



wwPDB X-ray Structure Validation Summary Report ⓘ

May 27, 2020 – 01:16 am BST

PDB ID : 6OIY
Title : Structure of Escherichia coli bound to dGTP
Authors : Barnes, C.O.; Wu, Y.; Calero, G.
Deposited on : 2019-04-09
Resolution : 3.29 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

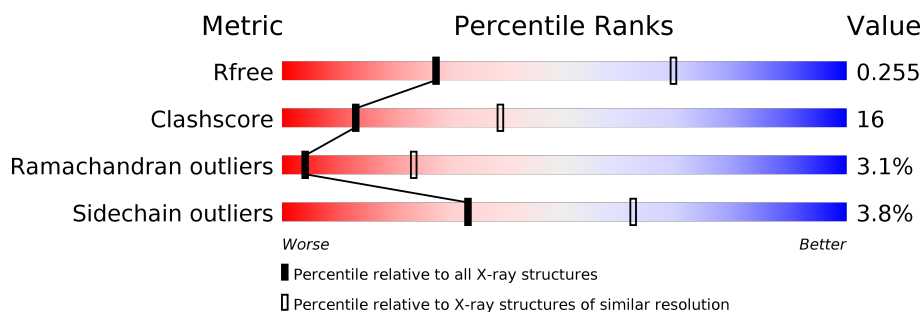
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.29 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	505	63% 31% . .
1	B	505	50% 43% 5% .
1	C	505	73% 24% .
1	D	505	69% 27% . . .
1	E	505	69% 25% . .
1	F	505	74% 22% . .

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	DGT	B	601	-	-	X	-

2 Entry composition [i](#)

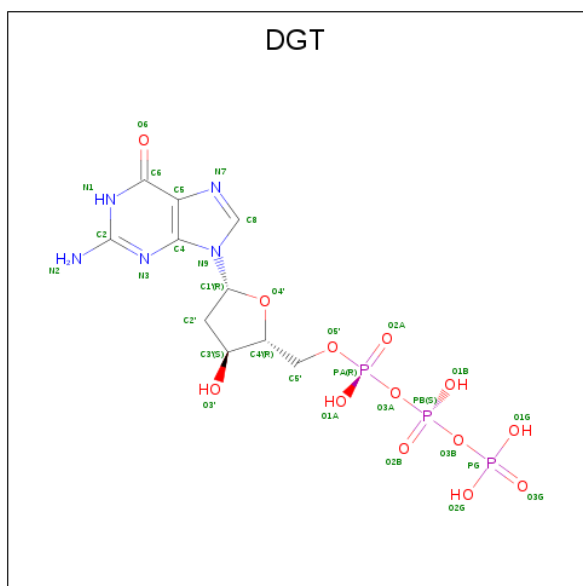
There are 4 unique types of molecules in this entry. The entry contains 24717 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Deoxyguanosinetriphosphate triphosphohydrolase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	503	Total	C	N	O	S	0	0	0
			4115	2627	737	735	16			
1	B	503	Total	C	N	O	S	0	0	0
			4112	2624	737	735	16			
1	C	503	Total	C	N	O	S	0	0	0
			4112	2624	737	735	16			
1	D	492	Total	C	N	O	S	0	0	0
			4058	2592	726	724	16			
1	E	492	Total	C	N	O	S	0	0	0
			4058	2592	726	724	16			
1	F	492	Total	C	N	O	S	0	0	0
			4058	2592	726	724	16			

- Molecule 2 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DGT) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	C	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	D	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	E	1	Total	C	N	O	P	0	0
			31	10	5	13	3		
2	F	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

- 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		

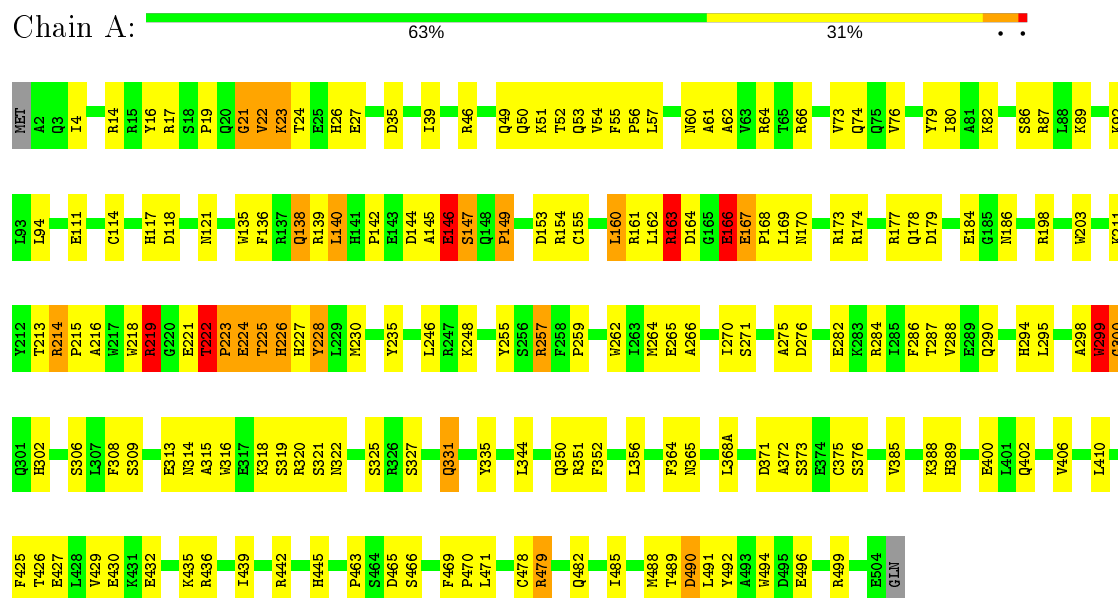
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	O	0	0
			2	2		
4	B	3	Total	O	0	0
			3	3		
4	C	3	Total	O	0	0
			3	3		
4	D	1	Total	O	0	0
			1	1		
4	E	1	Total	O	0	0
			1	1		
4	F	2	Total	O	0	0
			2	2		

