



wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 07:59 pm BST

PDB ID : 5FV9
Title : Crystal structure of GalNAc-T2 in complex with compound 16d
Authors : Ghirardello, M.; Rivas, M.; Lacetera, A.; Delso, I.; Lira-Navarrete, E.; Tejero, T.; Martin-Santamaria, S.; Hurtado-Guerrero, R.; Merino, P.
Deposited on : 2016-02-03
Resolution : 2.07 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

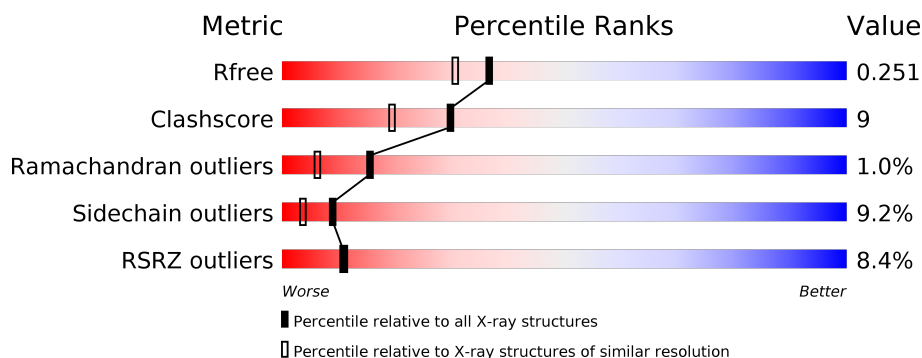
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.07 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	2684 (2.08-2.04)
Clashscore	141614	2801 (2.08-2.04)
Ramachandran outliers	138981	2768 (2.08-2.04)
Sidechain outliers	138945	2768 (2.08-2.04)
RSRZ outliers	127900	2646 (2.08-2.04)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	571	<div> <div>5%</div> <div> <div>66%</div> <div>13%</div> <div>• •</div> <div>15%</div> </div> </div>
1	B	571	<div> <div>5%</div> <div> <div>67%</div> <div>14%</div> <div>•</div> <div>15%</div> </div> </div>
1	C	571	<div> <div>7%</div> <div> <div>66%</div> <div>14%</div> <div>•</div> <div>16%</div> </div> </div>
1	D	571	<div> <div>11%</div> <div> <div>58%</div> <div>16%</div> <div>• •</div> <div>22%</div> </div> </div>
1	E	571	<div> <div>5%</div> <div> <div>67%</div> <div>14%</div> <div>•</div> <div>15%</div> </div> </div>
1	F	571	<div> <div>9%</div> <div> <div>63%</div> <div>17%</div> <div>• •</div> <div>16%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	UDP	B	1572	-	-	X	-
4	EDO	A	1573	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 24993 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.

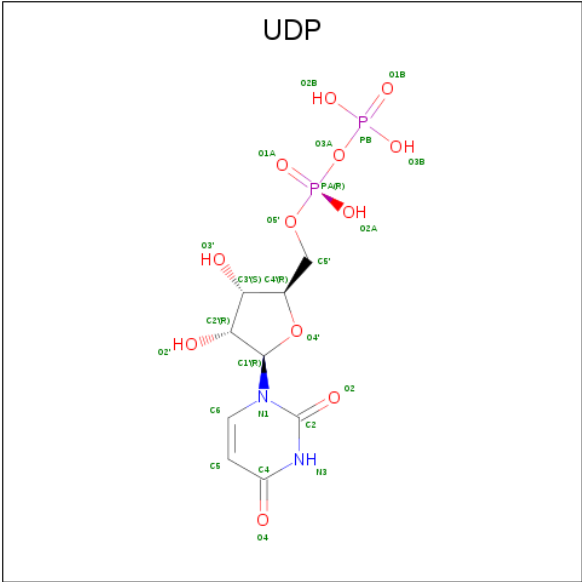
- Molecule 1 is a protein called GALNAC-T2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	483	Total	C	N	O	S	20	2	0
			3896	2451	709	712	24			
1	B	484	Total	C	N	O	S	20	1	0
			3894	2450	709	711	24			
1	C	482	Total	C	N	O	S	20	3	0
			3897	2451	712	710	24			
1	D	443	Total	C	N	O	S	20	0	0
			3575	2256	645	652	22			
1	E	483	Total	C	N	O	S	20	1	0
			3885	2445	705	711	24			
1	F	481	Total	C	N	O	S	20	1	0
			3869	2432	706	707	24			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	516	ASP	ASN	engineered mutation	UNP Q10471
B	516	ASP	ASN	engineered mutation	UNP Q10471
C	516	ASP	ASN	engineered mutation	UNP Q10471
D	516	ASP	ASN	engineered mutation	UNP Q10471
E	516	ASP	ASN	engineered mutation	UNP Q10471
F	516	ASP	ASN	engineered mutation	UNP Q10471

- Molecule 2 is URIDINE-5'-DIPHOSPHATE (three-letter code: UDP) (formula: C₉H₁₄N₂O₁₂P₂).

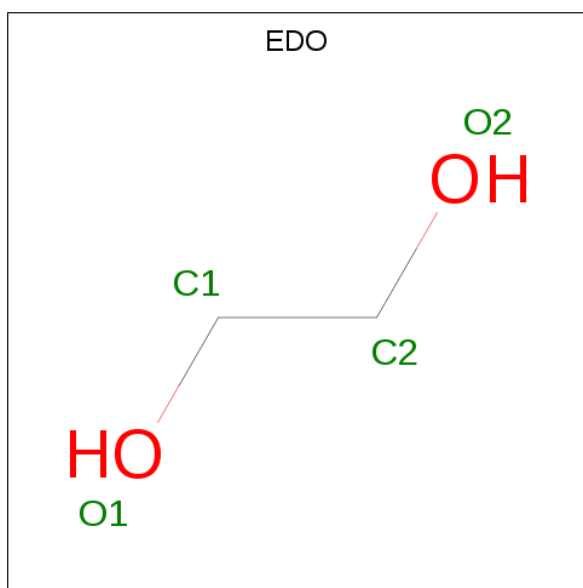


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	B	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	C	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	D	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	E	1	Total	C	N	O	P	0	0
			25	9	2	12	2		
2	F	1	Total	C	N	O	P	0	0
			25	9	2	12	2		

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	1	Total	Mn	0	0
			1	1		
3	E	1	Total	Mn	0	0
			1	1		
3	B	1	Total	Mn	0	0
			1	1		
3	C	1	Total	Mn	0	0
			1	1		
3	A	1	Total	Mn	0	0
			1	1		
3	F	1	Total	Mn	0	0
			1	1		

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



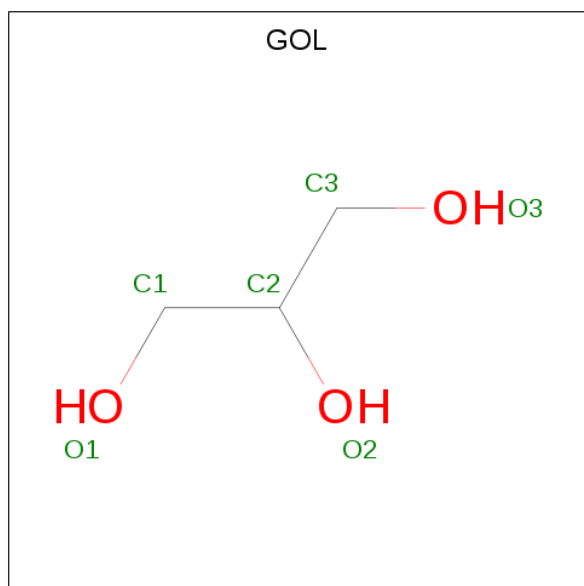
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	C	1	Total	C	O	0	0
			4	2	2		
4	C	1	Total	C	O	0	0
			4	2	2		
4	E	1	Total	C	O	0	0
			4	2	2		
4	E	1	Total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

Continued on next page...

Continued from previous page...

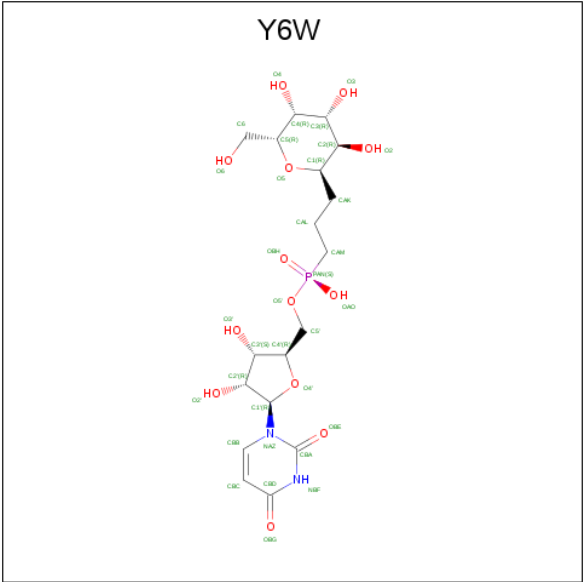
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	F	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 6 is [(2R,3S,4R,5R)-5-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-3,4-dihydroxytetrahydrofuran-2-yl]methyl hydrogen (S)-{3-[(2R,3R,4R,5R,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yl]propyl}phosphonate (three-letter code: Y6W) (formula: $C_{18}H_{29}N_2O_{13}P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	N	O	P	0	0
			34	18	2	13	1		
6	E	1	Total	C	N	O	P	0	0
			34	18	2	13	1		
6	F	1	Total	C	N	O	P	0	0
			34	18	2	13	1		

