



# Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 05:55 pm BST

PDB ID : 3G6X  
Title : Ternary complex of DNA Polymerase iota:DNA:dGTP with an abasic site at the templating position  
Authors : Nair, D.T.; Aggarwal, A.K.  
Deposited on : 2009-02-09  
Resolution : 2.08 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](mailto: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

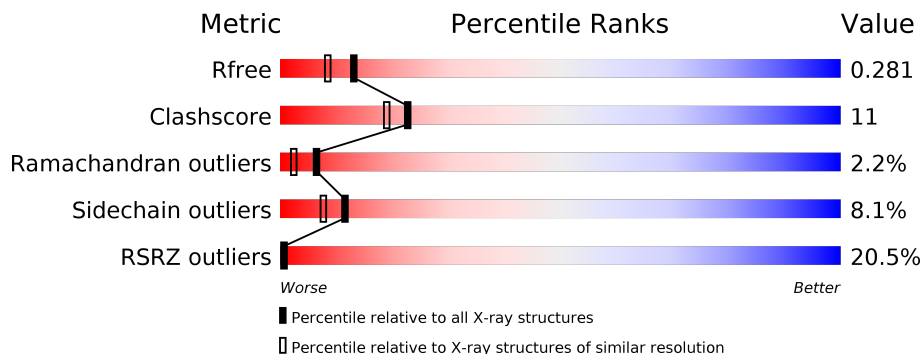
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.08 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	6189 (2.10-2.06)
Clashscore	141614	6738 (2.10-2.06)
Ramachandran outliers	138981	6663 (2.10-2.06)
Sidechain outliers	138945	6664 (2.10-2.06)
RSRZ outliers	127900	6057 (2.10-2.06)

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  $\geq 3$ , 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	420	
2	P	7	
3	T	11	

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 3422 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 DNA polymerase iota.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	373	2876	1809	504	542	21	0	0	0

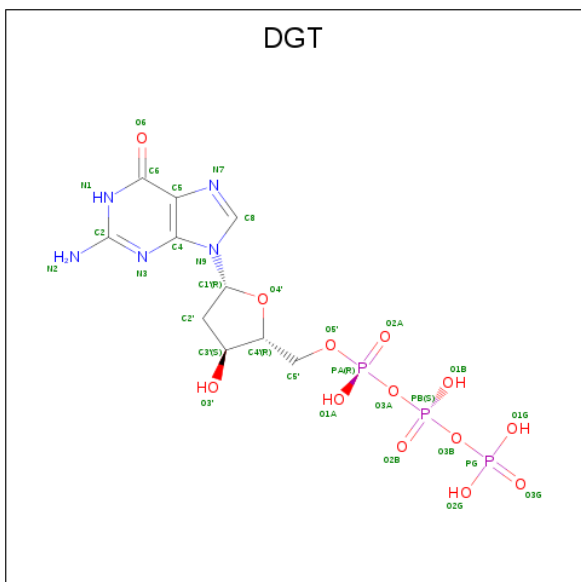
- Molecule 2 is a DNA chain called Primer DNA strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	P	7	139	67	29	37	6	0	0	0

- Molecule 3 is a DNA chain called Template DNA strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	T	8	155	73	25	49	8	0	0	0

- Molecule 4 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DGT) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	31	10	5	13	3	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mg		
5	A	2	2	2	0	0