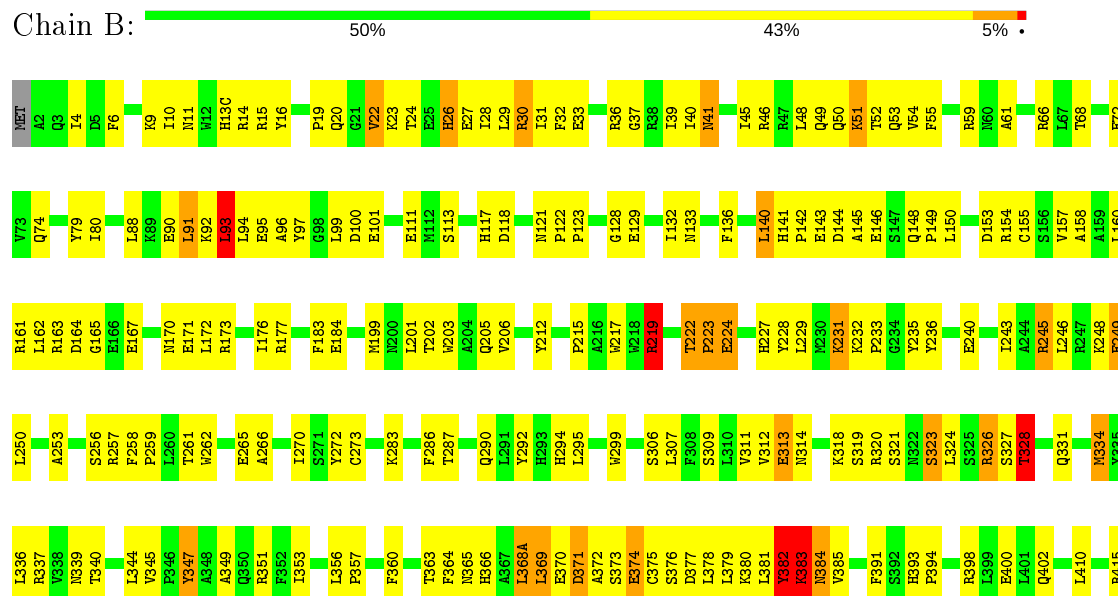
3 Residue-property plots

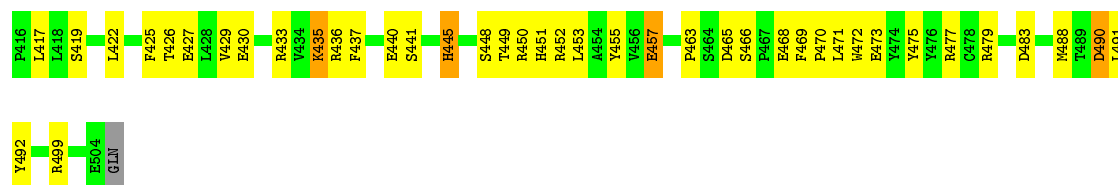
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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



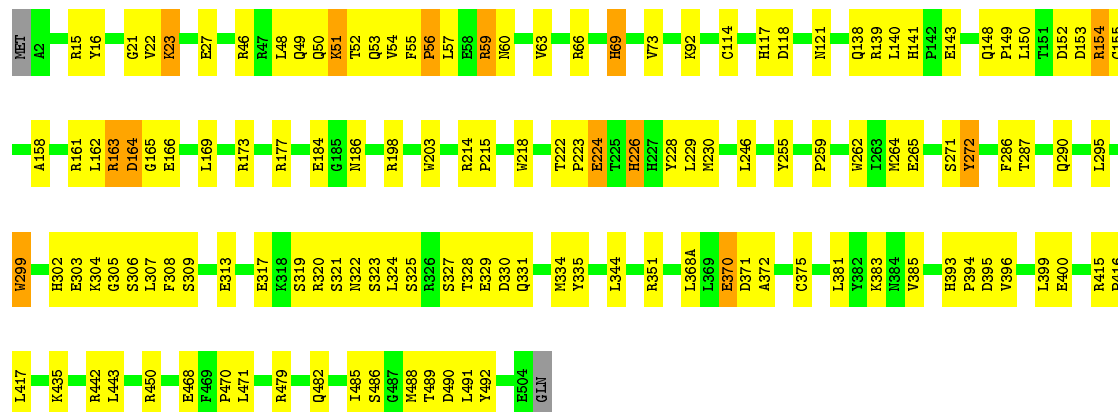
• Molecule 1: Deoxyguanosinetriphosphate triphosphohydrolase





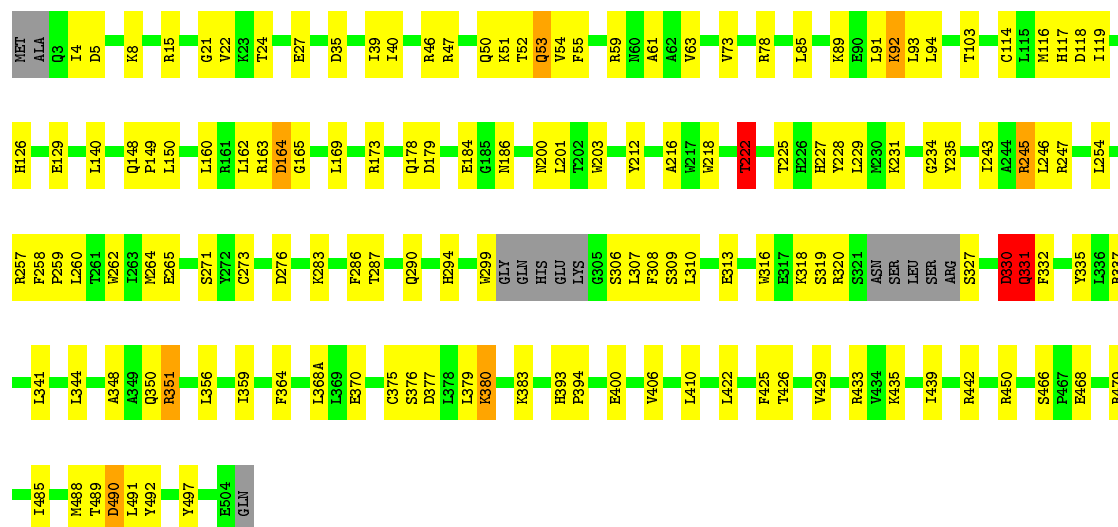
- Molecule 1: Deoxyguanosinetriphosphate triphosphohydrolase

Chain C: 73% 24% .



- Molecule 1: Deoxyguanosinetriphosphate triphosphohydrolase

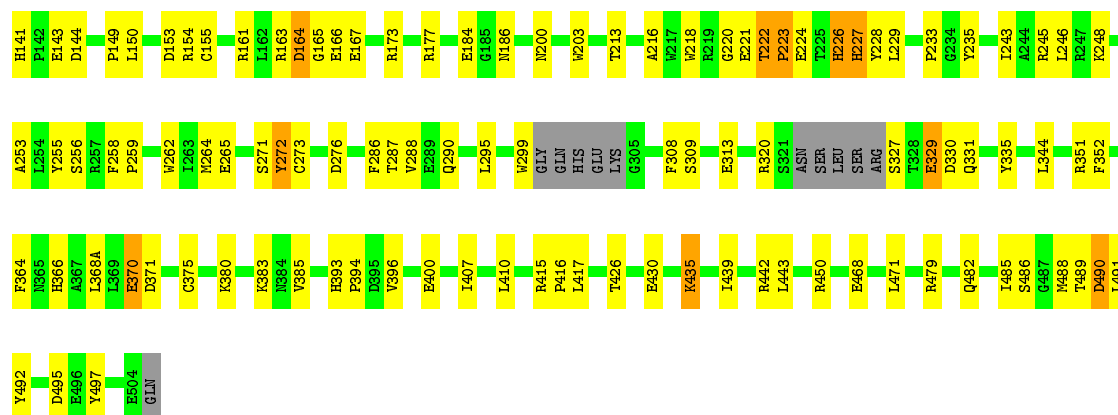
Chain D: 69% 27% . . .