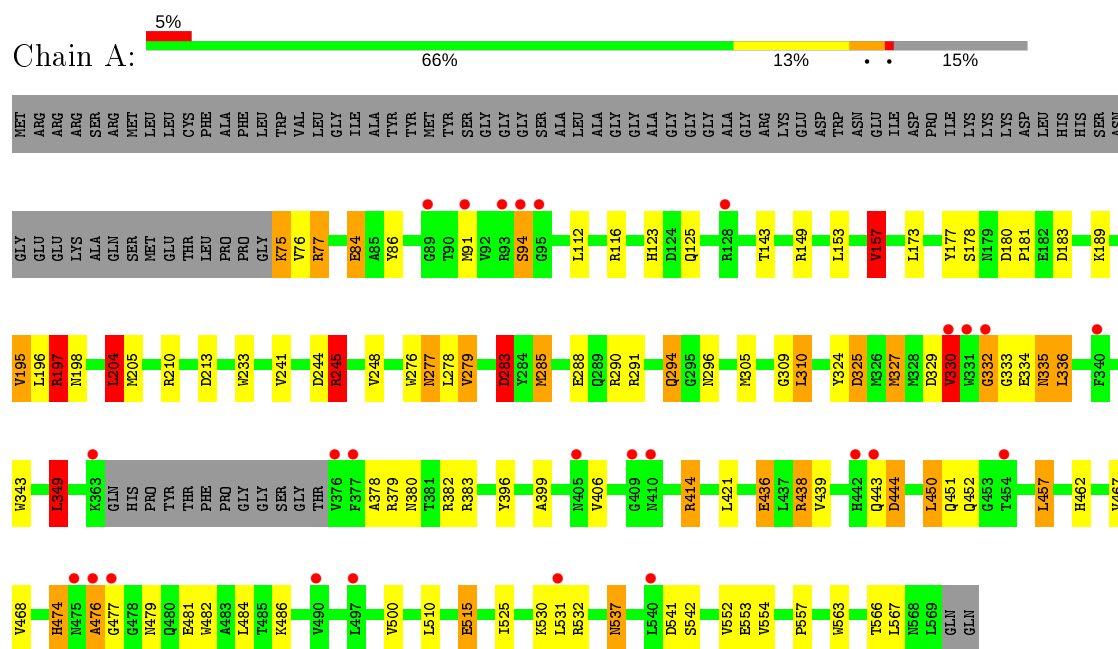
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	346	Total	O	0	0
			346	346		
7	B	299	Total	O	0	0
			299	299		
7	C	306	Total	O	0	0
			306	306		
7	D	184	Total	O	0	0
			184	184		
7	E	290	Total	O	0	0
			290	290		
7	F	208	Total	O	0	0
			208	208		

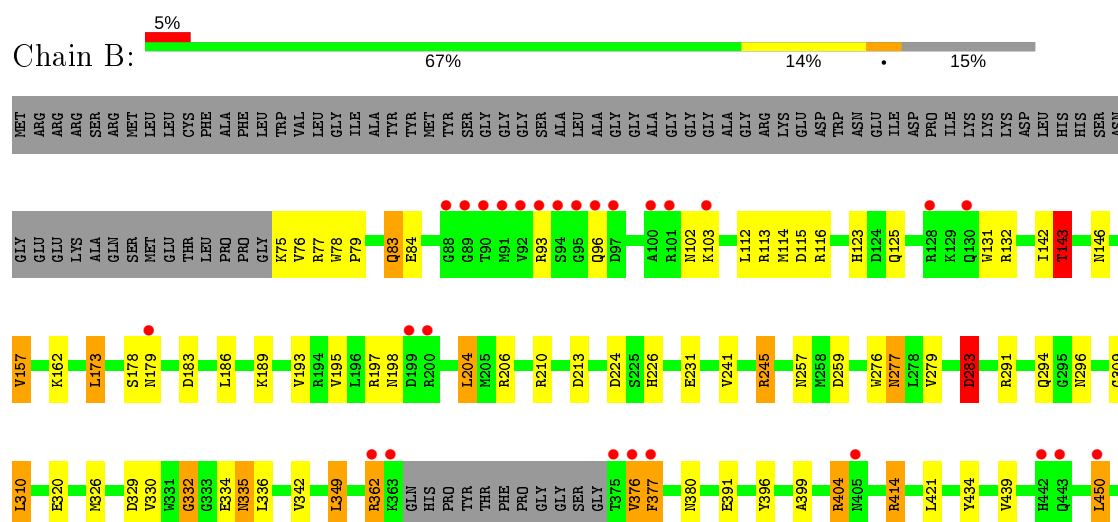
3 Residue-property plots [i](#)

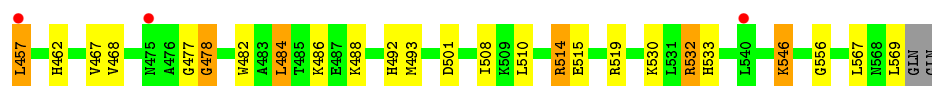
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: GALNAC-T2