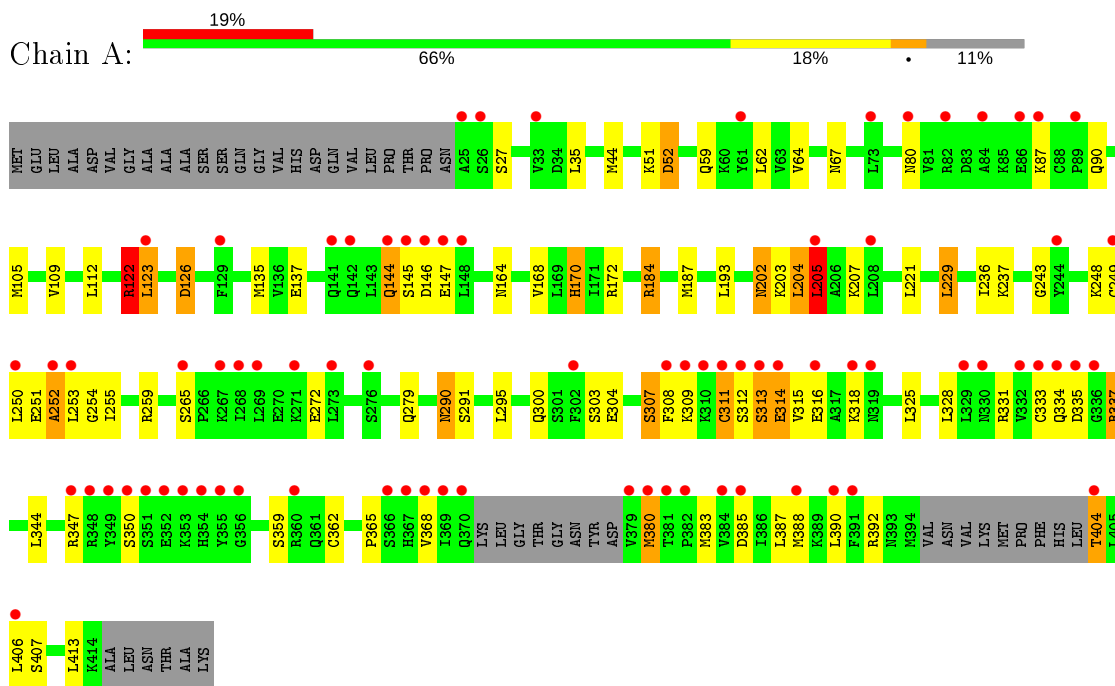
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	198	198	198	0	0
6	P	6	6	6	0	0
6	T	15	15	15	0	0

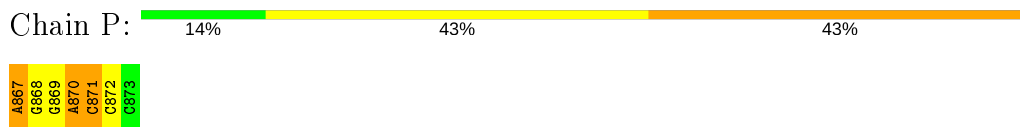
### 3 Residue-property plots [i](#)

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: DNA polymerase iota



- Molecule 2: Primer DNA strand



- Molecule 3: Template DNA strand



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	98.11Å 98.11Å 202.82Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.08 36.61 – 2.08	Depositor EDS
% Data completeness (in resolution range)	96.1 (20.00-2.08) 96.1 (36.61-2.08)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.60 (at 2.08Å)	Xtrriage
Refinement program	REFMAC 5.2.0019, CNS	Depositor
R, $R_{free}$	0.222 , 0.265 0.229 , 0.281	Depositor DCC
$R_{free}$ test set	2813 reflections (8.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.4	Xtrriage
Anisotropy	0.303	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 60.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	3422	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: DOC, MG, 3DR, 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	1.10	7/2914 (0.2%)	1.16	17/3934 (0.4%)
2	P	1.75	2/136 (1.5%)	2.94	20/208 (9.6%)
3	T	1.91	1/160 (0.6%)	2.79	18/245 (7.3%)
All	All	1.19	10/3210 (0.3%)	1.44	55/4387 (1.3%)

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

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	313	SER	CB-OG	9.86	1.55	1.42
1	A	392	ARG	CZ-NH1	9.19	1.45	1.33
3	T	842	DG	C3'-O3'	-6.60	1.35	1.44
1	A	122	ARG	CD-NE	-6.24	1.35	1.46
1	A	184	ARG	CD-NE	-5.89	1.36	1.46
1	A	249	CYS	CB-SG	5.66	1.91	1.82
2	P	868	DG	N1-C2	5.21	1.42	1.37
1	A	122	ARG	CG-CD	5.16	1.64	1.51
2	P	870	DA	N9-C4	5.15	1.41	1.37
1	A	207	LYS	CE-NZ	-5.06	1.36	1.49