- Molecule 1: Deoxyguanosinetriphosphate triphosphohydrolase

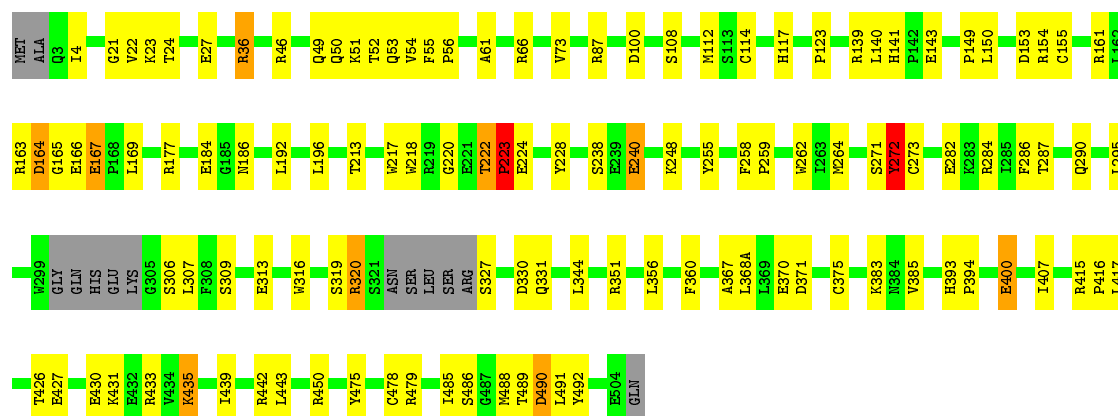
Chain E: 69% 25% . . .





• Molecule 1: Deoxyguanosinetriphosphate triphosphohydrolase

Chain F: 74% 22% • •



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	192.18 Å 192.18 Å 299.60 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.87 – 3.29 49.15 – 3.29	Depositor EDS
% Data completeness (in resolution range)	98.3 (46.87-3.29) 99.8 (49.15-3.29)	Depositor EDS
R_{merge}	0.25	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.01 (at 3.33 Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, R_{free}	0.205 , 0.241 0.247 , 0.255	Depositor DCC
R_{free} test set	2552 reflections (3.00%)	wwPDB-VP
Wilson B-factor (Å ²)	151.0	Xtriage
Anisotropy	0.223	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 136.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	24717	wwPDB-VP
Average B, all atoms (Å ²)	160.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.09% 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: MN, DGT

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	0.75	5/4216 (0.1%)	0.76	8/5705 (0.1%)
1	B	0.59	4/4213 (0.1%)	0.87	14/5699 (0.2%)
1	C	0.64	1/4213 (0.0%)	0.70	1/5699 (0.0%)
1	D	0.52	1/4157 (0.0%)	0.73	11/5618 (0.2%)
1	E	0.59	3/4157 (0.1%)	0.73	8/5618 (0.1%)
1	F	0.61	3/4157 (0.1%)	0.69	2/5618 (0.0%)
All	All	0.62	17/25113 (0.1%)	0.75	44/33957 (0.1%)

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	3
1	B	0	2
1	C	0	1
1	D	0	2
1	E	0	4
1	F	0	1
All	All	0	13

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	334	MET	CG-SD	-13.84	1.45	1.81
1	A	160	LEU	CG-CD1	10.65	1.91	1.51
1	A	479	ARG	CZ-NH2	-6.96	1.24	1.33
1	E	329	GLU	CD-OE1	-6.72	1.18	1.25
1	D	332	PHE	CB-CG	6.64	1.62	1.51

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	383	LYS	CD-CE-NZ	-15.62	75.78	111.70
1	E	17	ARG	CB-CG-CD	11.09	140.42	111.60
1	B	334	MET	CA-CB-CG	10.59	131.31	113.30
1	E	17	ARG	CG-CD-NE	10.34	133.50	111.80
1	B	140	LEU	CA-CB-CG	9.81	137.87	115.30

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	163	ARG	Sidechain
1	A	219	ARG	Sidechain
1	A	222	THR	Peptide
1	B	219	ARG	Sidechain
1	B	382	TYR	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	4115	0	4033	161	0
1	B	4112	0	4027	231	0
1	C	4112	0	4031	106	0
1	D	4058	0	4003	123	0
1	E	4058	0	4000	108	0
1	F	4058	0	4002	82	0
2	A	31	0	12	5	0
2	B	31	0	12	11	0
2	C	31	0	12	3	0
2	D	31	0	12	6	0
2	E	31	0	12	2	0
2	F	31	0	12	4	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	2	0	0	0	0
4	B	3	0	0	1	0
4	C	3	0	0	1	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	2	0	0	0	0
All	All	24717	0	24168	799	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 16.

The worst 5 of 799 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:160:LEU:CD1	1:A:160:LEU:CG	1.91	1.47
1:E:127:PHE:CE2	1:E:400:GLU:OE1	1.70	1.43
1:F:218:TRP:NE1	1:F:220:GLY:O	1.62	1.29
2:F:601:DGT:C4'	2:F:601:DGT:O4'	1.67	1.28
2:C:601:DGT:O4'	2:C:601:DGT:C4'	1.66	1.27

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	501/505 (99%)	448 (89%)	33 (7%)	20 (4%)	3	18
1	B	501/505 (99%)	423 (84%)	55 (11%)	23 (5%)	2	15
1	C	501/505 (99%)	444 (89%)	38 (8%)	19 (4%)	3	19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	486/505 (96%)	442 (91%)	33 (7%)	11 (2%)	6	29
1	E	486/505 (96%)	446 (92%)	31 (6%)	9 (2%)	8	34
1	F	486/505 (96%)	442 (91%)	35 (7%)	9 (2%)	8	34
All	All	2961/3030 (98%)	2645 (89%)	225 (8%)	91 (3%)	4	23

5 of 91 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	22	VAL
1	A	23	LYS
1	A	57	LEU
1	A	61	ALA
1	A	92	LYS

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	431/450 (96%)	408 (95%)	23 (5%)	22	53
1	B	431/450 (96%)	410 (95%)	21 (5%)	25	56
1	C	431/450 (96%)	416 (96%)	15 (4%)	36	64
1	D	431/450 (96%)	420 (97%)	11 (3%)	46	71
1	E	431/450 (96%)	417 (97%)	14 (3%)	39	67
1	F	431/450 (96%)	416 (96%)	15 (4%)	36	64
All	All	2586/2700 (96%)	2487 (96%)	99 (4%)	33	62

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	15	ARG
1	C	435	LYS
1	F	248	LYS
1	C	163	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	264	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	121	ASN
1	B	384	ASN
1	C	69	HIS
1	B	74	GLN
1	C	53	GLN

5.3.3 RNA [i](#)

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 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	$\# Z > 2$
2	DGT	A	601	3	26,33,33	1.56	4 (15%)	32,52,52	2.63	9 (28%)
2	DGT	C	601	-	26,33,33	4.05	11 (42%)	32,52,52	1.79	7 (21%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	DGT	E	601	-	26,33,33	1.51	7 (26%)	32,52,52	2.04	9 (28%)
2	DGT	D	602	-	26,33,33	1.55	6 (23%)	32,52,52	2.37	12 (37%)
2	DGT	B	601	-	26,33,33	1.60	5 (19%)	32,52,52	2.53	15 (46%)
2	DGT	F	601	3	26,33,33	4.00	11 (42%)	32,52,52	1.64	5 (15%)