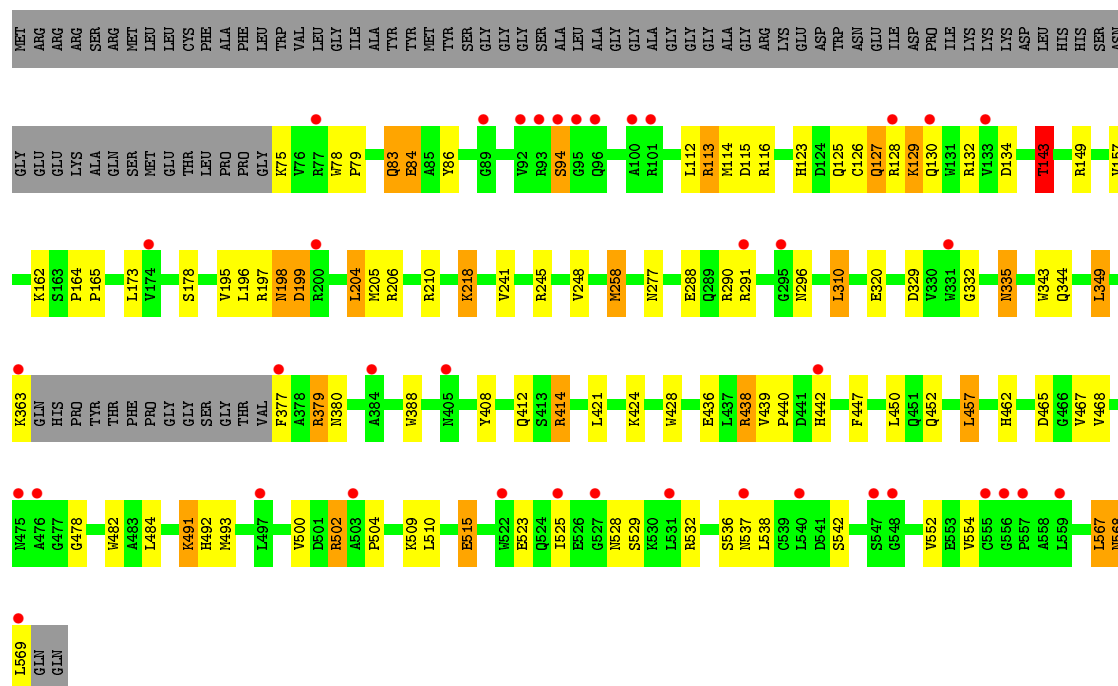


• Molecule 1: GALNAC-T2

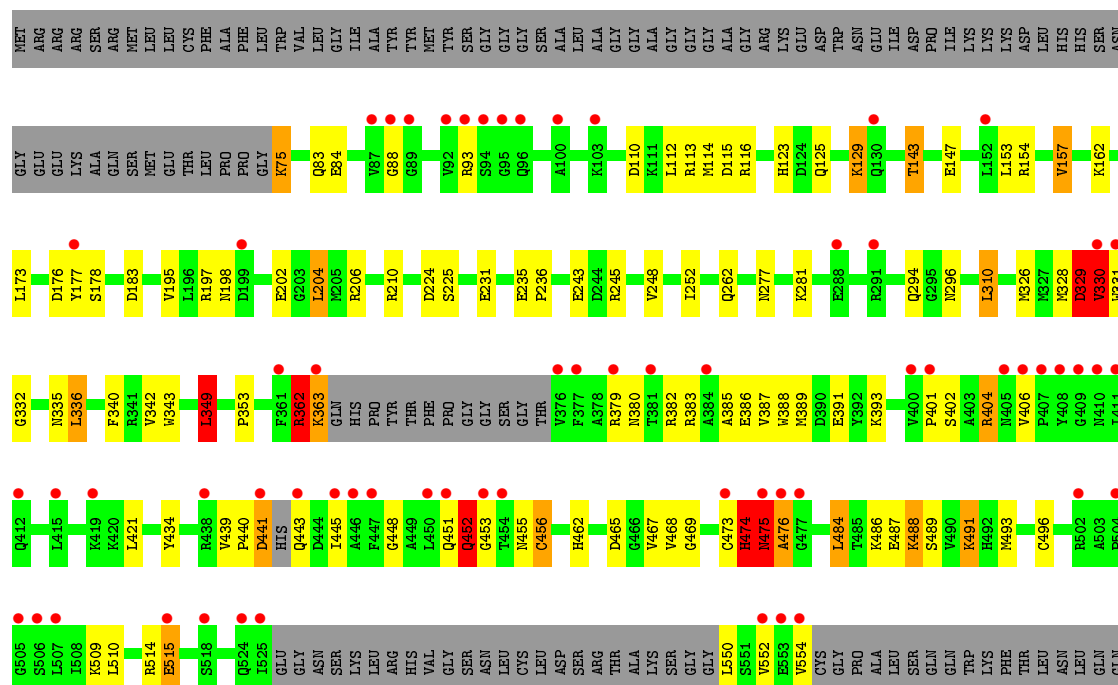




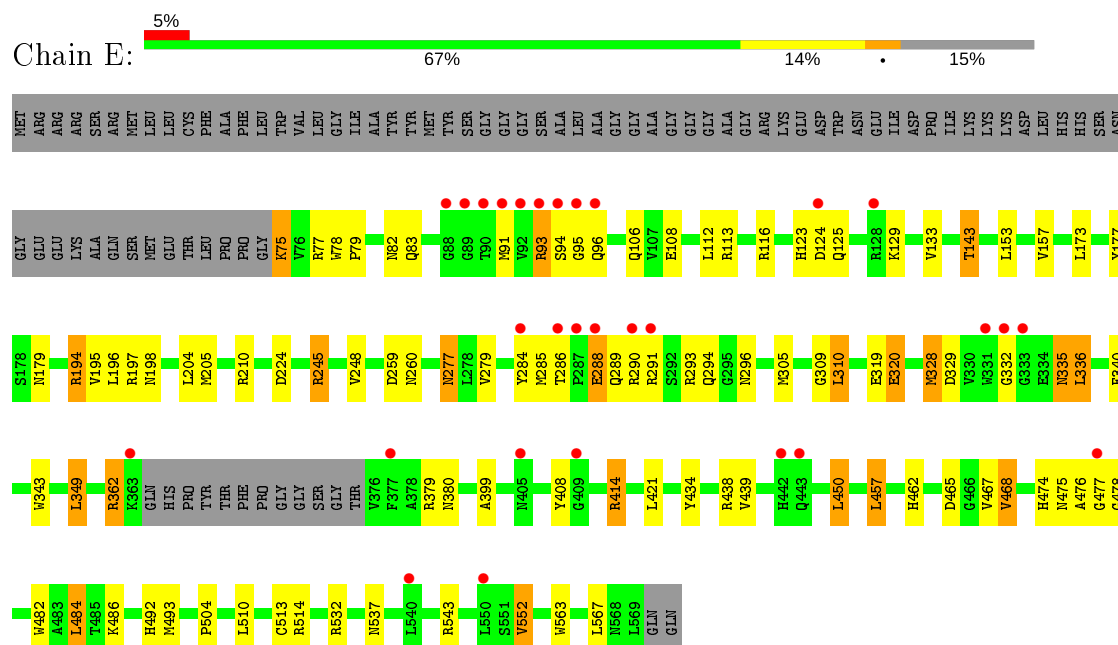
• Molecule 1: GALNAC-T2



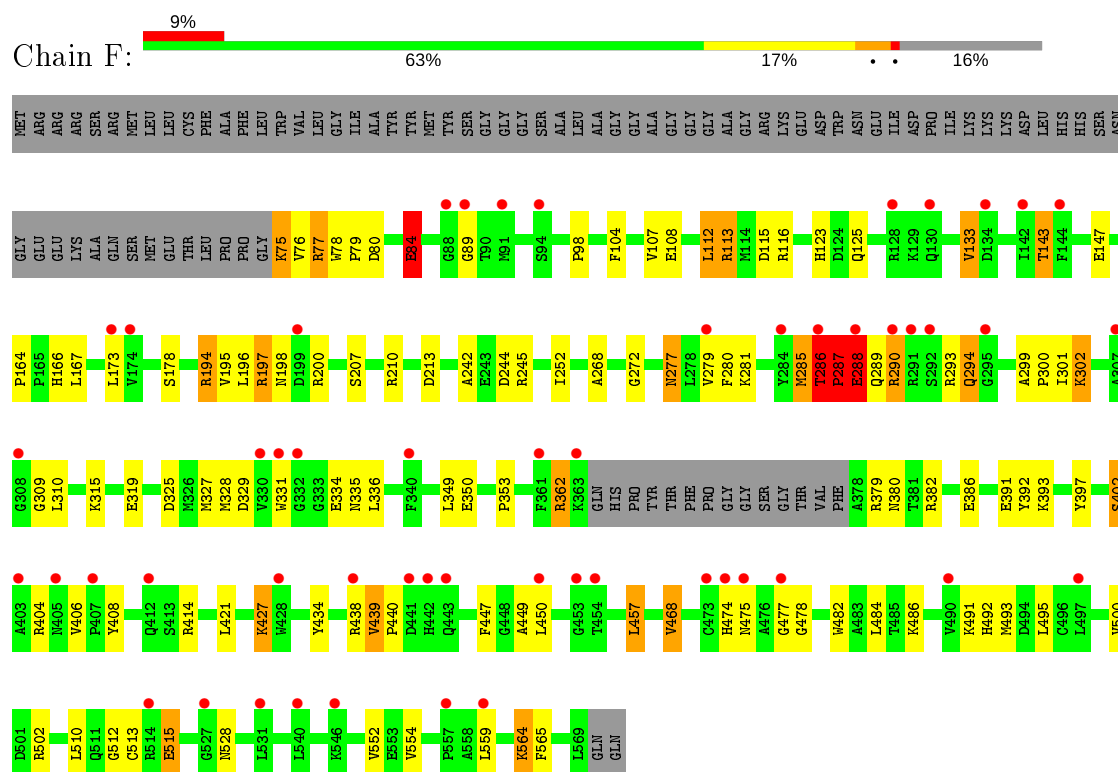
• Molecule 1: GALNAC-T2



• Molecule 1: GALNAC-T2



• Molecule 1: GALNAC-T2



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	116.49Å 121.75Å 250.16Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	250.15 – 2.07 19.99 – 2.07	Depositor EDS
% Data completeness (in resolution range)	99.8 (250.15-2.07) 99.9 (19.99-2.07)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.13 (at 2.07Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.190 , 0.247 0.198 , 0.251	Depositor DCC
R_{free} test set	5938 reflections (2.75%)	wwPDB-VP
Wilson B-factor (Å ²)	38.9	Xtriage
Anisotropy	0.213	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 48.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.013 for k,h,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	24993	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.06% 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: GOL, UDP, Y6W, EDO, MN

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	0.99	7/3982 (0.2%)	1.09	26/5380 (0.5%)
1	B	0.96	6/3980 (0.2%)	1.08	21/5378 (0.4%)
1	C	0.93	4/3986 (0.1%)	1.02	15/5384 (0.3%)
1	D	0.87	4/3653 (0.1%)	0.98	13/4935 (0.3%)
1	E	1.06	5/3971 (0.1%)	1.18	27/5366 (0.5%)
1	F	1.00	4/3954 (0.1%)	1.06	23/5342 (0.4%)
All	All	0.97	30/23526 (0.1%)	1.07	125/31785 (0.4%)

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	2
1	B	0	1
1	C	0	2
1	D	0	4
1	E	0	1
1	F	0	4
All	All	0	14

The worst 5 of 30 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	515	GLU	CB-CG	-31.88	0.91	1.52
1	E	486	LYS	CB-CG	-28.54	0.75	1.52
1	A	515	GLU	CB-CG	-18.04	1.17	1.52
1	C	515	GLU	CB-CG	-16.08	1.21	1.52
1	C	84	GLU	CB-CG	-14.73	1.24	1.52

The worst 5 of 125 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	197	ARG	NE-CZ-NH2	-23.21	108.69	120.30
1	E	197	ARG	NE-CZ-NH1	18.01	129.31	120.30
1	B	75	LYS	CA-CB-CG	13.77	143.68	113.40
1	F	75	LYS	CB-CG-CD	13.08	145.61	111.60
1	A	486	LYS	CB-CG-CD	-12.86	78.16	111.60

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	330	VAL	Peptide
1	A	476	ALA	Peptide
1	B	332	GLY	Peptide
1	C	129	LYS	Peptide
1	C	198	ASN	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	3896	0	3829	79	0
1	B	3894	0	3831	77	0
1	C	3897	0	3835	61	0
1	D	3575	0	3512	75	0
1	E	3885	0	3815	56	1
1	F	3869	0	3806	69	1
2	A	25	0	11	3	0
2	B	25	0	4	8	0
2	C	25	0	11	1	0
2	D	25	0	11	3	0
2	E	25	0	8	2	0
2	F	25	0	4	6	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	12	0	18	2	0
4	B	8	0	12	0	0
4	C	8	0	12	0	0
4	E	16	0	24	2	0
4	F	12	0	18	0	0
5	A	12	0	16	0	0
5	C	6	0	8	0	0
5	F	12	0	16	0	0
6	B	34	0	0	15	0
6	E	34	0	0	5	0
6	F	34	0	0	13	0
7	A	346	0	0	16	0
7	B	299	0	0	16	0
7	C	306	0	0	10	0
7	D	184	0	0	6	0
7	E	290	0	0	14	1
7	F	208	0	0	9	1
All	All	24993	0	22801	431	2

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 9.

The worst 5 of 431 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:205:MET:SD	1:A:334:GLU:HB3	1.50	1.50
1:A:205:MET:SD	1:A:334:GLU:CB	2.24	1.23
1:C:291[B]:ARG:HG3	1:C:291[B]:ARG:HH11	1.10	1.13
1:B:362:ARG:HH21	6:B:1570:Y6W:CAL	1.66	1.08
1:D:329:ASP:O	1:D:330:VAL:HG12	1.51	1.08

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:E:2235:HOH:O	7:F:2129:HOH:O[1_655]	1.89	0.31
1:E:320:GLU:OE2	1:F:84:GLU:OE2[3_554]	2.15	0.05

5.3 Torsion angles

5.3.1 Protein backbone

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	481/571 (84%)	462 (96%)	15 (3%)	4 (1%)	19	9
1	B	481/571 (84%)	459 (95%)	19 (4%)	3 (1%)	25	15
1	C	481/571 (84%)	453 (94%)	25 (5%)	3 (1%)	25	15
1	D	435/571 (76%)	402 (92%)	26 (6%)	7 (2%)	9	2
1	E	480/571 (84%)	460 (96%)	17 (4%)	3 (1%)	25	15
1	F	478/571 (84%)	447 (94%)	24 (5%)	7 (2%)	10	3
All	All	2836/3426 (83%)	2683 (95%)	126 (4%)	27 (1%)	15	6

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	330	VAL
1	B	332	GLY
1	B	477	GLY
1	C	199	ASP
1	D	440	PRO

5.3.2 Protein sidechains

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	420/485 (87%)	379 (90%)	41 (10%)	8	3
1	B	420/485 (87%)	387 (92%)	33 (8%)	12	5
1	C	420/485 (87%)	380 (90%)	40 (10%)	8	3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	385/485 (79%)	344 (89%)	41 (11%)	6	2
1	E	419/485 (86%)	388 (93%)	31 (7%)	13	6
1	F	417/485 (86%)	376 (90%)	41 (10%)	8	3
All	All	2481/2910 (85%)	2254 (91%)	227 (9%)	9	4

