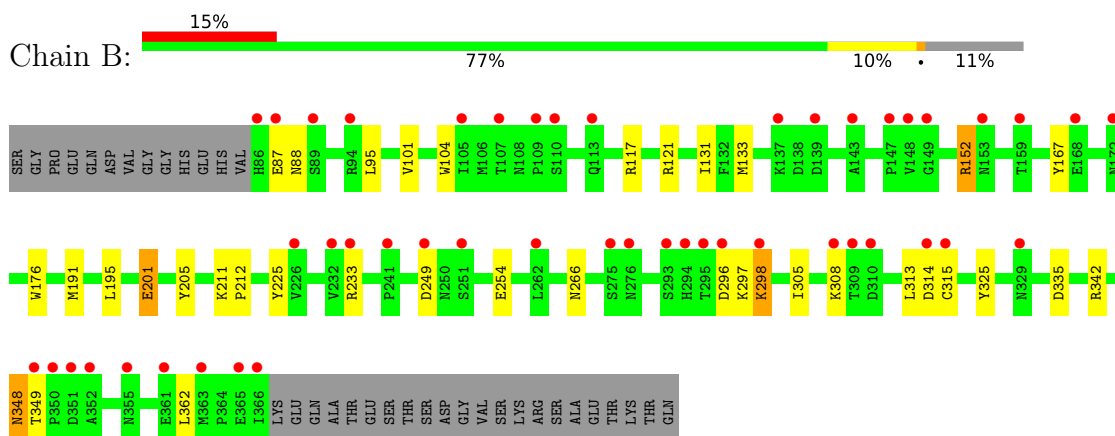
3 Residue-property plots [i](#)

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.

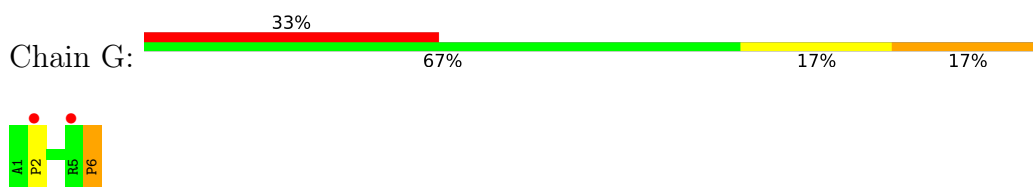
- Molecule 1: Glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1



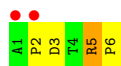
- Molecule 1: Glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1



- Molecule 2: Mucin-1



- Molecule 2: Mucin-1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	50.24Å 80.44Å 71.95Å 90.00° 93.19° 90.00°	Depositor
Resolution (Å)	20.00 – 2.40 19.99 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-2.40) 100.0 (19.99-2.40)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.59 (at 2.41Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.175 , 0.235 0.183 , 0.238	Depositor DCC
R_{free} test set	842 reflections (3.75%)	wwPDB-VP
Wilson B-factor (Å ²)	30.8	Xtriage
Anisotropy	0.141	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4830	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.91% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: UDP, EDO, MN, A2G, LPD

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	0.70	0/2354	0.88	0/3192
1	B	0.67	0/2362	0.86	0/3204
2	F	1.61	0/35	2.61	2/46 (4.3%)
2	G	1.40	1/35 (2.9%)	1.77	0/46
All	All	0.70	1/4786 (0.0%)	0.90	2/6488 (0.0%)

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
1	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	2	PRO	CA-C	-5.03	1.42	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	5	ARG	NE-CZ-NH2	12.10	126.35	120.30
2	F	5	ARG	NE-CZ-NH1	-6.16	117.22	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	315	CYS	Peptide
1	B	315	CYS	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	2284	0	2202	10	0
1	B	2294	0	2207	24	0
2	F	43	0	39	2	0
2	G	43	0	39	1	0
3	A	25	0	11	0	0
3	B	25	0	11	1	0
4	A	4	0	6	0	0
4	B	4	0	5	0	0
4	F	4	0	6	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
6	F	14	0	12	0	0
6	G	14	0	12	0	0
7	A	34	0	0	0	0
7	B	37	0	0	0	0
7	F	3	0	0	0	0
All	All	4830	0	4550	33	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 4.