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	DGT	A	601	3	-	9/18/34/34	0/3/3/3
2	DGT	C	601	-	-	1/18/34/34	0/3/3/3
2	DGT	E	601	-	-	1/18/34/34	0/3/3/3
2	DGT	D	602	-	-	5/18/34/34	0/3/3/3
2	DGT	B	601	-	-	5/18/34/34	0/3/3/3
2	DGT	F	601	3	-	2/18/34/34	0/3/3/3

The worst 5 of 44 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	601	DGT	O4'-C4'	9.86	1.67	1.45
2	C	601	DGT	O4'-C4'	9.62	1.66	1.45
2	C	601	DGT	C4-N3	8.74	1.49	1.35
2	F	601	DGT	C4-N3	8.27	1.48	1.35
2	C	601	DGT	C6-C5	7.15	1.53	1.41

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	601	DGT	C2-N3-C4	7.59	124.03	115.36
2	D	602	DGT	C2-N3-C4	7.40	123.81	115.36
2	A	601	DGT	C6-C5-C4	-7.26	113.86	120.80
2	B	601	DGT	C6-C5-C4	-5.81	115.25	120.80
2	C	601	DGT	N3-C2-N1	-5.65	119.69	127.22

There are no chirality outliers.

5 of 23 torsion outliers are listed below:

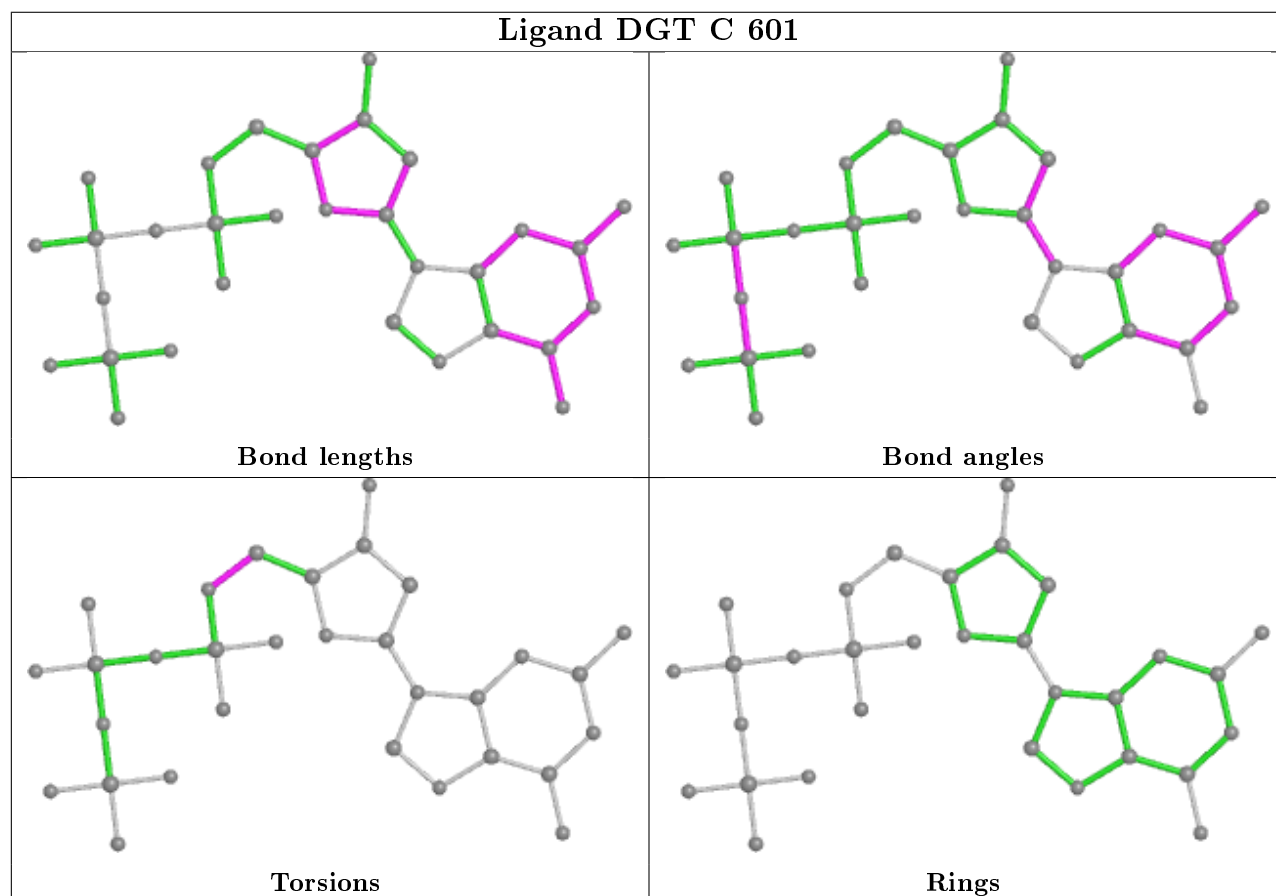
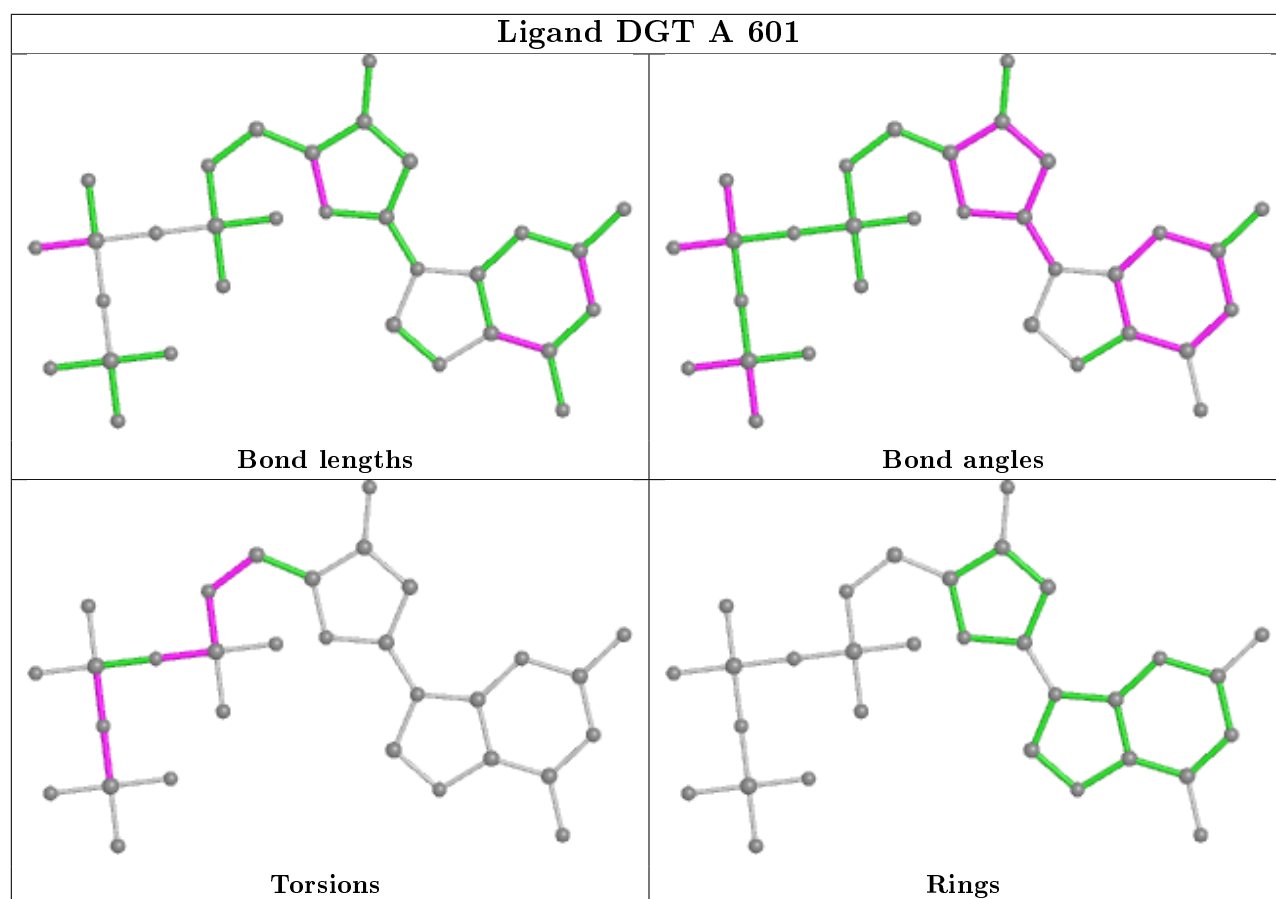
Mol	Chain	Res	Type	Atoms
2	A	601	DGT	PB-O3B-PG-O1G
2	A	601	DGT	C5'-O5'-PA-O1A
2	A	601	DGT	C4'-C5'-O5'-PA
2	D	602	DGT	C4'-C5'-O5'-PA
2	B	601	DGT	C5'-O5'-PA-O1A