5 of 227 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	457	LEU
1	D	277	ASN
1	F	421	LEU
1	C	491	LYS
1	D	93	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 99 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	277	ASN
1	D	83	GLN
1	F	296	ASN
1	C	335	ASN
1	C	452	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 34 ligands modelled in this entry, 6 are monoatomic - leaving 28 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
5	GOL	F	1577	-	5,5,5	0.56	0	5,5,5	0.79	0
2	UDP	B	1572	3	20,26,26	1.05	1 (5%)	25,40,40	1.09	2 (8%)
2	UDP	C	1570	3	20,26,26	1.27	2 (10%)	25,40,40	1.32	2 (8%)
5	GOL	C	1574	-	5,5,5	0.57	0	5,5,5	0.47	0
4	EDO	A	1574	-	3,3,3	0.44	0	2,2,2	0.58	0
2	UDP	D	1555	3	20,26,26	1.05	2 (10%)	25,40,40	1.38	3 (12%)
4	EDO	B	1574	-	3,3,3	0.40	0	2,2,2	0.38	0
4	EDO	F	1575	-	3,3,3	0.53	0	2,2,2	0.25	0
4	EDO	E	1575	-	3,3,3	0.65	0	2,2,2	1.05	0
4	EDO	F	1574	-	3,3,3	0.61	0	2,2,2	0.17	0
4	EDO	A	1572	-	3,3,3	0.41	0	2,2,2	0.29	0
4	EDO	A	1573	-	3,3,3	0.60	0	2,2,2	0.51	0
6	Y6W	E	1572	-	32,36,36	1.89	7 (21%)	37,53,53	1.98	10 (27%)
4	EDO	C	1573	-	3,3,3	0.51	0	2,2,2	0.33	0
4	EDO	E	1574	-	3,3,3	0.52	0	2,2,2	0.05	0
5	GOL	A	1575	-	5,5,5	0.33	0	5,5,5	1.18	0
2	UDP	F	1570	3	20,26,26	1.22	2 (10%)	25,40,40	1.18	1 (4%)
6	Y6W	B	1570	3	32,36,36	1.46	4 (12%)	37,53,53	1.13	3 (8%)
4	EDO	C	1572	-	3,3,3	0.51	0	2,2,2	0.38	0
4	EDO	E	1573	-	3,3,3	0.88	0	2,2,2	0.72	0
4	EDO	E	1576	-	3,3,3	0.23	0	2,2,2	0.65	0
2	UDP	E	1570	3	20,26,26	1.13	2 (10%)	25,40,40	1.16	2 (8%)
4	EDO	F	1573	-	3,3,3	0.45	0	2,2,2	0.18	0
4	EDO	B	1573	-	3,3,3	0.48	0	2,2,2	0.35	0
6	Y6W	F	1572	-	32,36,36	1.74	8 (25%)	37,53,53	1.66	9 (24%)
5	GOL	A	1576	-	5,5,5	0.41	0	5,5,5	0.31	0
5	GOL	F	1576	-	5,5,5	0.41	0	5,5,5	0.75	0
2	UDP	A	1570	3	20,26,26	1.28	3 (15%)	25,40,40	1.43	4 (16%)

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
5	GOL	F	1577	-	-	2/4/4/4	-
2	UDP	B	1572	3	-	2/14/32/32	0/2/2/2
2	UDP	C	1570	3	-	3/14/32/32	0/2/2/2
5	GOL	C	1574	-	-	2/4/4/4	-
4	EDO	A	1574	-	-	1/1/1/1	-
2	UDP	D	1555	3	-	3/14/32/32	0/2/2/2
4	EDO	B	1574	-	-	0/1/1/1	-
4	EDO	F	1575	-	-	0/1/1/1	-
4	EDO	E	1575	-	-	1/1/1/1	-
4	EDO	F	1574	-	-	1/1/1/1	-
4	EDO	A	1572	-	-	1/1/1/1	-
4	EDO	A	1573	-	-	0/1/1/1	-
6	Y6W	E	1572	-	-	8/17/55/55	0/3/3/3
4	EDO	C	1573	-	-	1/1/1/1	-
4	EDO	E	1574	-	-	0/1/1/1	-
5	GOL	A	1575	-	-	2/4/4/4	-
2	UDP	F	1570	3	-	3/14/32/32	0/2/2/2
6	Y6W	B	1570	3	-	7/17/55/55	0/3/3/3
4	EDO	C	1572	-	-	0/1/1/1	-
4	EDO	E	1573	-	-	1/1/1/1	-
4	EDO	E	1576	-	-	0/1/1/1	-
2	UDP	E	1570	3	-	2/14/32/32	0/2/2/2
4	EDO	F	1573	-	-	0/1/1/1	-
4	EDO	B	1573	-	-	1/1/1/1	-
6	Y6W	F	1572	-	-	10/17/55/55	0/3/3/3
5	GOL	A	1576	-	-	2/4/4/4	-
5	GOL	F	1576	-	-	0/4/4/4	-
2	UDP	A	1570	3	-	0/14/32/32	0/2/2/2

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	E	1572	Y6W	O4'-C1'	5.19	1.48	1.41
6	B	1570	Y6W	OBG-CBD	4.76	1.36	1.24
6	E	1572	Y6W	OBG-CBD	4.56	1.36	1.24
2	C	1570	UDP	O4'-C1'	4.43	1.47	1.41
6	F	1572	Y6W	O4'-C1'	4.22	1.47	1.41

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	E	1572	Y6W	O5'-PAN-CAM	-4.63	91.88	104.11
6	E	1572	Y6W	C5-O5-C1	-4.59	105.14	113.16
2	A	1570	UDP	O3B-PB-O3A	4.30	119.04	104.64
2	D	1555	UDP	O3B-PB-O1B	3.87	125.82	110.68
6	F	1572	Y6W	O4'-C1'-C2'	-3.73	101.48	106.93

There are no chirality outliers.

5 of 53 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	F	1577	GOL	O1-C1-C2-C3
2	B	1572	UDP	PA-O3A-PB-O2B
2	B	1572	UDP	PA-O3A-PB-O3B
2	C	1570	UDP	PA-O3A-PB-O3B
5	C	1574	GOL	O1-C1-C2-C3

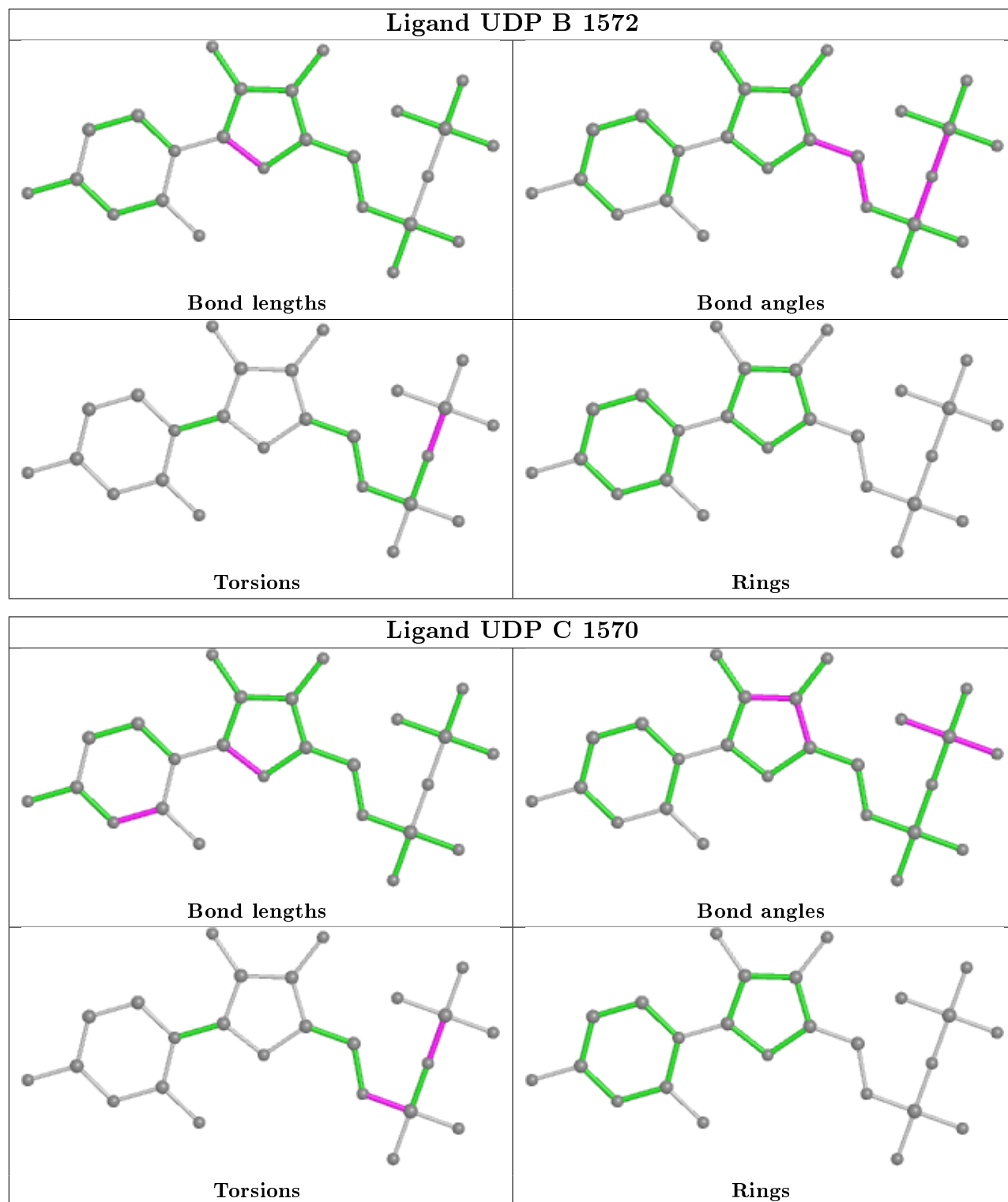
There are no ring outliers.