All (33) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:3:ASP:OD2	2:F:5:ARG:HD3	1.87	0.74
1:B:201:GLU:OE1	1:B:233:ARG:NH2	2.23	0.72
1:B:348:ASN:HD22	1:B:348:ASN:N	1.87	0.70
1:A:213:TYR:OH	2:F:2:PRO:HD2	1.96	0.66
1:B:152:ARG:CG	1:B:152:ARG:HH11	2.09	0.66
1:B:152:ARG:HH11	1:B:152:ARG:HG2	1.63	0.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:101:VAL:HG11	1:B:191:MET:CE	2.31	0.60
1:B:152:ARG:CG	1:B:152:ARG:NH1	2.65	0.60
1:A:348:ASN:OD1	1:B:348:ASN:ND2	2.35	0.59
1:A:101:VAL:HG11	1:A:191:MET:CE	2.36	0.56
1:B:348:ASN:HD22	1:B:348:ASN:H	1.55	0.55
1:B:325:TYR:CE2	2:G:6:LPD:HD2	2.42	0.54
1:A:211:LYS:N	1:A:212:PRO:CD	2.74	0.51
1:B:117:ARG:HD2	1:B:121:ARG:NH2	2.27	0.49
1:B:348:ASN:N	1:B:348:ASN:ND2	2.58	0.48
1:B:101:VAL:HG11	1:B:191:MET:HE2	1.94	0.48
1:B:211:LYS:N	1:B:212:PRO:CD	2.77	0.47
1:B:297:LYS:O	1:B:298:LYS:HB2	2.17	0.44
1:B:104:TRP:HA	1:B:131:ILE:O	2.18	0.44
1:B:176:TRP:CG	1:B:195:LEU:HD13	2.52	0.44
1:A:176:TRP:CG	1:A:195:LEU:HD13	2.52	0.44
1:B:152:ARG:NH2	3:B:401:UDP:O2B	2.45	0.44
1:B:225:TYR:OH	1:B:254:GLU:OE1	2.25	0.43
1:A:101:VAL:HG11	1:A:191:MET:HE2	1.99	0.43
1:B:348:ASN:ND2	1:B:348:ASN:H	2.15	0.43
1:B:104:TRP:CE3	1:B:133:MET:HG3	2.53	0.42
1:B:152:ARG:NH1	1:B:152:ARG:HG3	2.34	0.42
1:A:104:TRP:HA	1:A:131:ILE:O	2.19	0.42
1:A:104:TRP:CE3	1:A:133:MET:HG3	2.55	0.41
1:B:167:TYR:HE1	1:B:233:ARG:HG2	1.84	0.41
1:A:185:TYR:O	1:A:322:SER:HA	2.20	0.40
1:A:87:GLU:HB3	1:B:349:THR:OG1	2.22	0.40
1:B:95:LEU:HD23	1:B:95:LEU:HA	1.88	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	278/316 (88%)	263 (95%)	13 (5%)	2 (1%)	22	32
1	B	279/316 (88%)	264 (95%)	14 (5%)	1 (0%)	34	48
2	F	4/6 (67%)	4 (100%)	0	0	100	100
2	G	4/6 (67%)	4 (100%)	0	0	100	100
All	All	565/644 (88%)	535 (95%)	27 (5%)	3 (0%)	29	41

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	296	ASP
1	A	294	HIS
1	B	296	ASP

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	247/277 (89%)	234 (95%)	13 (5%)	22	37
1	B	248/277 (90%)	232 (94%)	16 (6%)	17	27
2	F	4/4 (100%)	4 (100%)	0	100	100
2	G	4/4 (100%)	4 (100%)	0	100	100
All	All	503/562 (90%)	474 (94%)	29 (6%)	20	32

All (29) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	121	ARG
1	A	201	GLU
1	A	205	TYR
1	A	249	ASP
1	A	296	ASP
1	A	303	GLN
1	A	305	ILE
1	A	308	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	311	GLU
1	A	314	ASP
1	A	335	ASP
1	A	342	ARG
1	A	351	ASP
1	B	87	GLU
1	B	88	ASN
1	B	152	ARG
1	B	201	GLU
1	B	205	TYR
1	B	249	ASP
1	B	266	ASN
1	B	298	LYS
1	B	305	ILE
1	B	308	LYS
1	B	313	LEU
1	B	314	ASP
1	B	335	ASP
1	B	342	ARG
1	B	348	ASN
1	B	362	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	88	ASN
1	B	348	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](#)

2 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
2	LPD	G	6	2	8,8,8	2.28	1 (12%)	10,10,10	1.33	2 (20%)
2	LPD	F	6	2	8,8,8	1.53	1 (12%)	10,10,10	1.41	2 (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	LPD	G	6	2	-	2/4/11/11	0/1/1/1
2	LPD	F	6	2	-	2/4/11/11	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	6	LPD	CA-C	-6.28	1.40	1.52
2	F	6	LPD	CA-C	-3.79	1.45	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	6	LPD	O-C-N2	-3.35	117.17	123.00
2	G	6	LPD	C-CA-N	2.25	116.08	111.88
2	G	6	LPD	CG-CD-N	-2.20	99.26	105.69
2	F	6	LPD	O-C-CA	2.11	123.36	120.27

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	G	6	LPD	O-C-CA-N
2	G	6	LPD	N2-C-CA-N
2	F	6	LPD	O-C-CA-N
2	F	6	LPD	N2-C-CA-N