There are no ring outliers.

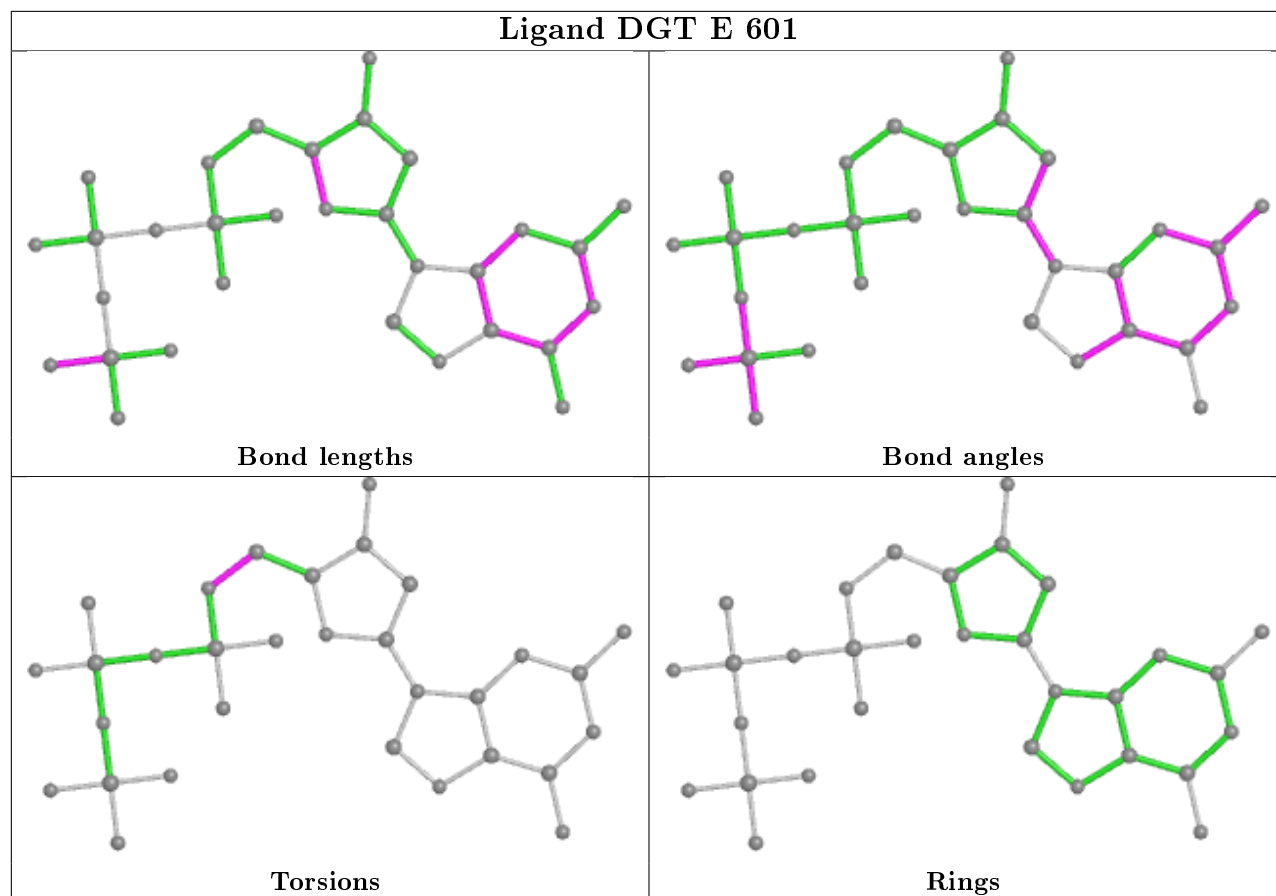
6 monomers are involved in 31 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	DGT	5	0
2	C	601	DGT	3	0
2	E	601	DGT	2	0
2	D	602	DGT	6	0
2	B	601	DGT	11	0
2	F	601	DGT	4	0