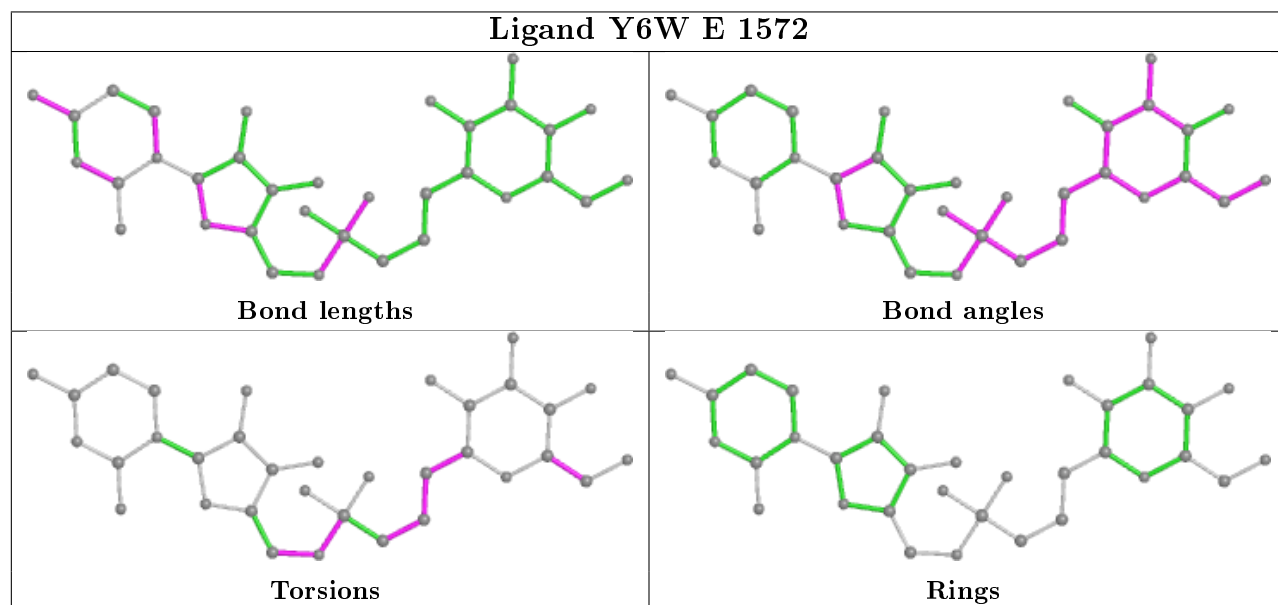
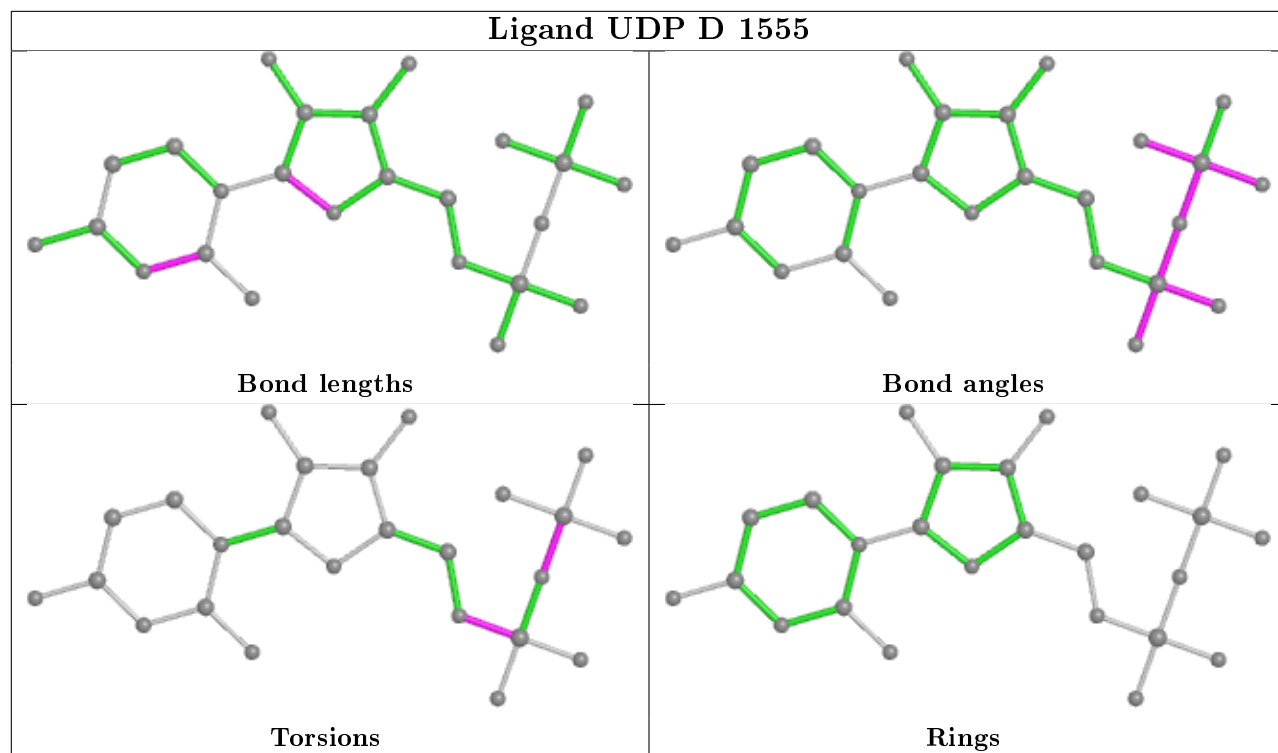
11 monomers are involved in 50 short contacts:

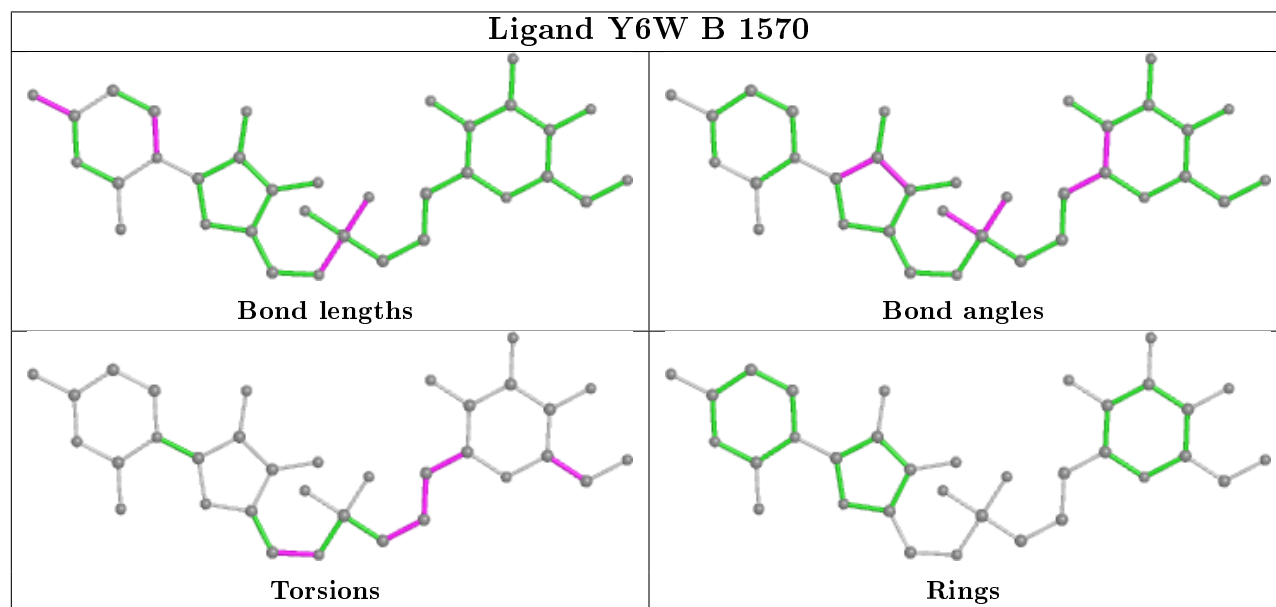
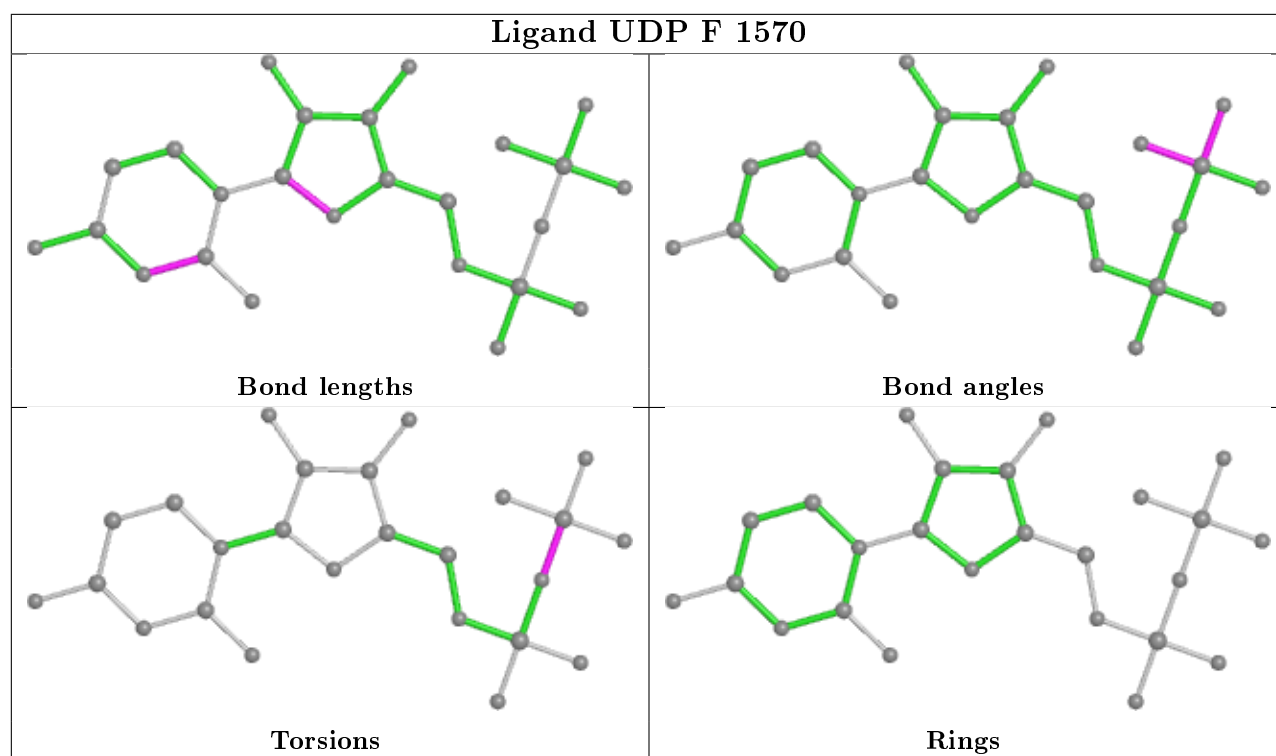
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1572	UDP	8	0
2	C	1570	UDP	1	0
2	D	1555	UDP	3	0
4	E	1575	EDO	2	0
4	A	1573	EDO	2	0
6	E	1572	Y6W	5	0
2	F	1570	UDP	6	0
6	B	1570	Y6W	15	0
2	E	1570	UDP	2	0
6	F	1572	Y6W	13	0
2	A	1570	UDP	3	0

