



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2021 – 02:21 PM EDT

PDB ID : 7JLG  
Title : Human PrimPol extending from the erroneous primer base A opposite the 8-oxoguanine lesion  
Authors : Rechkoblit, O.; Aggarwal, A.K.  
Deposited on : 2020-07-29  
Resolution : 2.07 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 : 2.20  
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) : 2.20

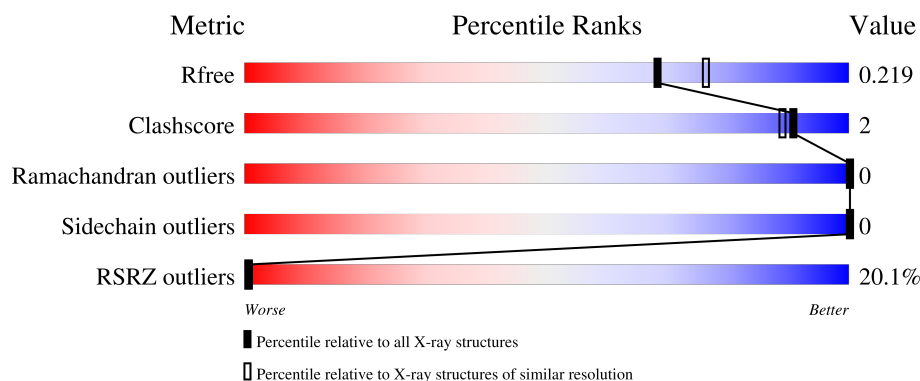
# 1 Overall quality at a glance

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 of 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	354	<div> <div>15%</div> <div>73%</div> <div>24%</div> </div>
1	B	354	<div> <div>11%</div> <div>66%</div> <div>29%</div> </div>
2	C	17	<div> <div>41%</div> <div>71%</div> <div>12%</div> <div>18%</div> </div>
2	G	17	<div> <div>35%</div> <div>88%</div> <div>6%</div> <div>6%</div> </div>
3	D	13	<div> <div>54%</div> <div>85%</div> <div>15%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	H	13	

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
5	CA	A	402	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 9958 atoms, of which 4607 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-directed primase/polymerase protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	268	Total	C	H	N	O	S	0	0	0
			4151	1373	2040	356	373	9			
1	B	250	Total	C	H	N	O	S	0	0	0
			3960	1302	1955	332	363	8			

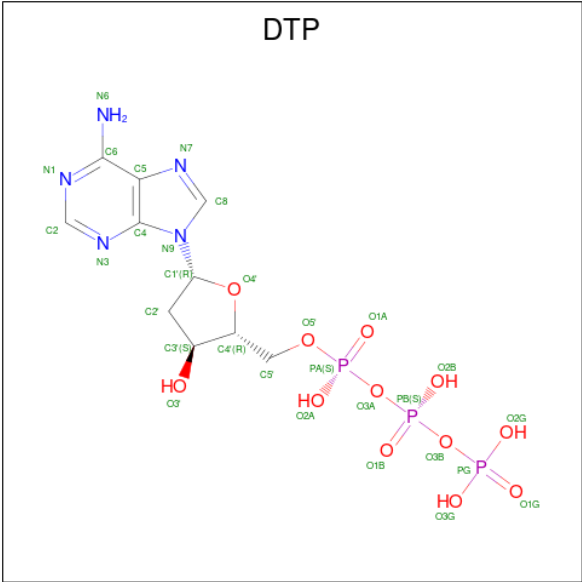
- Molecule 2 is a DNA chain called DNA (5'-D(P\*AP\*TP\*(8OG)P\*CP\*CP\*TP\*AP\*CP\*CP\*AP\*CP\*AP\*CP\*CP\*C)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	14	Total	C	H	N	O	P	0	0	0
			436	133	156	50	83	14			
2	G	16	Total	C	H	N	O	P	0	0	0
			494	151	179	56	93	15			