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	G	6	LPD	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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
3	UDP	A	401	5	20,26,26	0.95	0	25,40,40	1.33	3 (12%)
6	A2G	G	101	2	14,14,15	1.54	2 (14%)	17,19,21	0.86	0
6	A2G	F	102	2	14,14,15	1.30	1 (7%)	17,19,21	1.21	1 (5%)
4	EDO	F	101	-	3,3,3	0.21	0	2,2,2	0.27	0
3	UDP	B	401	5	20,26,26	1.08	1 (5%)	25,40,40	1.19	3 (12%)
4	EDO	B	402	-	3,3,3	0.79	0	2,2,2	0.21	0
4	EDO	A	402	-	3,3,3	0.59	0	2,2,2	0.96	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	UDP	A	401	5	-	1/14/32/32	0/2/2/2
6	A2G	G	101	2	-	0/6/23/26	0/1/1/1
6	A2G	F	102	2	-	0/6/23/26	0/1/1/1
4	EDO	F	101	-	-	1/1/1/1	-
3	UDP	B	401	5	-	3/14/32/32	0/2/2/2
4	EDO	B	402	-	-	0/1/1/1	-
4	EDO	A	402	-	-	1/1/1/1	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	G	101	A2G	O4-C4	4.68	1.54	1.43
6	F	102	A2G	C8-C7	2.86	1.56	1.50
3	B	401	UDP	O4'-C4'	-2.08	1.40	1.45
6	G	101	A2G	O3-C3	2.02	1.47	1.43

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	401	UDP	O2B-PB-O3A	2.67	113.59	104.64
3	A	401	UDP	O3A-PB-O1B	-2.56	97.01	111.19
6	F	102	A2G	C1-O5-C5	2.54	115.63	112.19
3	B	401	UDP	C5-C4-N3	-2.30	118.25	123.31
3	B	401	UDP	C6-N1-C2	-2.14	117.81	121.20
3	B	401	UDP	O3B-PB-O2B	2.12	115.73	107.64
3	A	401	UDP	O4'-C1'-C2'	-2.04	103.94	106.93

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	401	UDP	C2'-C1'-N1-C6
3	B	401	UDP	O4'-C1'-N1-C6
3	A	401	UDP	C4'-C5'-O5'-PA
4	F	101	EDO	O1-C1-C2-O2
3	B	401	UDP	C3'-C4'-C5'-O5'
4	A	402	EDO	O1-C1-C2-O2

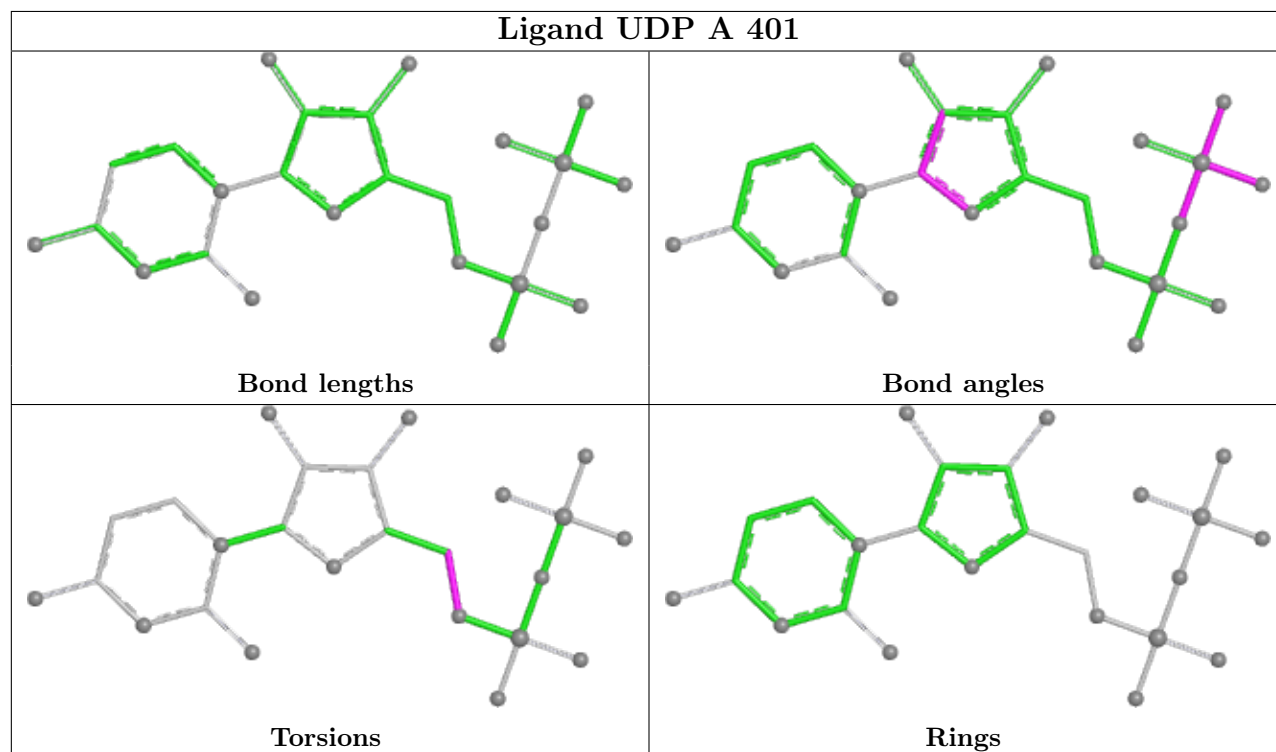
There are no ring outliers.