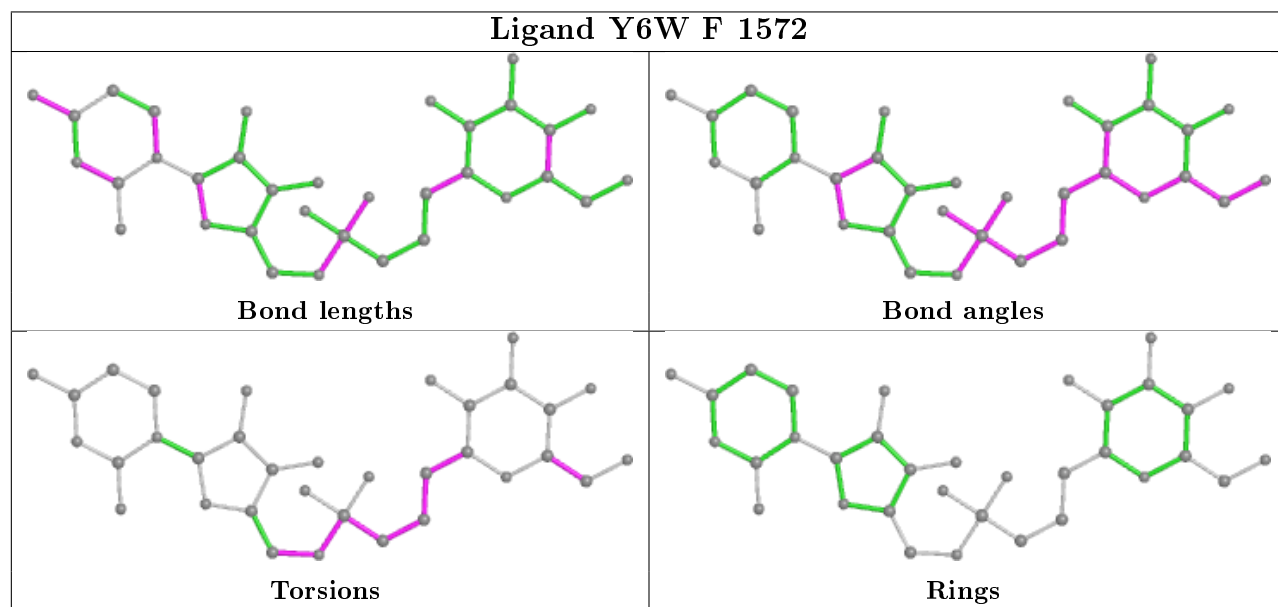
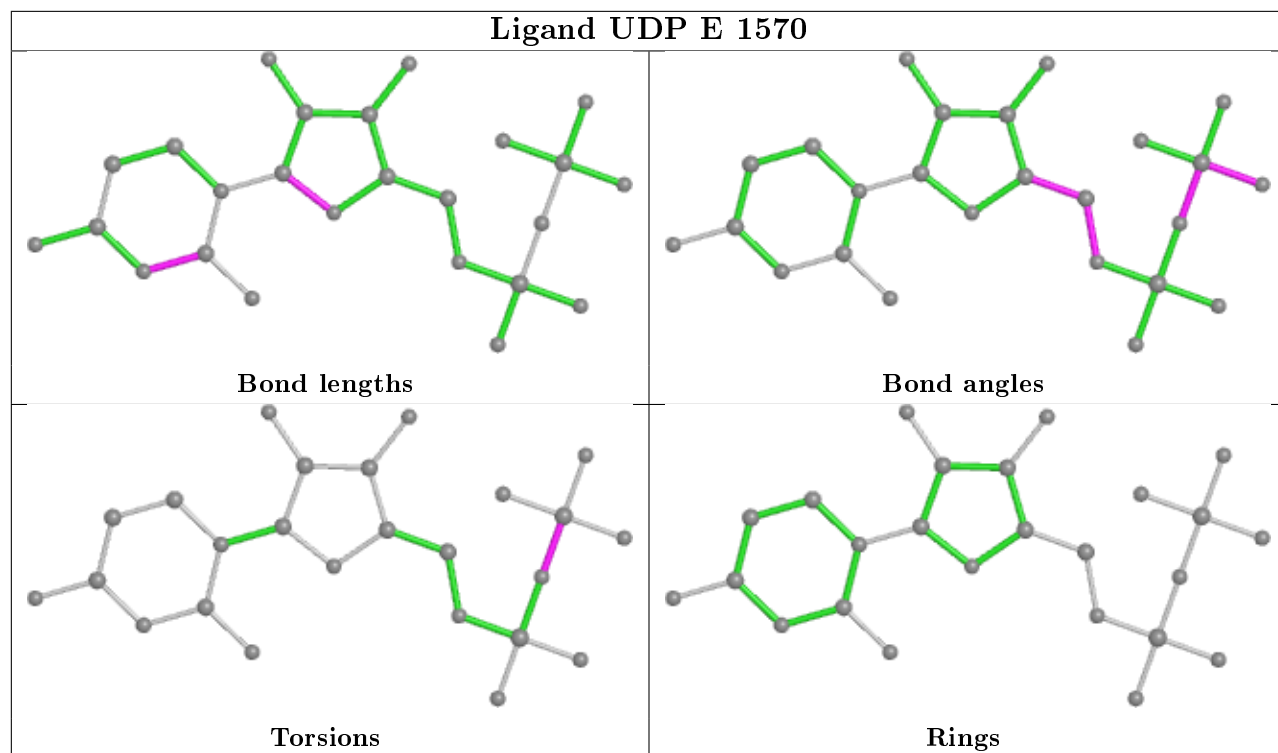
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

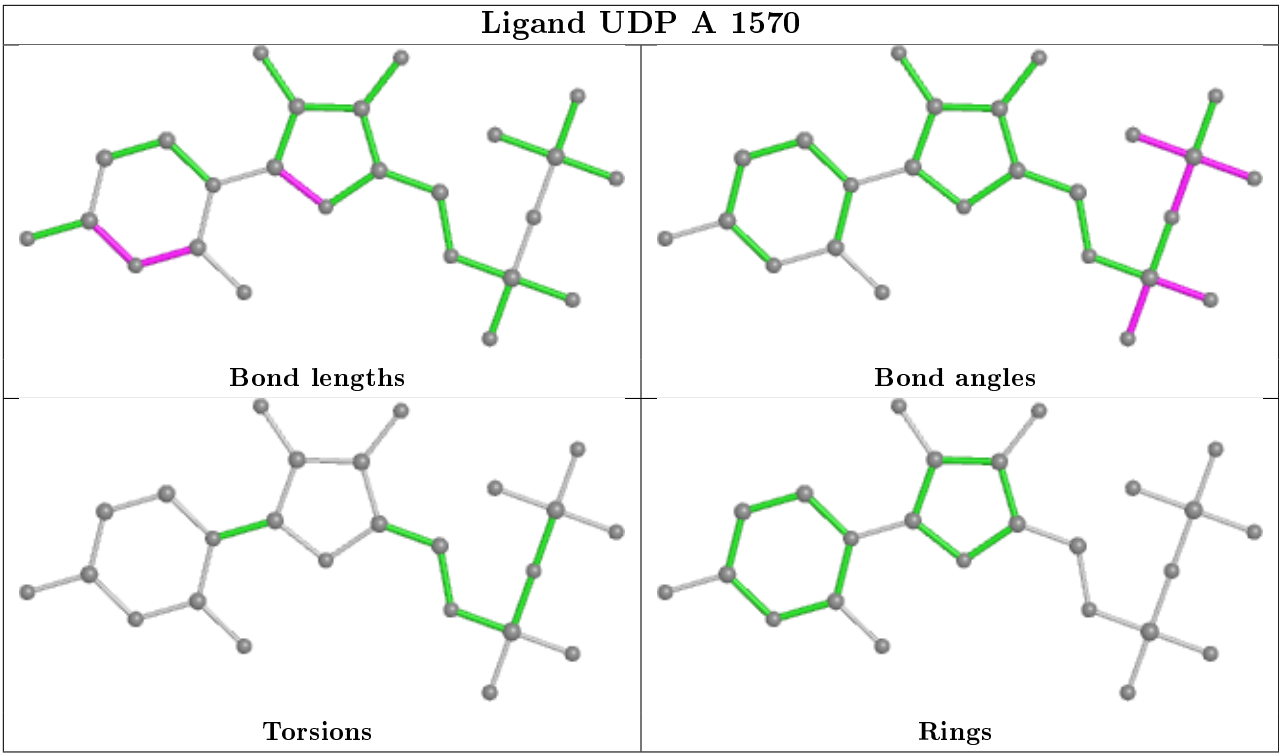
any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.











5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	E	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	E	319:GLU	C	320:GLU	N	1.19
1	E	513:CYS	C	514:ARG	N	1.15

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	483/571 (84%)	0.22	26 (5%)	25 26	23, 42, 67, 107	5 (1%)
1	B	484/571 (84%)	0.27	30 (6%)	20 21	26, 41, 73, 147	5 (1%)
1	C	482/571 (84%)	0.34	39 (8%)	12 12	26, 45, 70, 110	5 (1%)
1	D	443/571 (77%)	0.67	63 (14%)	2 2	27, 52, 87, 107	5 (1%)
1	E	483/571 (84%)	0.31	29 (6%)	21 22	26, 40, 67, 105	5 (1%)
1	F	481/571 (84%)	0.65	53 (11%)	5 5	29, 55, 82, 96	5 (1%)
All	All	2856/3426 (83%)	0.41	240 (8%)	11 11	23, 45, 79, 147	30 (1%)

The worst 5 of 240 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	525	ILE	10.2
1	B	93	ARG	10.1
1	D	331	TRP	9.0
1	A	377	PHE	7.3
1	B	92	VAL	7.0

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 carbohydrates in this entry.

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, 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.