- Molecule 3 is a DNA chain called DNA (5'-D(P\*GP\*TP\*GP\*TP\*GP\*GP\*TP\*AP\*GP\*GP\*A)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	D	11	Total	C	H	N	O	P	0	0	0
			359	110	125	46	67	11			
3	H	9	Total	C	H	N	O	P	0	0	0
			294	90	102	39	54	9			

- Molecule 4 is 2'-DEOXYADENOSINE 5'-TRIPHOSPHATE (three-letter code: DTP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>12</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
4	A	1	Total	C	H	N	O	P	0	0
			42	10	12	5	12	3		
4	B	1	Total	C	H	N	O	P	0	0
			42	10	12	5	12	3		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Ca	0	0
			2	2		
5	B	2	Total	Ca	0	0
			2	2		

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	B	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	B	1	Total	C	H	O	0	0
			14	3	8	3		
7	B	1	Total	C	H	O	0	0
			14	3	8	3		

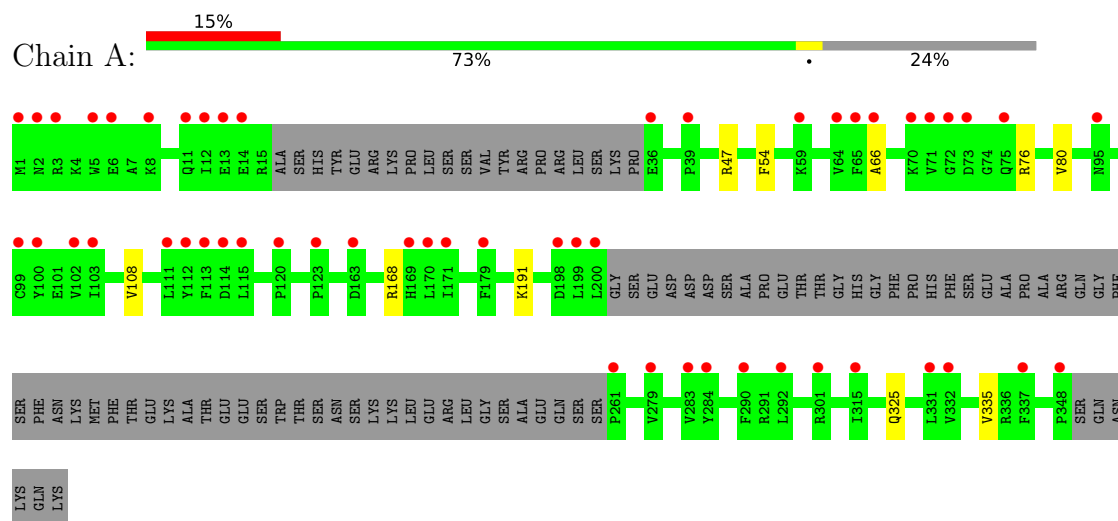
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	48	Total 48	O 48	0	0
8	C	2	Total 2	O 2	0	0
8	B	73	Total 73	O 73	0	0
8	G	6	Total 6	O 6	0	0
8	H	2	Total 2	O 2	0	0

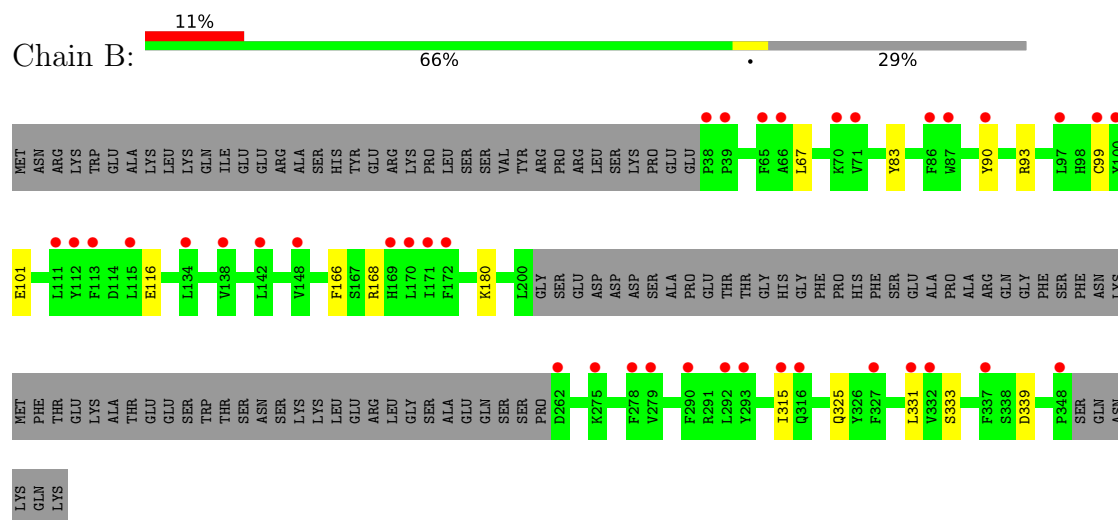
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