All (55) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	122	ARG	NE-CZ-NH2	-28.80	105.90	120.30
1	A	122	ARG	NE-CZ-NH1	19.38	129.99	120.30
1	A	184	ARG	NE-CZ-NH2	-13.11	113.75	120.30
2	P	870	DA	O4'-C1'-N9	12.82	116.97	108.00
3	T	843	DG	O5'-P-OP2	-12.32	94.62	105.70
1	A	122	ARG	CD-NE-CZ	9.57	136.99	123.60
2	P	869	DG	O4'-C1'-N9	9.22	114.45	108.00
3	T	846	DC	O4'-C1'-N1	-8.53	102.03	108.00
3	T	847	DT	C4-C5-C7	8.51	124.10	119.00
2	P	867	DA	O4'-C1'-N9	8.42	113.90	108.00
2	P	867	DA	O4'-C4'-C3'	-8.16	101.10	106.00
3	T	841	DG	C5-C6-N1	8.07	115.54	111.50
2	P	870	DA	N1-C6-N6	-8.04	113.78	118.60
2	P	868	DG	N3-C2-N2	-8.01	114.30	119.90
2	P	869	DG	N1-C6-O6	-7.59	115.34	119.90
2	P	872	DC	O4'-C4'-C3'	-7.52	101.49	106.00
1	A	205	LEU	CB-CG-CD1	7.49	123.74	111.00
1	A	126	ASP	CB-CG-OD1	-7.39	111.65	118.30
1	A	122	ARG	CG-CD-NE	-7.21	96.66	111.80
3	T	847	DT	C6-C5-C7	-7.21	118.58	122.90
1	A	123	LEU	CB-CG-CD2	7.11	123.09	111.00
3	T	845	DC	P-O3'-C3'	6.95	128.04	119.70
3	T	842	DG	O5'-P-OP2	-6.84	99.54	105.70
2	P	870	DA	O5'-P-OP2	-6.82	99.56	105.70
3	T	843	DG	P-O5'-C5'	6.80	131.79	120.90
3	T	841	DG	C5-C6-O6	-6.78	124.53	128.60
3	T	844	DT	C6-C5-C7	-6.74	118.86	122.90
3	T	841	DG	C2-N3-C4	6.55	115.17	111.90
1	A	392	ARG	NE-CZ-NH1	-6.49	117.06	120.30
1	A	135	MET	CG-SD-CE	-6.32	90.09	100.20
3	T	841	DG	O4'-C1'-N9	6.26	112.38	108.00
1	A	184	ARG	NE-CZ-NH1	6.21	123.40	120.30
2	P	868	DG	O4'-C1'-N9	6.20	112.34	108.00
2	P	871	DC	C4'-C3'-C2'	-6.12	97.59	103.10
3	T	842	DG	N1-C6-O6	-6.11	116.23	119.90
1	A	203	LYS	CD-CE-NZ	-5.84	98.27	111.70
2	P	870	DA	O4'-C4'-C3'	5.77	109.46	106.00
2	P	870	DA	C3'-C2'-C1'	5.72	109.36	102.50
2	P	869	DG	C5-C6-O6	5.63	131.98	128.60
3	T	842	DG	N1-C2-N2	-5.60	111.16	116.20
2	P	870	DA	OP1-P-O3'	5.53	117.36	105.20
2	P	870	DA	C2-N3-C4	5.51	113.35	110.60
1	A	337	ARG	NE-CZ-NH1	5.42	123.01	120.30

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	842	DG	P-O5'-C5'	5.41	129.55	120.90
1	A	259	ARG	NE-CZ-NH2	-5.39	117.61	120.30
2	P	870	DA	C4-C5-N7	-5.36	108.02	110.70
2	P	870	DA	N9-C4-C5	5.29	107.92	105.80
3	T	845	DC	P-O5'-C5'	5.18	129.18	120.90
1	A	337	ARG	NE-CZ-NH2	-5.13	117.74	120.30
1	A	204	LEU	CB-CG-CD2	5.12	119.71	111.00
2	P	871	DC	O4'-C4'-C3'	-5.12	102.45	104.50
2	P	871	DC	OP1-P-OP2	-5.10	111.95	119.60
3	T	844	DT	C4-C5-C7	5.08	122.05	119.00
1	A	311	CYS	N-CA-C	5.08	124.72	111.00
3	T	842	DG	O4'-C4'-C3'	-5.08	102.47	104.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	122	ARG	Sidechain

## 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	2876	0	2899	70	0
2	P	139	0	79	3	1
3	T	155	0	88	2	1
4	A	31	0	12	0	0
5	A	2	0	0	0	0
6	A	198	0	0	4	0
6	P	6	0	0	0	0
6	T	15	0	0	0	0
All	All	3422	0	3078	72	1

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 11.