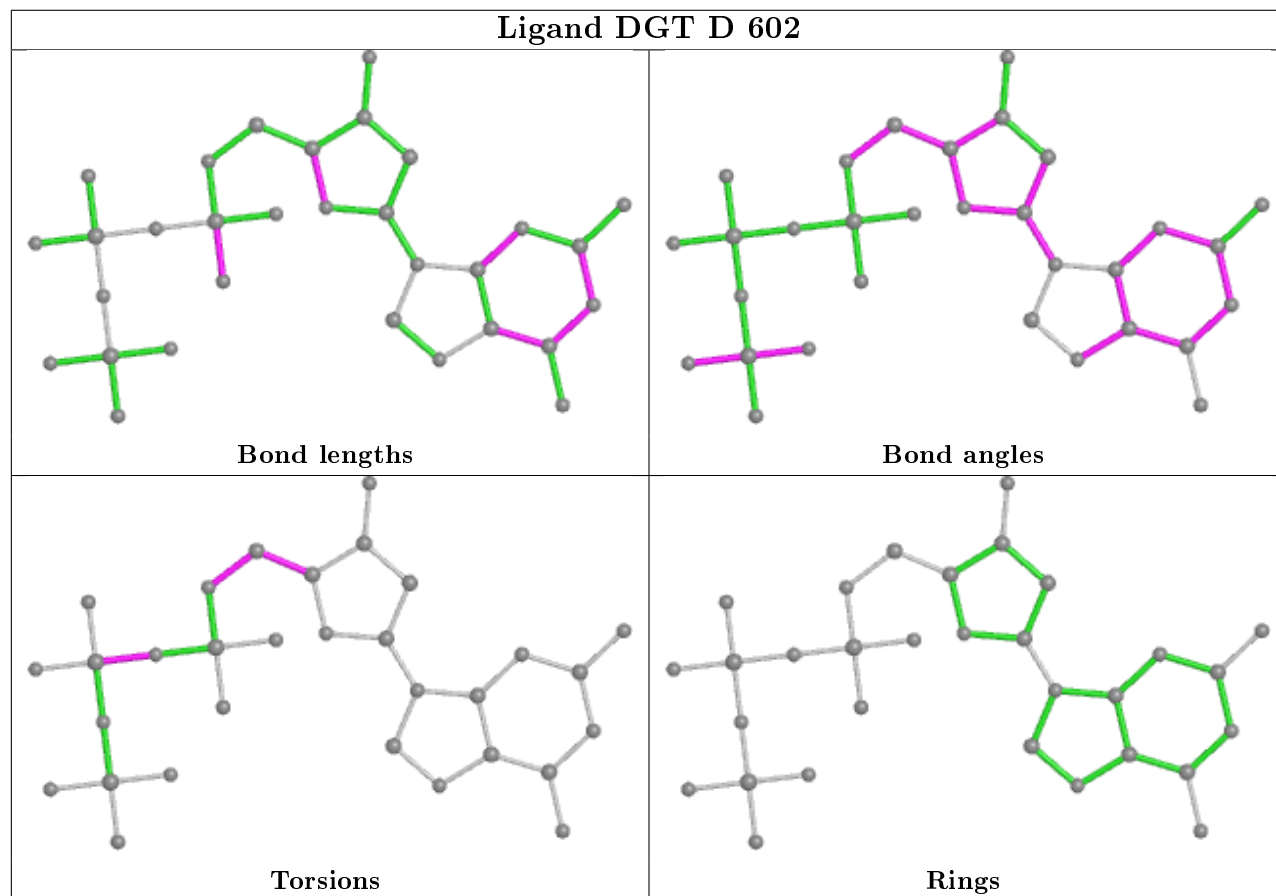
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included.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

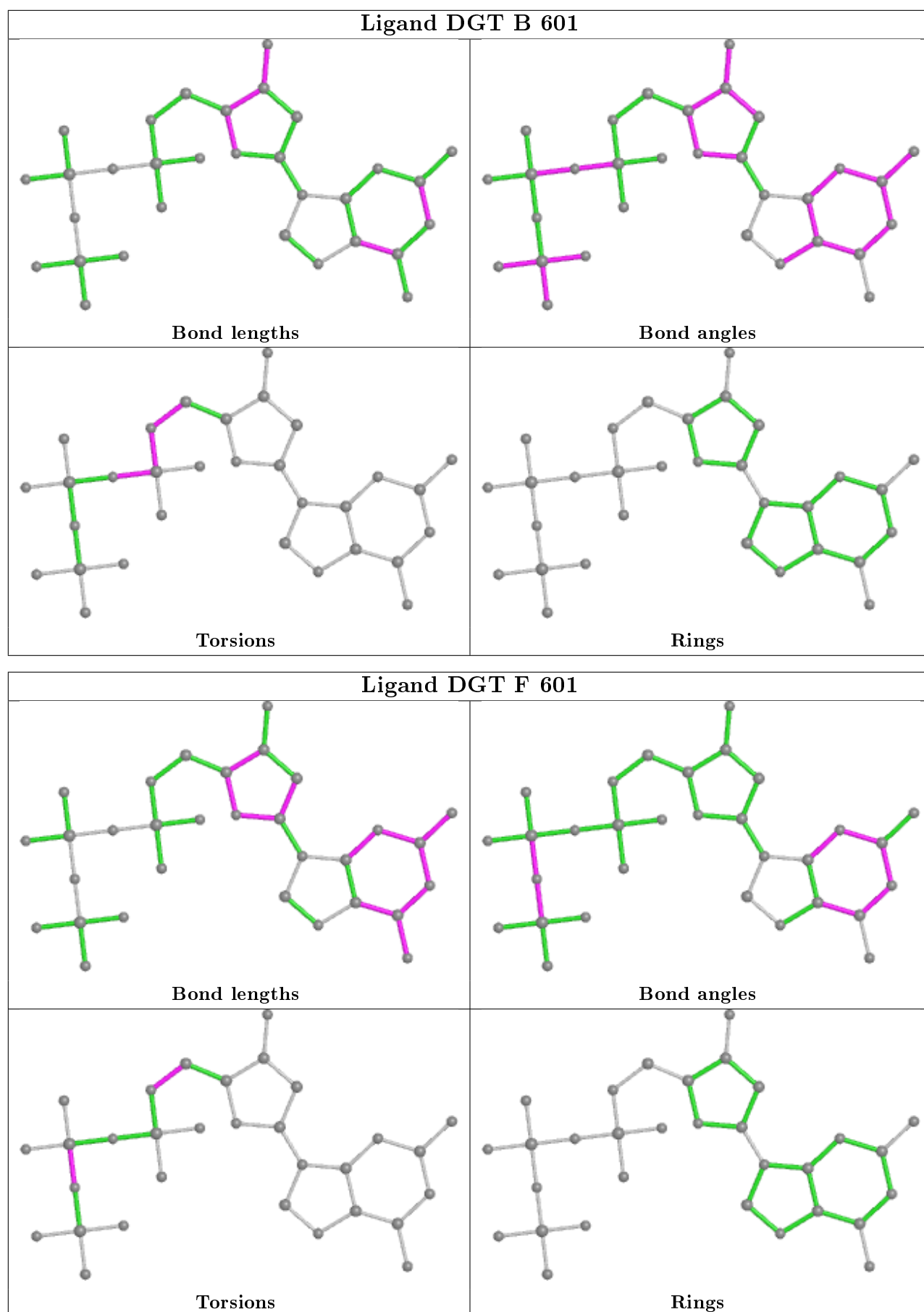


Ligand DGT E 601



Ligand DGT D 602





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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