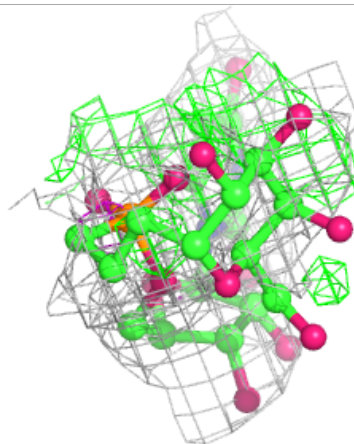
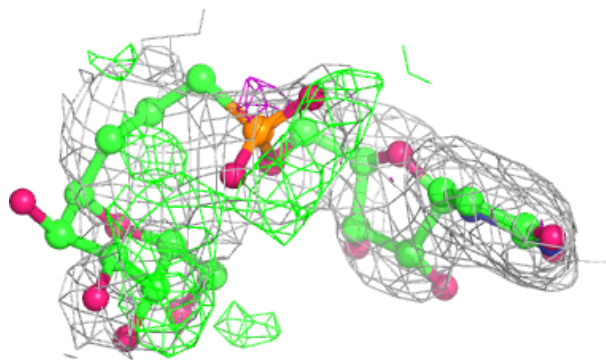
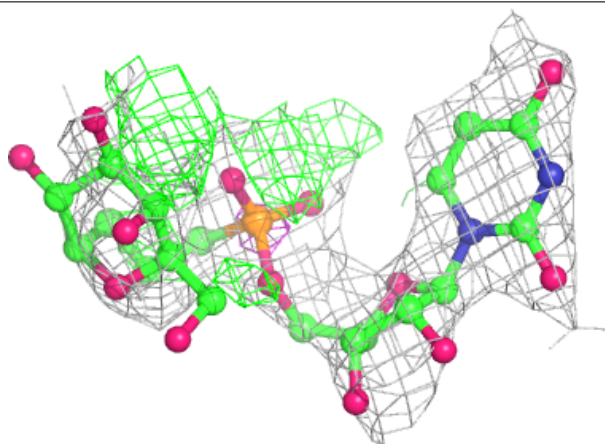
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	EDO	A	1574	4/4	0.57	0.35	72,75,80,80	0
5	GOL	A	1576	6/6	0.64	0.34	74,80,86,89	0
4	EDO	E	1574	4/4	0.72	0.22	63,65,72,73	0
4	EDO	C	1573	4/4	0.72	0.29	65,72,73,77	0
4	EDO	F	1574	4/4	0.75	0.23	72,73,76,78	0
4	EDO	F	1573	4/4	0.77	0.27	72,74,76,79	0
4	EDO	E	1575	4/4	0.78	0.16	44,47,51,62	0
4	EDO	F	1575	4/4	0.79	0.24	66,69,69,71	0
4	EDO	A	1573	4/4	0.79	0.41	64,65,66,67	0
4	EDO	B	1573	4/4	0.79	0.16	55,64,67,70	0
5	GOL	C	1574	6/6	0.79	0.14	50,60,63,65	0
4	EDO	A	1572	4/4	0.80	0.33	78,80,81,84	0
5	GOL	F	1577	6/6	0.81	0.14	66,67,69,72	0
5	GOL	A	1575	6/6	0.83	0.24	56,65,73,78	0
4	EDO	C	1572	4/4	0.85	0.22	52,55,56,60	0
5	GOL	F	1576	6/6	0.87	0.14	68,73,81,87	0
6	Y6W	B	1570	34/34	0.88	0.22	40,53,57,58	34
2	UDP	B	1572	25/25	0.89	0.18	35,53,57,58	25
2	UDP	F	1570	25/25	0.90	0.17	37,44,46,49	25
6	Y6W	F	1572	34/34	0.90	0.24	61,78,100,105	34
6	Y6W	E	1572	34/34	0.93	0.19	48,56,73,74	34
4	EDO	B	1574	4/4	0.94	0.15	50,54,55,56	0
4	EDO	E	1573	4/4	0.94	0.14	30,31,37,40	0
2	UDP	E	1570	25/25	0.94	0.15	29,35,42,42	25
2	UDP	D	1555	25/25	0.94	0.16	51,60,70,76	0
2	UDP	C	1570	25/25	0.95	0.15	40,49,55,59	0
4	EDO	E	1576	4/4	0.96	0.16	44,50,51,60	0
2	UDP	A	1570	25/25	0.97	0.13	37,44,50,51	0
3	MN	F	1571	1/1	0.99	0.09	34,34,34,34	0
3	MN	B	1571	1/1	0.99	0.10	35,35,35,35	0
3	MN	A	1571	1/1	0.99	0.07	27,27,27,27	0
3	MN	D	1556	1/1	0.99	0.08	41,41,41,41	0
3	MN	E	1571	1/1	1.00	0.10	29,29,29,29	0
3	MN	C	1571	1/1	1.00	0.08	30,30,30,30	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.

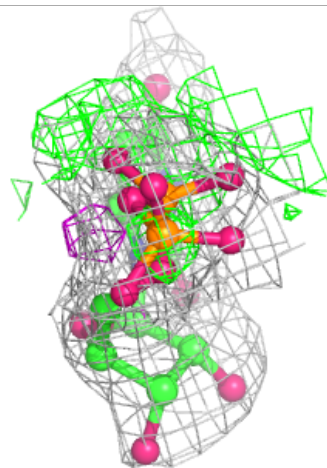
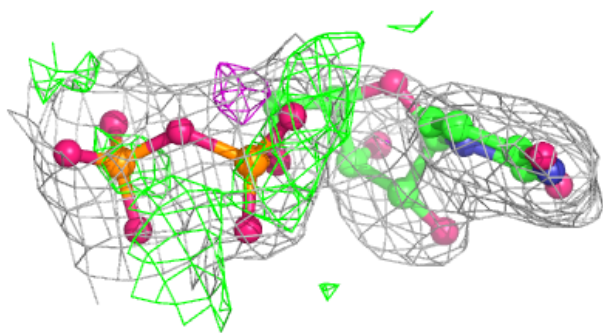
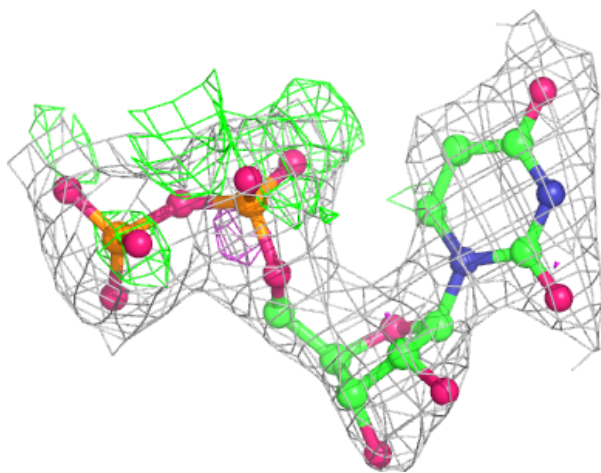
Electron density around Y6W B 1570:

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



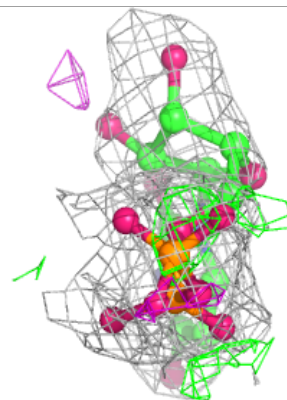
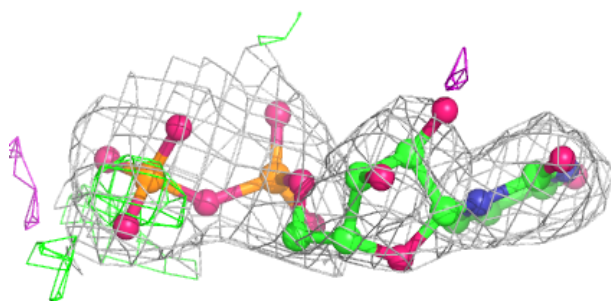
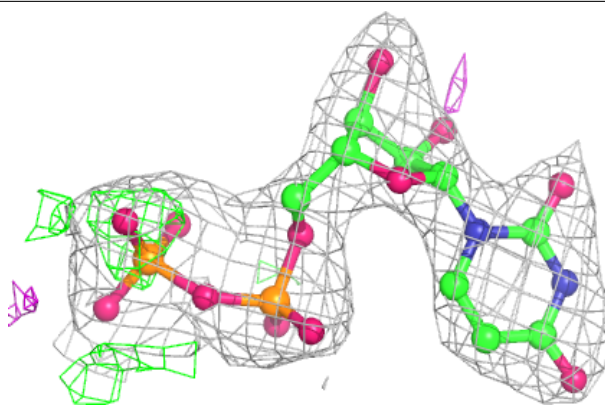
Electron density around UDP B 1572:

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

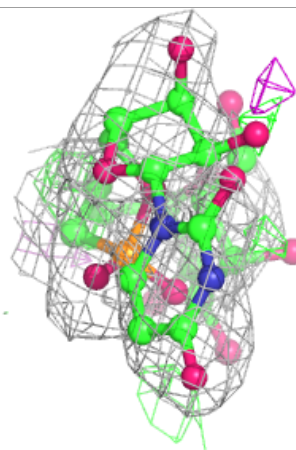
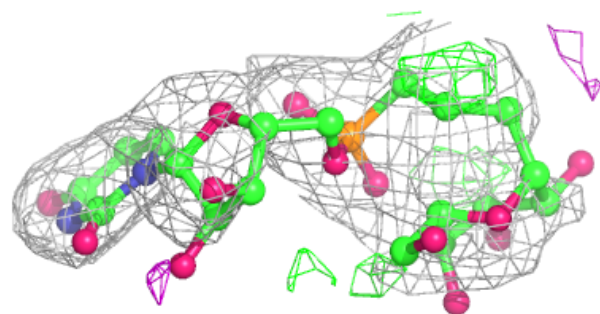
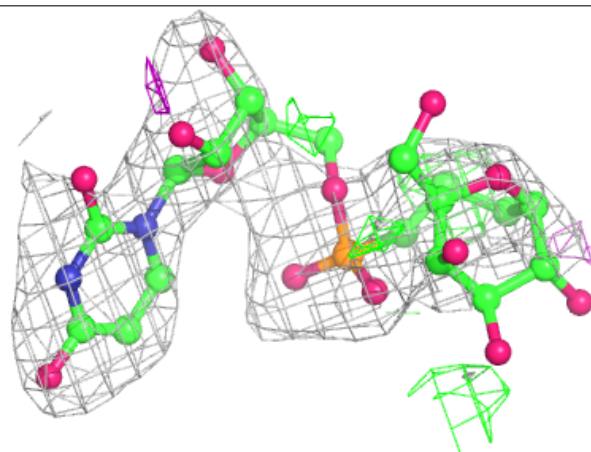


Electron density around UDP F 1570:

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

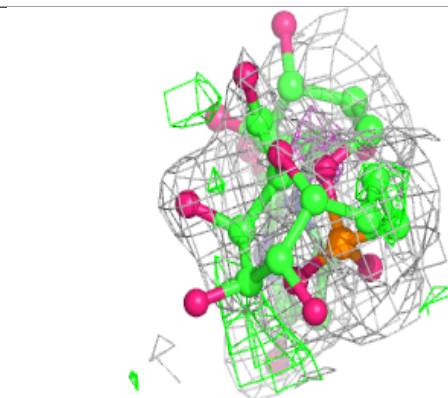
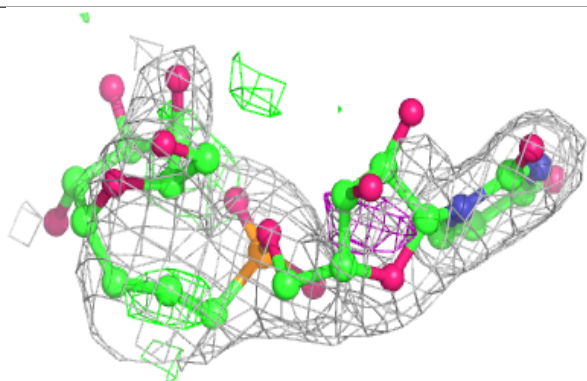
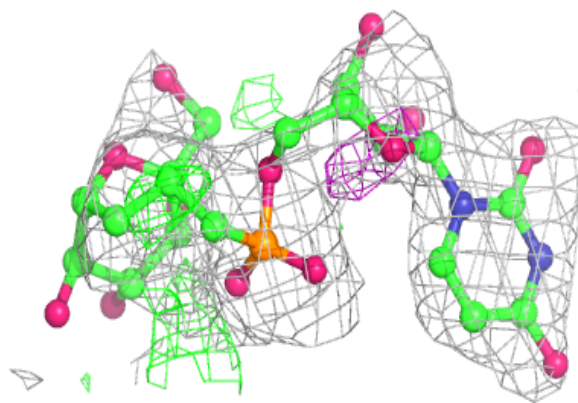
**Electron density around Y6W F 1572:**

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

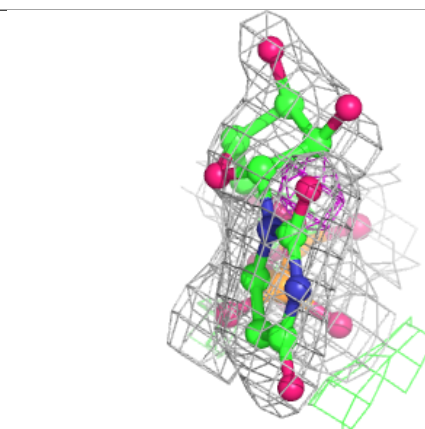
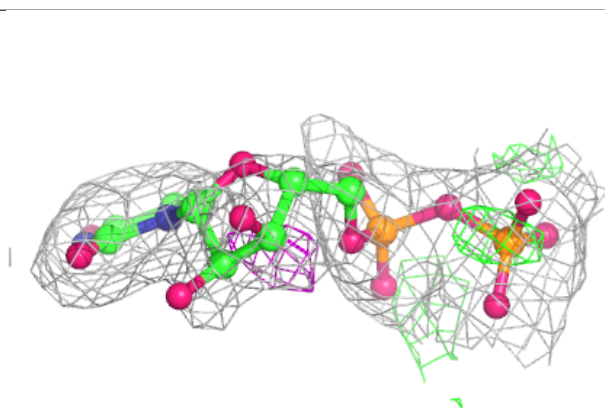
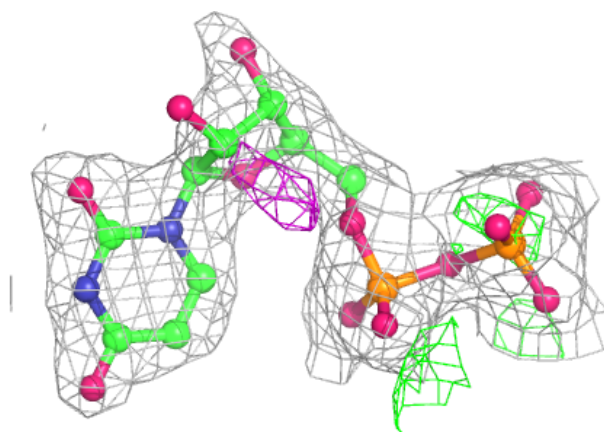


Electron density around Y6W E 1572:

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

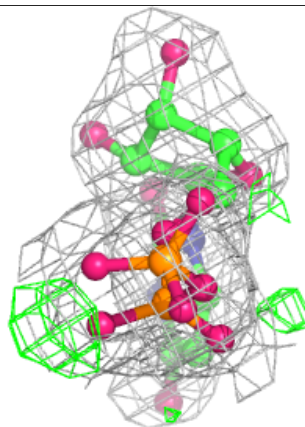
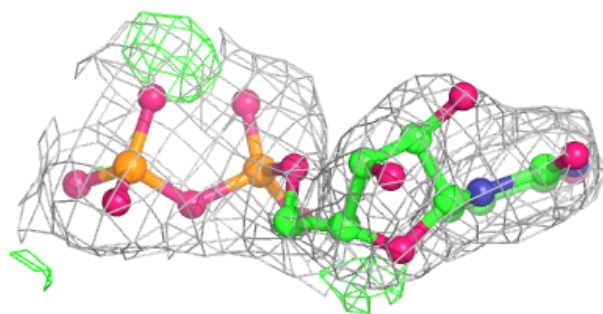
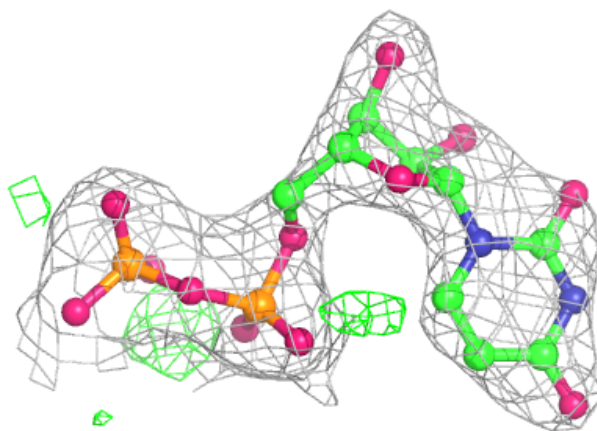
**Electron density around UDP E 1570:**

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



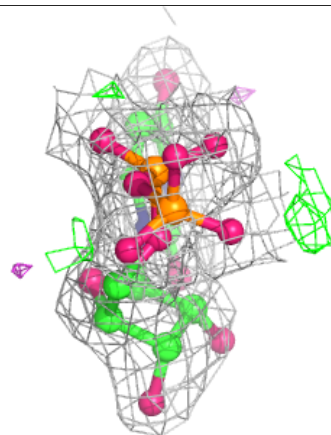
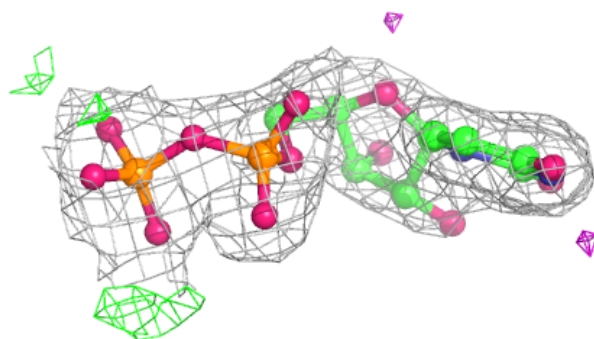
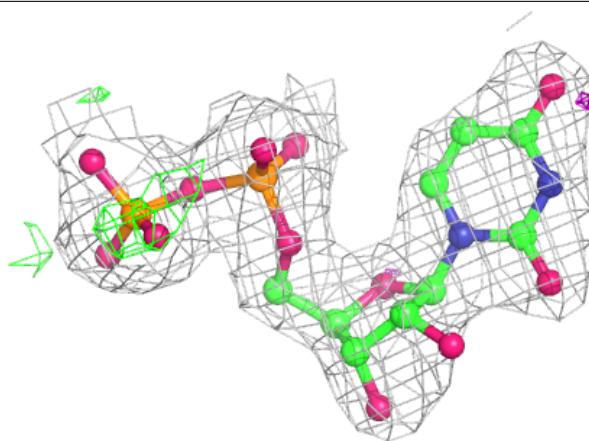
Electron density around UDP D 1555:

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



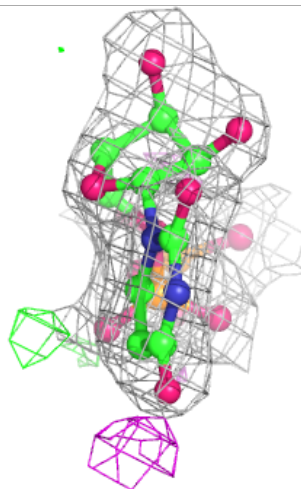
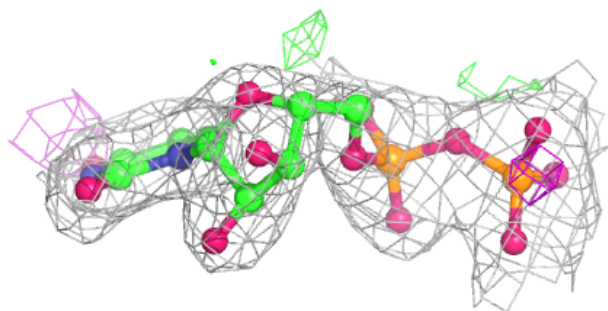
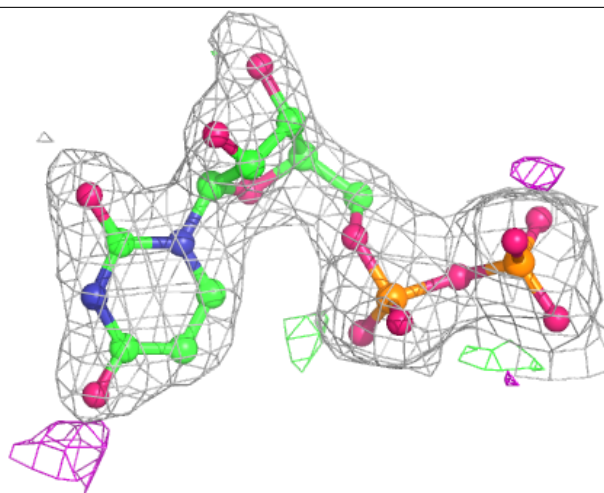
Electron density around UDP C 1570:

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



Electron density around UDP A 1570:

$2mF_o-DF_c$ (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.