All (72) 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:137:GLU:HG2	6:A:614:HOH:O	1.46	1.12
1:A:164:ASN:H	1:A:170:HIS:HD2	1.11	0.95
1:A:44:MET:HE2	1:A:51:LYS:HA	1.49	0.94
1:A:344:LEU:CD1	1:A:387:LEU:HD22	2.03	0.89
1:A:308:PHE:HE1	1:A:312:SER:HB2	1.38	0.88
1:A:308:PHE:HB2	1:A:311:CYS:HB2	1.62	0.82
1:A:236:ILE:HD12	1:A:250:LEU:HD13	1.63	0.80
1:A:331:ARG:HD2	6:A:597:HOH:O	1.81	0.78
1:A:164:ASN:H	1:A:170:HIS:CD2	2.00	0.77
1:A:344:LEU:HD11	1:A:387:LEU:HD22	1.66	0.77
1:A:44:MET:CE	1:A:67:ASN:HD22	2.00	0.73
1:A:308:PHE:HE1	1:A:312:SER:CB	2.07	0.68
1:A:308:PHE:CE1	1:A:312:SER:HB2	2.27	0.67
1:A:318:LYS:HB2	1:A:388:MET:CE	2.24	0.67
1:A:308:PHE:O	1:A:404:THR:HA	1.94	0.67
1:A:59:GLN:OE1	1:A:64:VAL:HG11	1.97	0.64
1:A:325:LEU:HD23	1:A:380:MET:HE1	1.80	0.62
1:A:304:GLU:HG3	1:A:328:LEU:HG	1.81	0.62
1:A:365:PRO:O	1:A:368:VAL:HG22	1.99	0.62
1:A:112:LEU:HD23	1:A:112:LEU:C	2.20	0.62
1:A:300:GLN:NE2	6:A:512:HOH:O	2.33	0.60
1:A:362:CYS:SG	1:A:390:LEU:HD21	2.41	0.60
1:A:290:ASN:HD22	1:A:290:ASN:C	2.04	0.60
1:A:202:ASN:ND2	1:A:205:LEU:H	2.00	0.58
1:A:313:SER:O	1:A:314:GLU:HB2	2.04	0.58
1:A:318:LYS:HB2	1:A:388:MET:HE1	1.87	0.57
1:A:325:LEU:HD23	1:A:380:MET:CE	2.34	0.57
1:A:236:ILE:HD12	1:A:250:LEU:CD1	2.34	0.56
1:A:335:ASP:OD2	1:A:337:ARG:HD3	2.06	0.56
1:A:221:LEU:HD22	1:A:229:LEU:HD12	1.86	0.55
1:A:44:MET:HE1	1:A:67:ASN:HD22	1.70	0.55
1:A:347:ARG:HD2	1:A:404:THR:CG2	2.37	0.54
1:A:168:VAL:HG22	1:A:172:ARG:HD2	1.91	0.53
1:A:308:PHE:HB2	1:A:311:CYS:CB	2.37	0.53
1:A:308:PHE:CD1	1:A:311:CYS:HB2	2.43	0.53
1:A:308:PHE:HD1	1:A:311:CYS:HB2	1.74	0.52
1:A:347:ARG:NH1	1:A:404:THR:OG1	2.43	0.52
1:A:164:ASN:N	1:A:170:HIS:HD2	1.94	0.52
1:A:105:MET:CG	1:A:193:LEU:HD11	2.40	0.51
1:A:144:GLN:HG3	1:A:147:GLU:OE2	2.11	0.51
1:A:304:GLU:O	1:A:407:SER:HB2	2.10	0.51
1:A:308:PHE:HZ	1:A:406:LEU:HD23	1.76	0.50