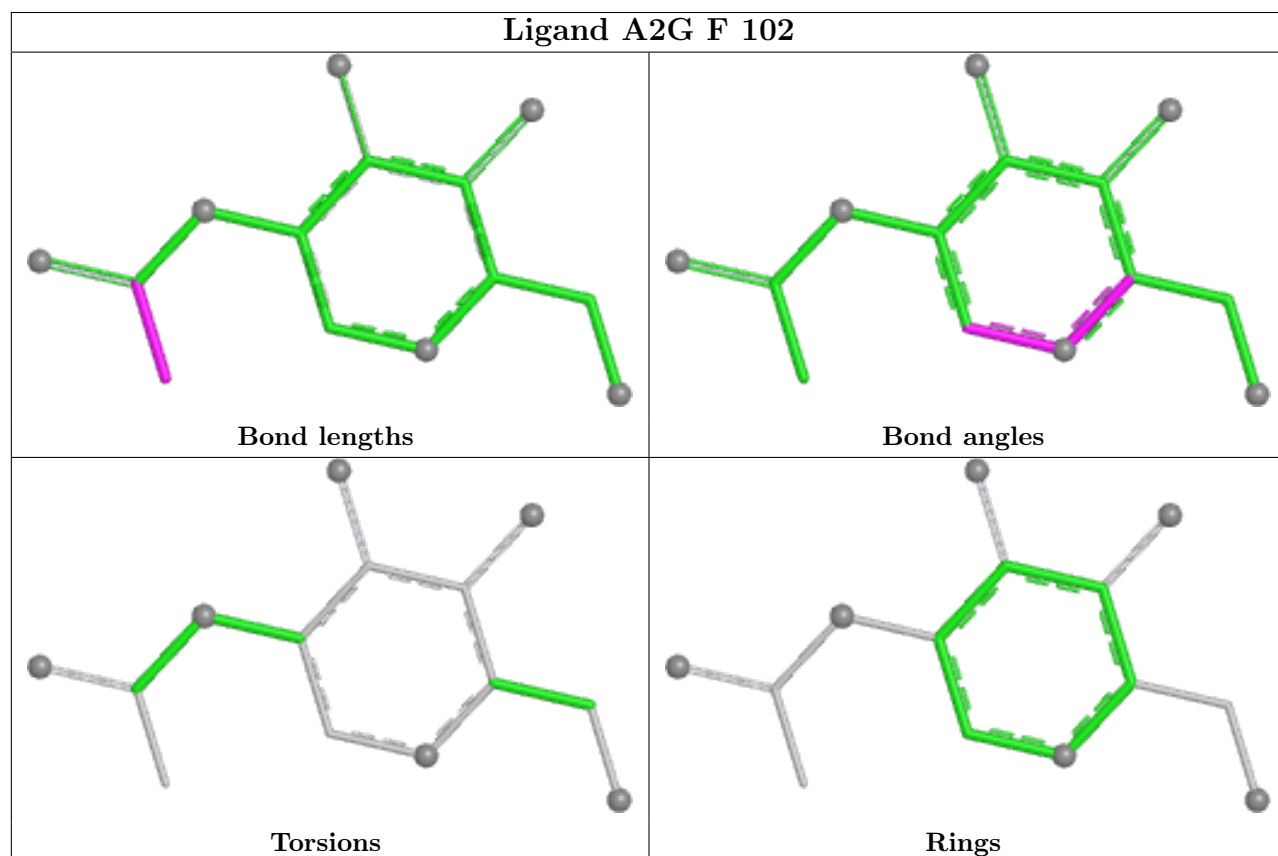
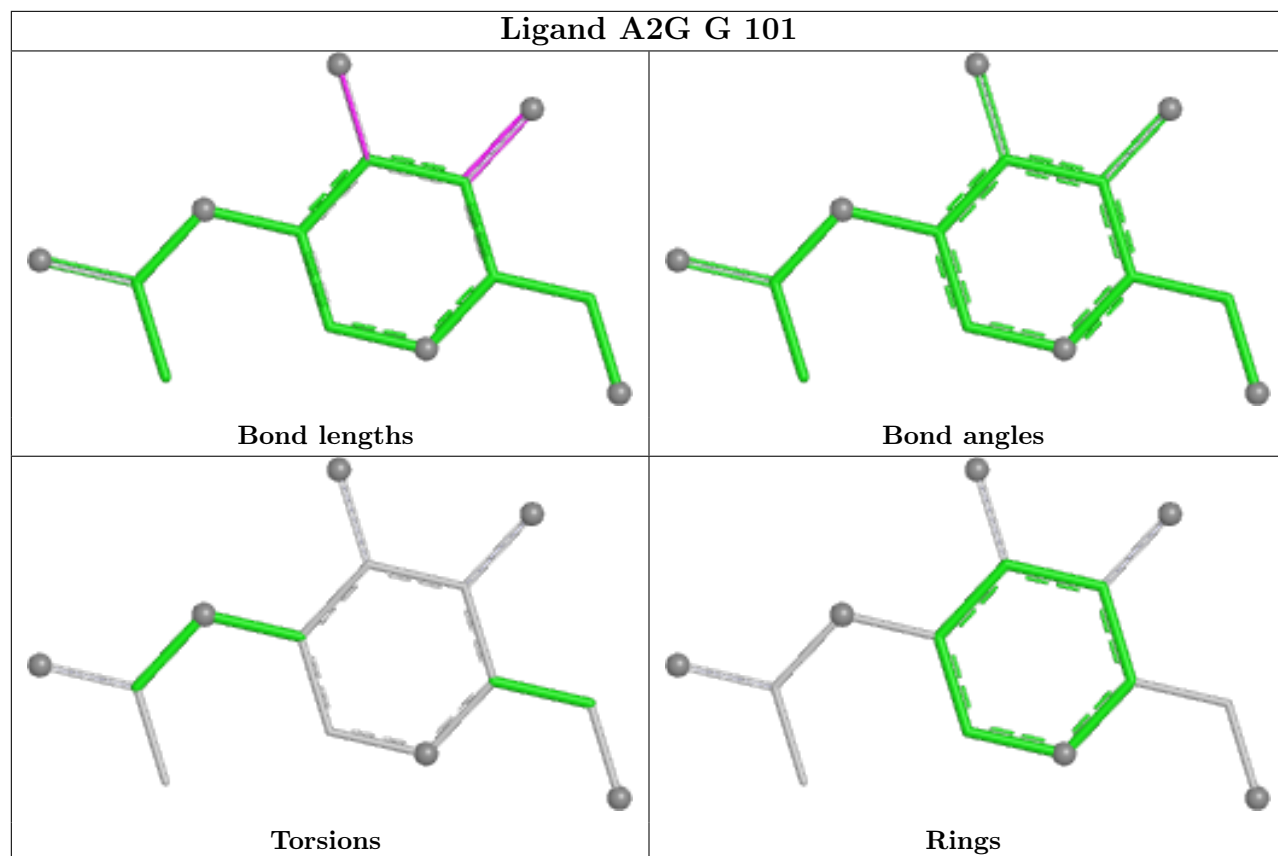
1 monomer is involved in 1 short contact:

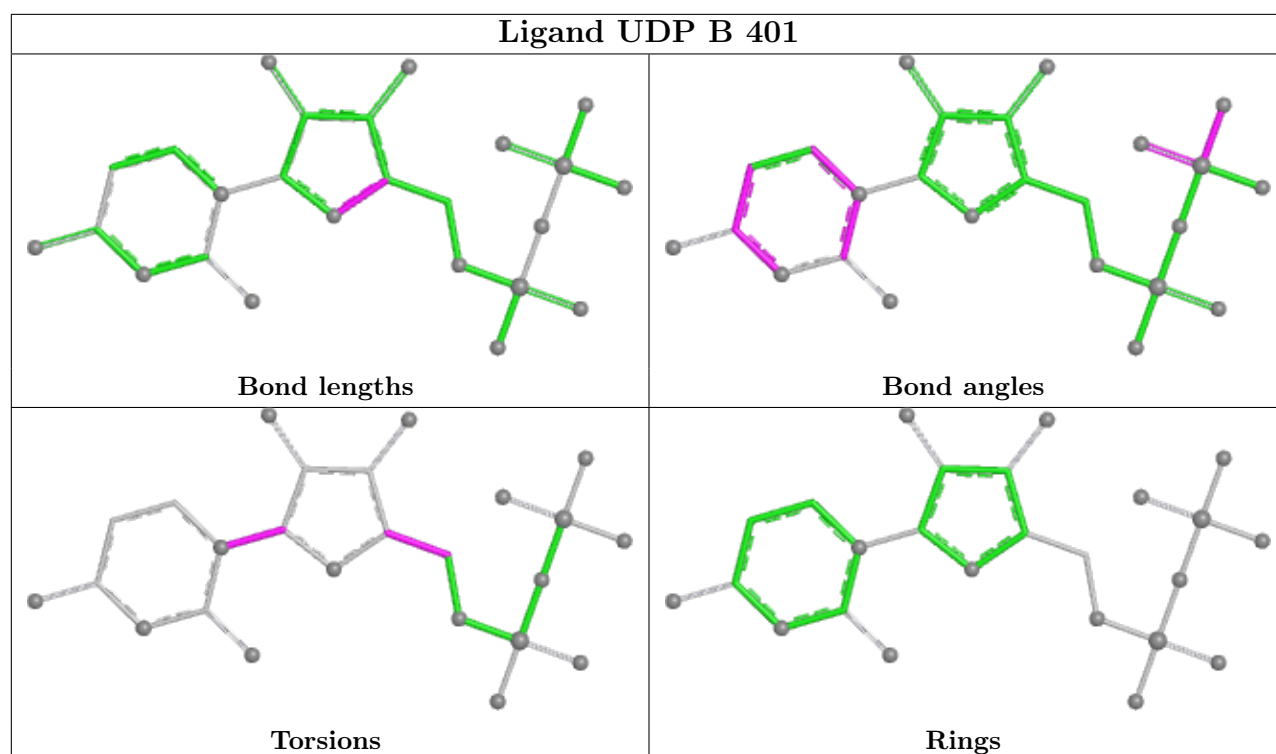
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	401	UDP	1	0

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 than 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.