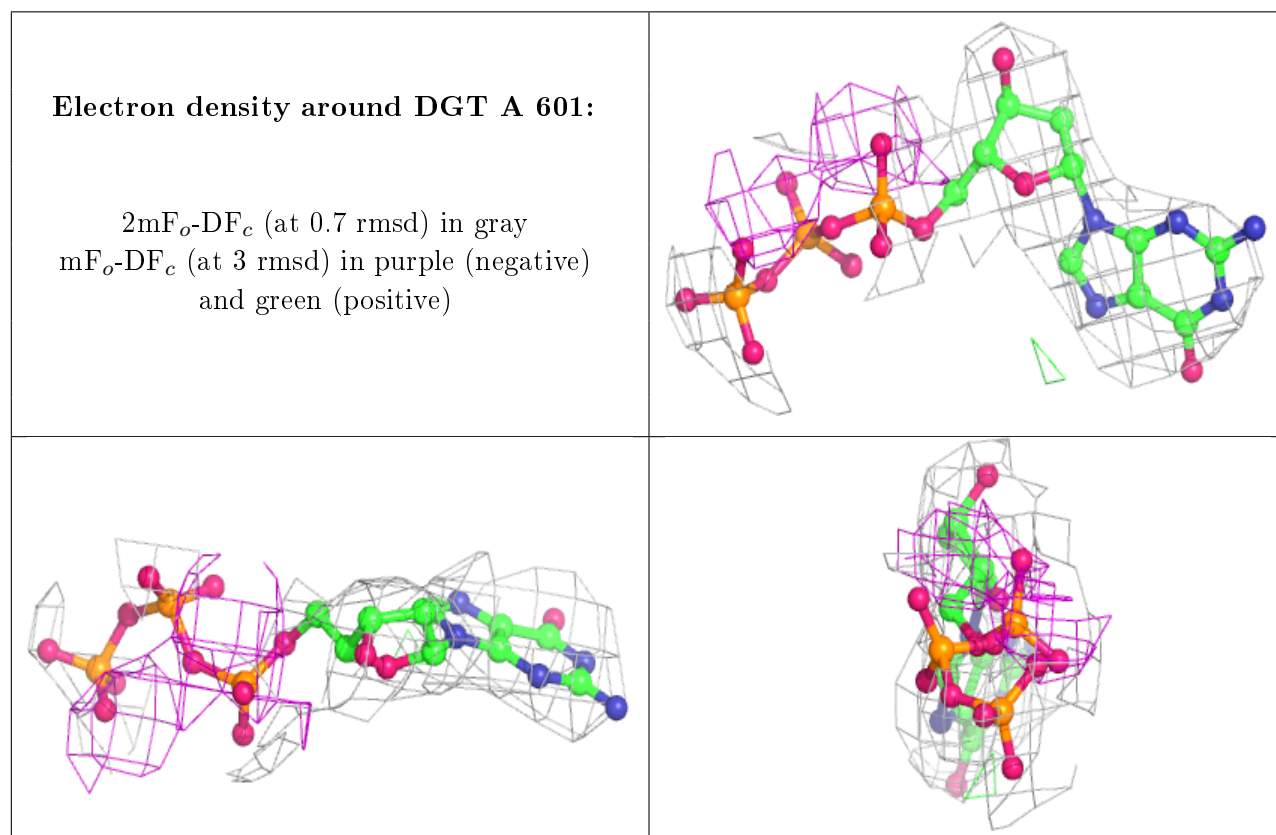
6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

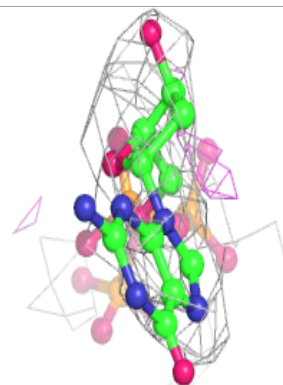
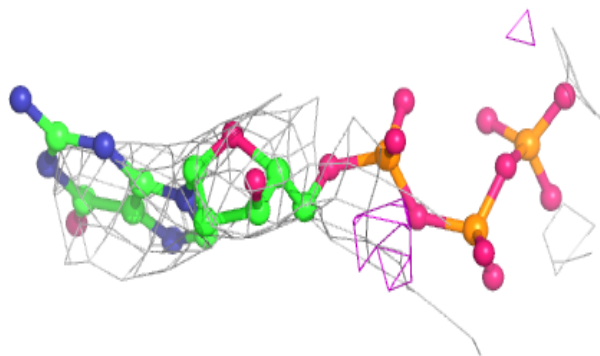
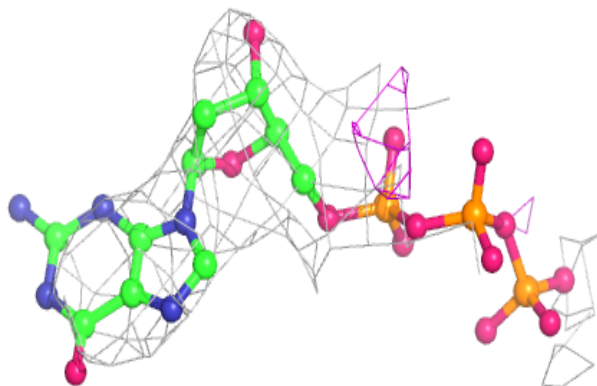
Unable to reproduce the depositors R factor - this section is therefore empty.

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.

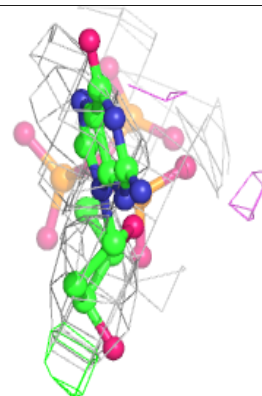
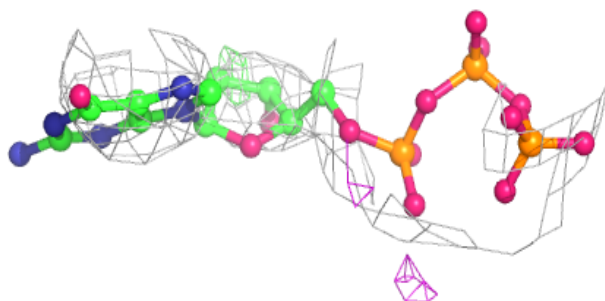
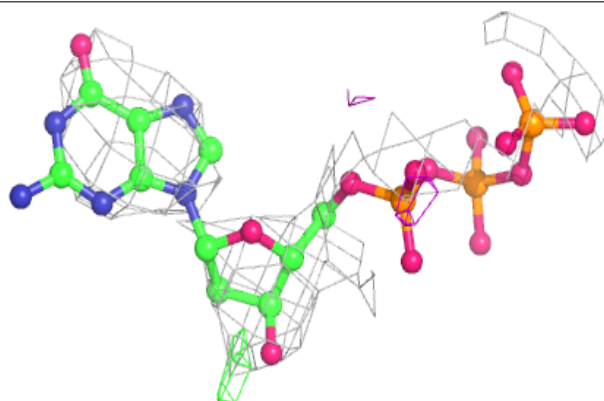


Electron density around DGT C 601:

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

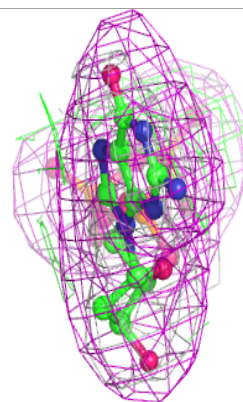
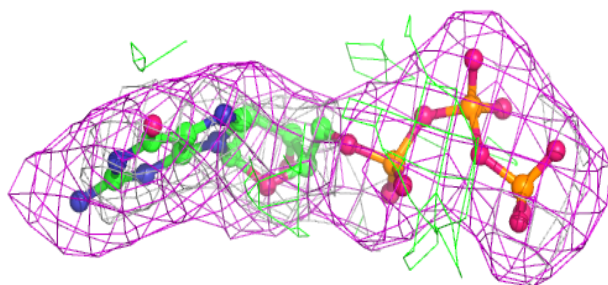
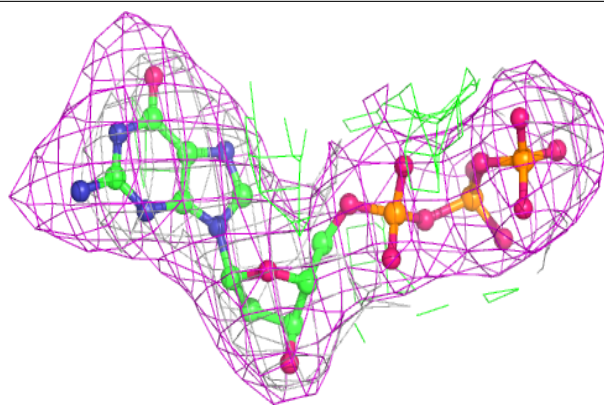
**Electron density around DGT E 601:**

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

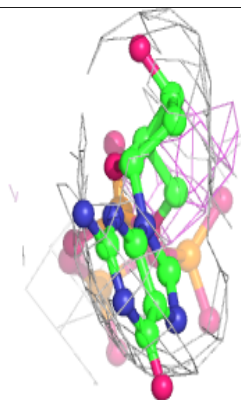
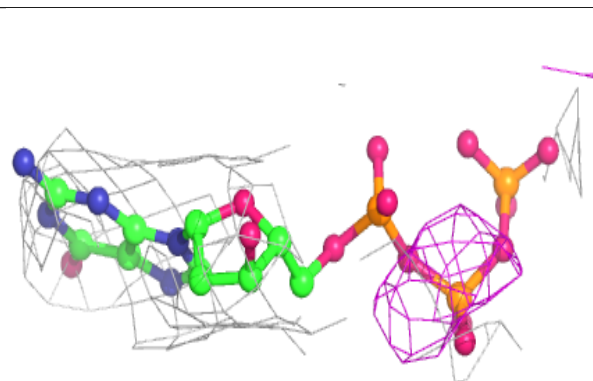
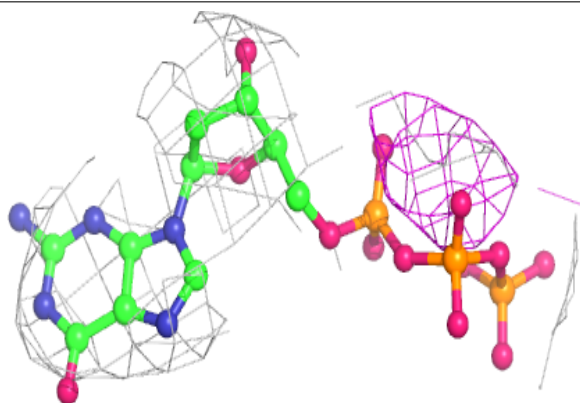


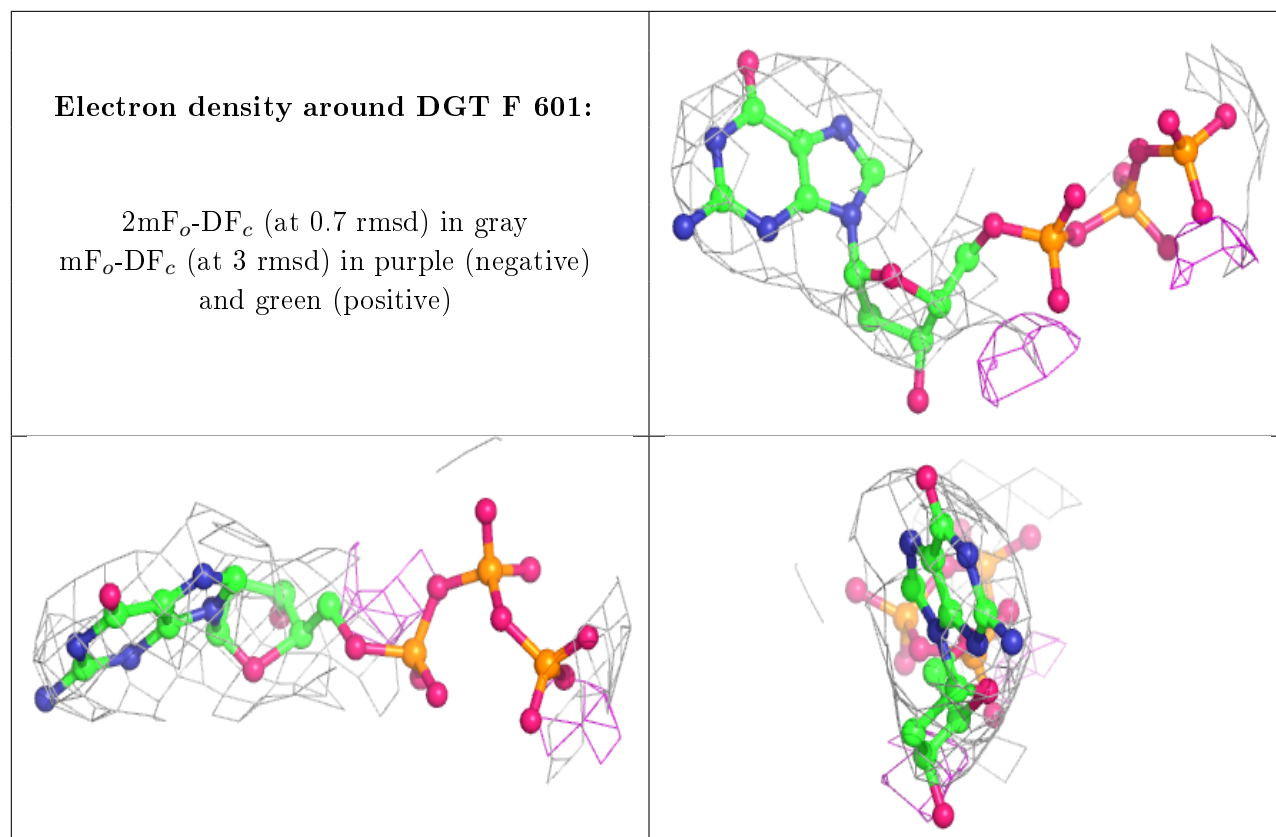
Electron density around DGT D 602:

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

**Electron density around DGT B 601:**

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





6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.