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:359:SER:O	2:P:867:DA:H3'	2.12	0.50
1:A:251:GLU:O	1:A:252:ALA:C	2.50	0.49
1:A:307:SER:HB2	3:T:841:DG:OP2	2.13	0.49
1:A:304:GLU:HG2	1:A:331:ARG:HH21	1.78	0.49
1:A:335:ASP:OD1	1:A:337:ARG:HD3	2.13	0.48
3:T:844:DT:H2''	3:T:845:DC:H5'	1.96	0.48
1:A:335:ASP:CG	1:A:337:ARG:HD3	2.35	0.47
1:A:251:GLU:O	1:A:254:GLY:N	2.42	0.47
1:A:105:MET:HE2	1:A:105:MET:HA	1.97	0.47
1:A:248:LYS:HA	1:A:248:LYS:HD3	1.56	0.47
1:A:290:ASN:HD22	1:A:291:SER:N	2.12	0.47
1:A:347:ARG:HD2	1:A:404:THR:HG23	1.97	0.47
1:A:122:ARG:HD3	6:A:425:HOH:O	2.15	0.47
1:A:184:ARG:HA	1:A:187:MET:HE2	1.97	0.46
1:A:44:MET:CE	1:A:67:ASN:ND2	2.76	0.45
1:A:105:MET:HB2	1:A:105:MET:HE3	1.86	0.44
1:A:318:LYS:HB2	1:A:388:MET:HE2	1.99	0.42
1:A:202:ASN:C	1:A:202:ASN:HD22	2.22	0.42
1:A:105:MET:CA	1:A:105:MET:HE2	2.48	0.42
1:A:347:ARG:HD2	1:A:404:THR:HG21	2.00	0.42
1:A:251:GLU:O	1:A:253:LEU:N	2.52	0.42
1:A:236:ILE:HD13	1:A:255:ILE:HG22	2.01	0.41
1:A:243:GLY:CA	2:P:871:DC:OP1	2.69	0.41
1:A:303:SER:O	1:A:331:ARG:NH2	2.53	0.41
1:A:253:LEU:HD11	1:A:272:GLU:HG2	2.03	0.41
1:A:144:GLN:C	1:A:146:ASP:H	2.23	0.41
1:A:325:LEU:HD21	1:A:383:MET:HB3	2.02	0.41
1:A:144:GLN:HG3	1:A:144:GLN:H	1.63	0.40
2:P:870:DA:H2''	2:P:871:DC:O4'	2.21	0.40
1:A:35:LEU:HD11	1:A:109:VAL:HG21	2.02	0.40

All (1) 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)
2:P:867:DA:O5'	3:T:847:DT:O3'[10_665]	1.87	0.33

## 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	367/420 (87%)	341 (93%)	18 (5%)	8 (2%)	6 2

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	315	VAL
1	A	334	GLN
1	A	252	ALA
1	A	314	GLU
1	A	316	GLU
1	A	333	CYS
1	A	52	ASP
1	A	350	SER

### 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	321/376 (85%)	295 (92%)	26 (8%)	11 8

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	27	SER
1	A	52	ASP
1	A	62	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	80	ASN
1	A	87	LYS
1	A	90	GLN
1	A	123	LEU
1	A	126	ASP
1	A	144	GLN
1	A	145	SER
1	A	170	HIS
1	A	202	ASN
1	A	204	LEU
1	A	205	LEU
1	A	229	LEU
1	A	237	LYS
1	A	265	SER
1	A	279	GLN
1	A	290	ASN
1	A	295	LEU
1	A	307	SER
1	A	309	LYS
1	A	380	MET
1	A	385	ASP
1	A	404	THR
1	A	413	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	58	GLN
1	A	67	ASN
1	A	170	HIS
1	A	202	ASN
1	A	262	GLN
1	A	290	ASN
1	A	334	GLN
1	A	393	ASN
1	A	412	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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
3	3DR	T	840	3	8,11,12	1.23	2 (25%)	9,14,17	1.78	3 (33%)
2	DOC	P	873	3,2	14,19,20	0.79	0	13,26,29	2.03	4 (30%)

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	3DR	T	840	3	-	2/3/15/16	0/1/1/1
2	DOC	P	873	3,2	-	0/4/18/19	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	T	840	3DR	O4'-C4'	-2.40	1.40	1.44
3	T	840	3DR	O5'-C5'	-2.04	1.39	1.44