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	279/316 (88%)	1.28	53 (18%) 1 1	18, 30, 55, 107	0
1	B	281/316 (88%)	1.16	48 (17%) 1 1	20, 30, 59, 95	0
2	F	5/6 (83%)	1.81	2 (40%) 0 0	24, 30, 31, 34	0
2	G	5/6 (83%)	2.95	2 (40%) 0 0	42, 61, 70, 79	0
All	All	570/644 (88%)	1.24	105 (18%) 1 1	18, 30, 60, 107	0

All (105) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	296	ASP	8.8
1	B	298	LYS	8.1
1	A	295	THR	7.8
1	B	295	THR	6.8
1	B	86	HIS	6.4
1	B	139	ASP	6.3
2	G	2	PRO	5.9
1	B	153	ASN	5.8
1	A	139	ASP	5.3
2	G	5	ARG	5.1
1	A	314	ASP	5.1
1	B	351	ASP	5.0
1	A	355	ASN	4.8
1	B	87	GLU	4.5
1	B	294	HIS	4.4
1	A	249	ASP	4.3
1	A	141	LEU	4.3
1	A	87	GLU	4.3
1	A	276	ASN	4.2
1	A	89	SER	4.1
2	F	1	ALA	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	90	THR	3.9
1	A	298	LYS	3.9
1	B	366	ILE	3.8
1	A	308	LYS	3.8
1	A	244	LYS	3.8
1	A	147	PRO	3.7
1	B	293	SER	3.7
1	B	168	GLU	3.6
1	A	297	LYS	3.5
1	B	94	ARG	3.5
1	B	249	ASP	3.5
1	A	294	HIS	3.4
1	A	140	GLU	3.4
1	B	143	ALA	3.4
1	B	365	GLU	3.3
1	A	142	ASP	3.3
1	A	94	ARG	3.2
1	A	275	SER	3.2
1	A	310	ASP	3.2
1	B	314	ASP	3.2
1	B	361	GLU	3.1
1	A	250	ASN	3.1
1	B	296	ASP	3.0
1	A	309	THR	3.0
1	A	113	GLN	3.0
1	B	352	ALA	3.0
1	A	105	ILE	2.9
1	B	105	ILE	2.9
1	A	153	ASN	2.9
1	B	310	ASP	2.9
1	B	147	PRO	2.9
1	A	168	GLU	2.9
1	A	213	TYR	2.8
1	A	258	ILE	2.8
1	A	243	PRO	2.8
1	B	308	LYS	2.8
1	A	277	GLY	2.8
1	B	276	ASN	2.8
1	A	138	ASP	2.7
1	B	110	SER	2.7
1	B	349	THR	2.7
1	B	251	SER	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	149	GLY	2.7
1	A	269	ALA	2.6
1	A	88	ASN	2.6
1	B	241	PRO	2.6
1	B	309	THR	2.6
1	A	212	PRO	2.6
1	B	109	PRO	2.5
1	A	246	CYS	2.5
1	A	365	GLU	2.5
1	A	315	CYS	2.4
1	A	241	PRO	2.4
2	F	2	PRO	2.4
1	A	329	ASN	2.4
1	A	354	PRO	2.4
1	A	104	TRP	2.4
1	B	275	SER	2.4
1	B	113	GLN	2.4
1	B	355	ASN	2.3
1	B	137	LYS	2.3
1	B	233	ARG	2.3
1	B	89	SER	2.3
1	A	95	LEU	2.3
1	B	148	VAL	2.3
1	A	291	ILE	2.3
1	A	270	GLY	2.3
1	A	240	LEU	2.3
1	A	313	LEU	2.3
1	A	293	SER	2.3
1	A	350	PRO	2.2
1	B	315	CYS	2.2
1	B	329	ASN	2.2
1	B	226	VAL	2.2
1	A	135	SER	2.2
1	A	364	PRO	2.2
1	B	232	VAL	2.1
1	B	172	ASN	2.1
1	B	363	MET	2.1
1	B	350	PRO	2.1
1	A	152[A]	ARG	2.1
1	B	262	LEU	2.1
1	B	107	THR	2.1
1	B	159	THR	2.0

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
2	LPD	F	6	8/8	0.81	0.25	24,25,27,29	0
2	LPD	G	6	8/8	0.84	0.19	53,57,58,60	0