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	T	840	3DR	O4'-C1'-C2'	-3.69	98.91	106.37
2	P	873	DOC	C3'-C2'-C1'	3.59	106.93	102.78
2	P	873	DOC	C2-N3-C4	3.56	119.95	116.34
2	P	873	DOC	O4'-C4'-C3'	3.50	110.61	104.80
2	P	873	DOC	C2'-C1'-N1	-2.48	107.82	112.48
3	T	840	3DR	O3'-C3'-C2'	-2.06	106.64	111.54
3	T	840	3DR	C1'-C2'-C3'	-2.04	100.90	103.20

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	T	840	3DR	C3'-C4'-C5'-O5'
3	T	840	3DR	O4'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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
4	DGT	A	875	5	26,33,33	1.48	4 (15%)	32,52,52	2.14	8 (25%)

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
4	DGT	A	875	5	-	1/18/34/34	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	875	DGT	C6-C5	3.82	1.47	1.41
4	A	875	DGT	C5-C4	3.02	1.48	1.40
4	A	875	DGT	O3'-C3'	2.93	1.49	1.43

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	875	DGT	C8-N7	2.16	1.38	1.34

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	875	DGT	C6-N1-C2	5.64	124.89	115.93
4	A	875	DGT	C5-C6-N1	-4.93	116.68	123.43
4	A	875	DGT	C2-N3-C4	4.17	120.11	115.36
4	A	875	DGT	C6-C5-C4	-3.84	117.13	120.80
4	A	875	DGT	N3-C2-N1	-3.63	122.38	127.22
4	A	875	DGT	O1A-PA-O2A	2.87	126.44	112.24
4	A	875	DGT	N2-C2-N3	2.39	121.69	117.79
4	A	875	DGT	PA-O3A-PB	-2.03	125.86	132.83

There are no chirality outliers.

All (1) torsion outliers are listed below:

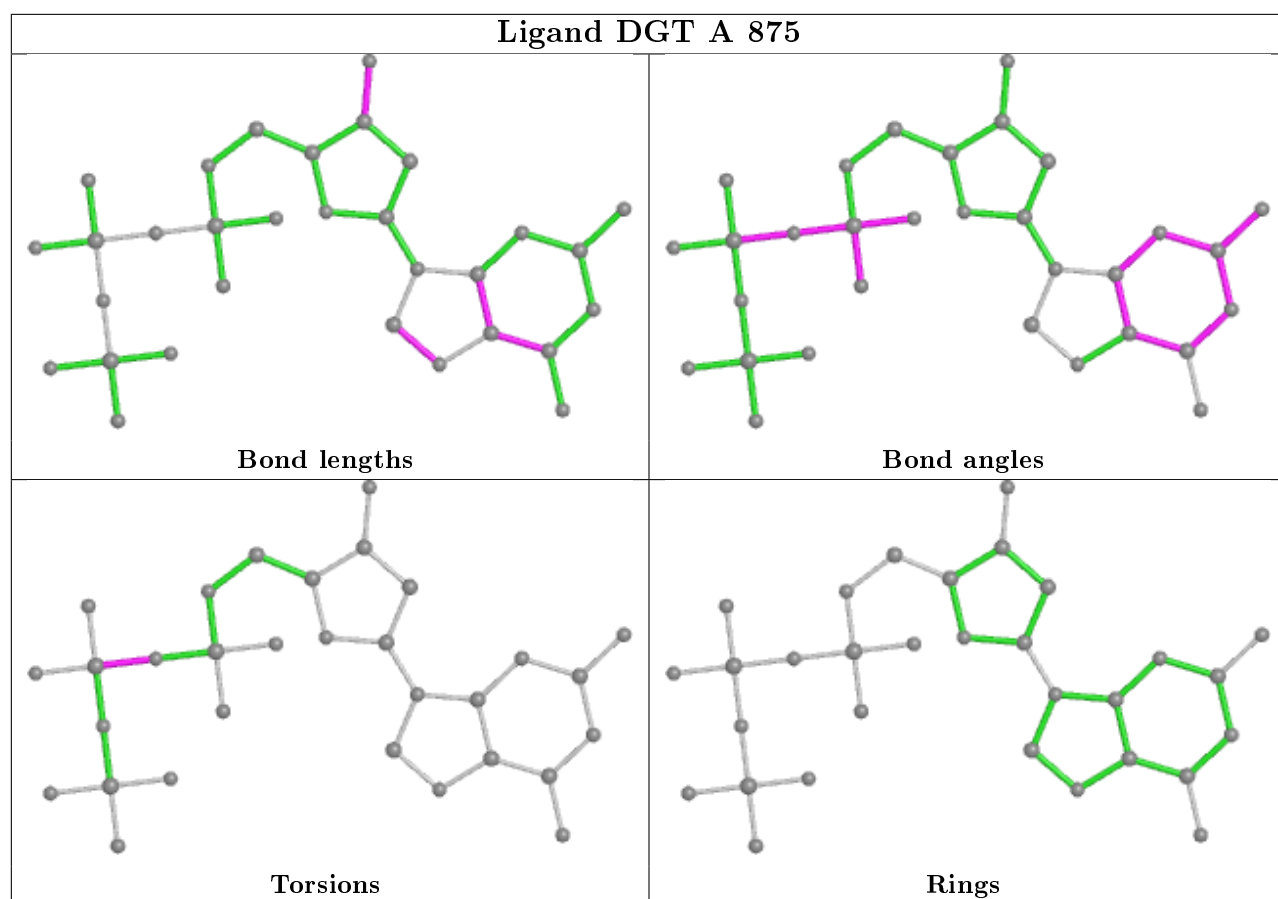
Mol	Chain	Res	Type	Atoms
4	A	875	DGT	PA-O3A-PB-O2B