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

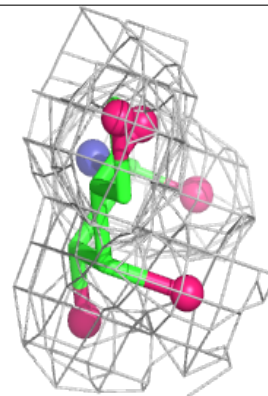
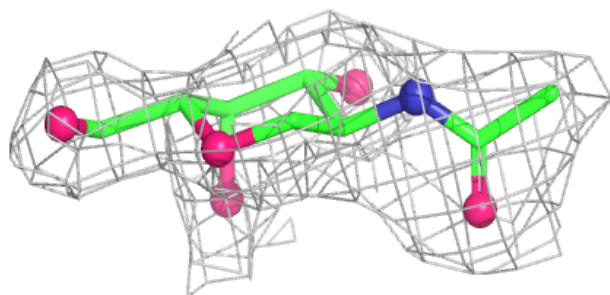
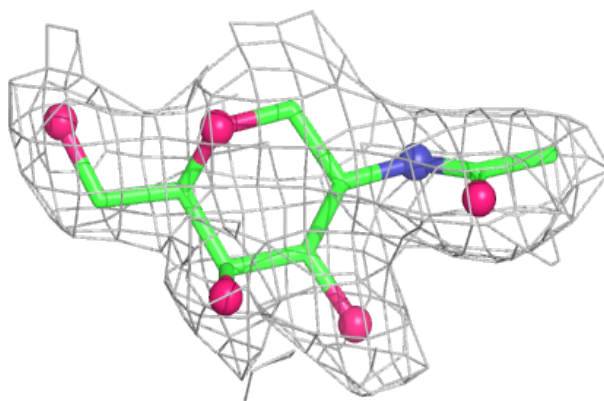
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	EDO	F	101	4/4	0.61	0.34	34,35,40,41	0
6	A2G	G	101	14/15	0.83	0.18	20,25,29,29	0
6	A2G	F	102	14/15	0.87	0.18	20,23,24,24	0
4	EDO	B	402	4/4	0.90	0.31	23,30,35,37	0
3	UDP	B	401	25/25	0.90	0.20	22,29,35,36	0
4	EDO	A	402	4/4	0.91	0.32	32,33,35,35	0
3	UDP	A	401	25/25	0.93	0.16	23,28,32,39	0
5	MN	A	403	1/1	0.99	0.08	22,22,22,22	0
5	MN	B	403	1/1	0.99	0.08	26,26,26,26	0

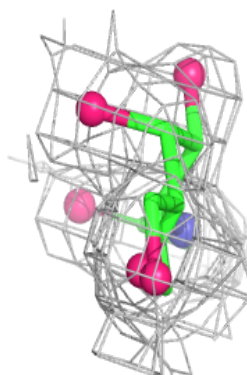
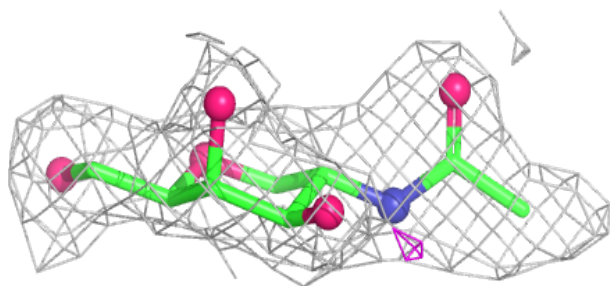
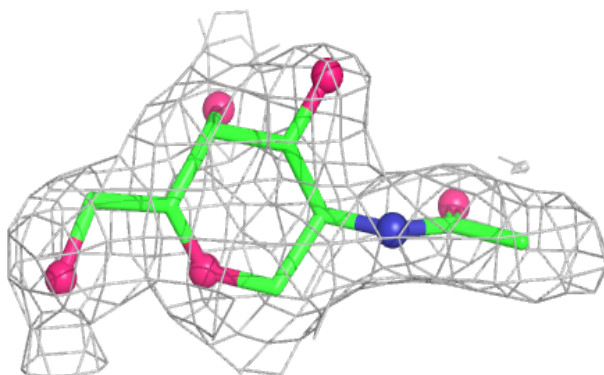
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around A2G G 101:

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

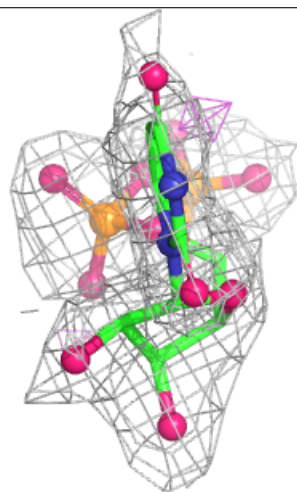
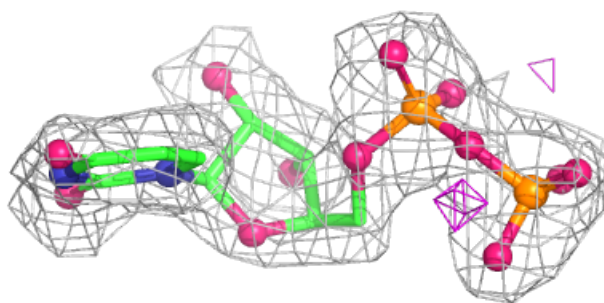
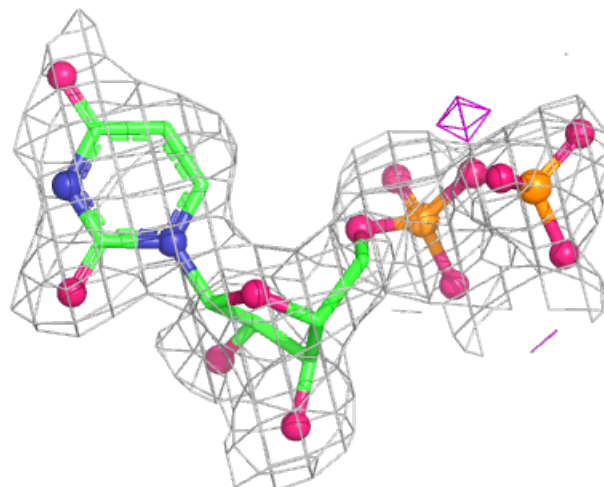
**Electron density around A2G F 102:**

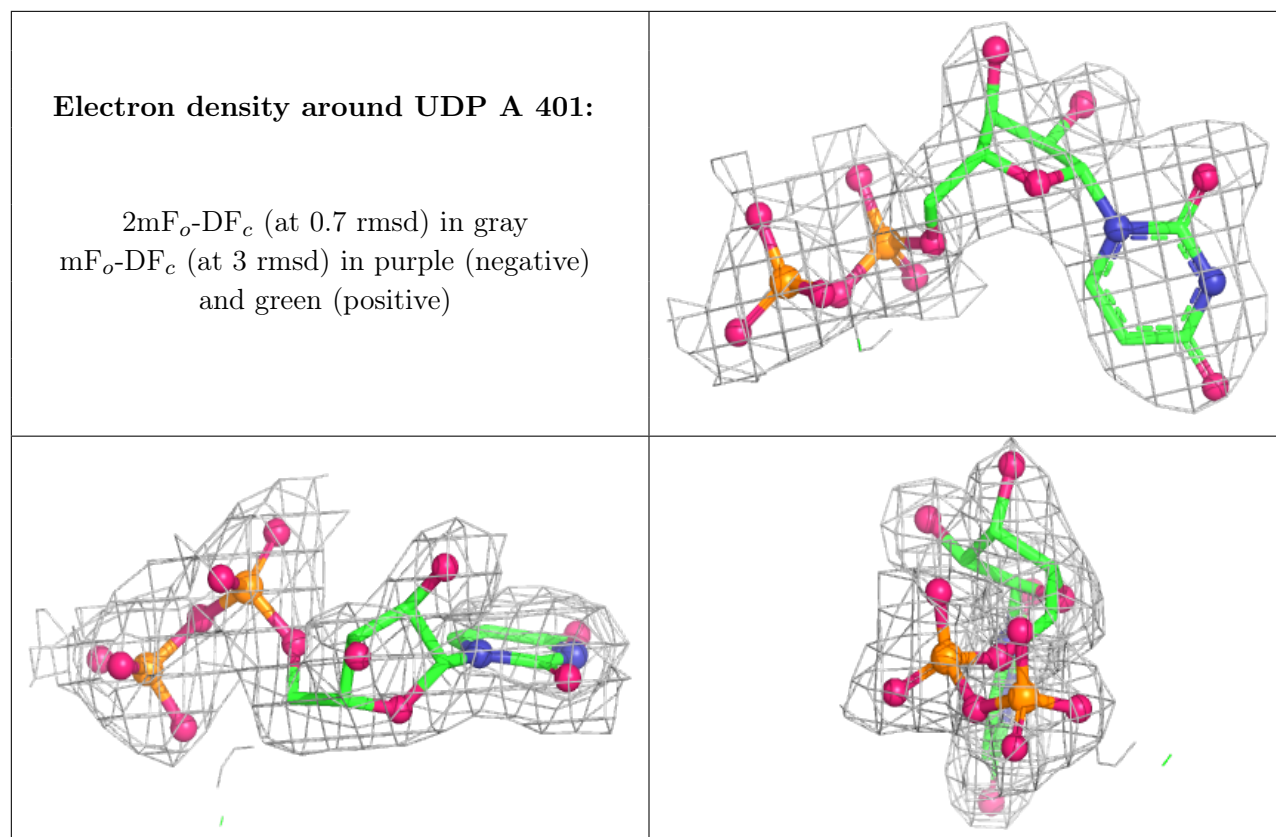
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around UDP B 401:

$2mF_o-DF_c$ (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.