There are no ring outliers.

No monomer is involved in short contacts.

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

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	373/420 (88%)	1.25	79 (21%) <b>0</b> <b>0</b>	10, 27, 49, 62	0
2	P	6/7 (85%)	0.51	0 <b>100</b> <b>100</b>	23, 29, 33, 35	0
3	T	7/11 (63%)	0.09	0 <b>100</b> <b>100</b>	15, 17, 23, 23	0
All	All	386/438 (88%)	1.22	79 (20%) <b>1</b> <b>1</b>	10, 27, 49, 62	0

All (79) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	354	HIS	13.6
1	A	25	ALA	11.8
1	A	355	TYR	11.1
1	A	351	SER	9.2
1	A	312	SER	8.3
1	A	349	TYR	8.0
1	A	350	SER	7.3
1	A	352	GLU	7.0
1	A	244	TYR	6.9
1	A	336	GLY	6.5
1	A	333	CYS	6.3
1	A	353	LYS	6.1
1	A	368	VAL	5.5
1	A	311	CYS	5.2
1	A	348	ARG	5.2
1	A	310	LYS	5.2
1	A	367	HIS	5.1
1	A	381	THR	4.9
1	A	145	SER	4.8
1	A	332	VAL	4.6
1	A	391	PHE	4.6
1	A	144	GLN	4.5
1	A	249	CYS	4.4

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	A	366	SER	4.4
1	A	390	LEU	4.2
1	A	273	LEU	4.0
1	A	316	GLU	3.9
1	A	308	PHE	3.9
1	A	319	ASN	3.9
1	A	380	MET	3.7
1	A	252	ALA	3.6
1	A	334	GLN	3.6
1	A	385	ASP	3.5
1	A	26	SER	3.4
1	A	384	VAL	3.4
1	A	335	ASP	3.3
1	A	89	PRO	3.2
1	A	268	ILE	3.2
1	A	313	SER	3.2
1	A	82	ARG	3.2
1	A	147	GLU	3.1
1	A	309	LYS	3.1
1	A	84	ALA	3.1
1	A	142	GLN	3.1
1	A	276	SER	3.1
1	A	73	LEU	3.1
1	A	87	LYS	3.0
1	A	86	GLU	2.9
1	A	388	MET	2.9
1	A	61	TYR	2.9
1	A	370	GLN	2.9
1	A	33	VAL	2.8
1	A	141	GLN	2.8
1	A	329	LEU	2.8
1	A	369	ILE	2.7
1	A	406	LEU	2.7
1	A	123	LEU	2.6
1	A	379	VAL	2.6
1	A	129	PHE	2.5
1	A	269	LEU	2.5
1	A	271	LYS	2.5
1	A	347	ARG	2.5
1	A	148	LEU	2.4
1	A	208	LEU	2.4
1	A	80	ASN	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	250	LEU	2.3
1	A	146	ASP	2.3
1	A	382	PRO	2.3
1	A	356	GLY	2.3
1	A	267	LYS	2.2
1	A	205	LEU	2.1
1	A	253	LEU	2.1
1	A	302	PHE	2.1
1	A	404	THR	2.1
1	A	314	GLU	2.1
1	A	265	SER	2.1
1	A	318	LYS	2.1
1	A	360	ARG	2.1
1	A	330	ASN	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	3DR	T	840	11/12	0.95	0.13	22,31,46,47	0
2	DOC	P	873	18/19	0.98	0.18	10,16,19,21	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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<sup>th</sup> 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.

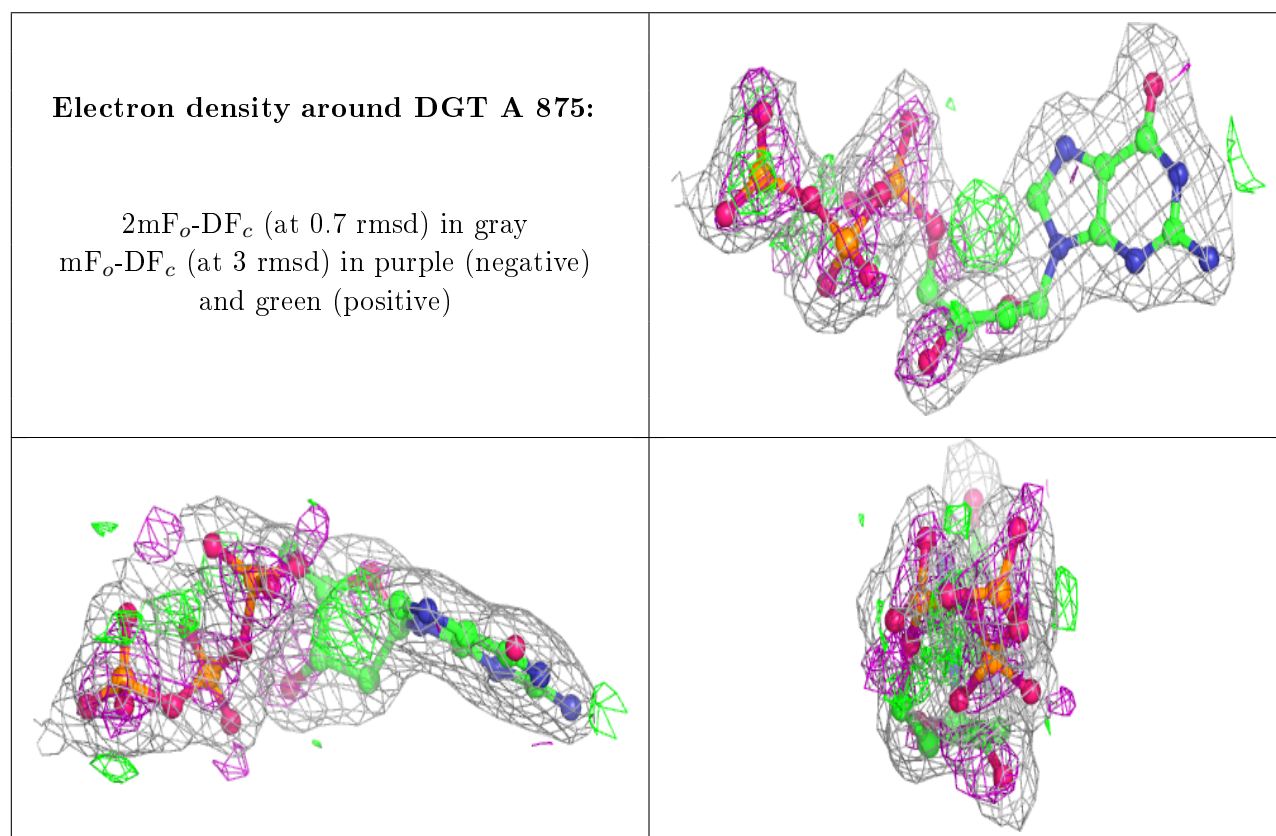
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	MG	A	872	1/1	0.89	0.64	59,59,59,59	0
4	DGT	A	875	31/31	0.97	0.07	8,16,21,22	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	MG	A	871	1/1	0.99	0.04	2,2,2,2	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.



## 6.5 Other polymers [i](#)

There are no such residues in this entry.