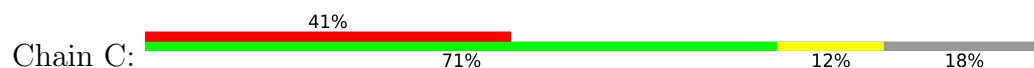
- Molecule 1: DNA-directed primase/polymerase protein



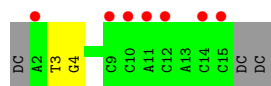
- Molecule 1: DNA-directed primase/polymerase protein



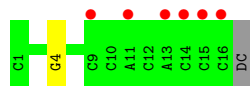
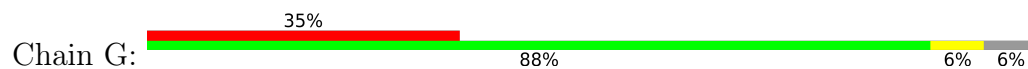
- Molecule 2: DNA (5'-D(P\*AP\*TP\*(8OG)P\*CP\*CP\*TP\*AP\*CP\*CP\*AP\*CP\*AP\*CP\*CP\*C)-3')



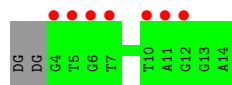
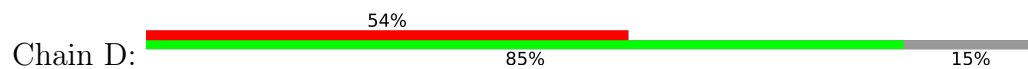




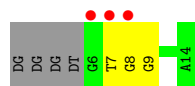
- Molecule 2: DNA (5'-D(P\*AP\*TP\*(8OG)P\*CP\*CP\*TP\*AP\*CP\*CP\*AP\*CP\*AP\*CP\*CP\*C)-3')



- Molecule 3: DNA (5'-D(P\*GP\*TP\*GP\*TP\*GP\*GP\*TP\*AP\*GP\*GP\*A)-3')



- Molecule 3: DNA (5'-D(P\*GP\*TP\*GP\*TP\*GP\*GP\*TP\*AP\*GP\*GP\*A)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.03Å 65.21Å 73.11Å 71.36° 84.31° 88.63°	Depositor
Resolution (Å)	26.81 – 2.07 26.81 – 2.07	Depositor EDS
% Data completeness (in resolution range)	97.0 (26.81-2.07) 97.0 (26.81-2.07)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.18 (at 2.08Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.201 , 0.219 0.201 , 0.219	Depositor DCC
$R_{free}$ test set	2494 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.0	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.40 , 60.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9958	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.52% 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: DTP, PEG, CA, GOL, 8OG

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	0.25	0/2161	0.43	0/2929
1	B	0.26	0/2052	0.44	0/2778
2	C	0.57	0/285	0.91	0/432
2	G	0.58	0/324	0.88	0/492
3	D	0.54	0/263	0.91	0/406
3	H	0.51	0/216	0.87	0/333
All	All	0.34	0/5301	0.57	0/7370

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2111	2040	2031	7	0
1	B	2005	1955	1950	9	0
2	C	280	156	157	1	0
2	G	315	179	180	0	0
3	D	234	125	125	0	0
3	H	192	102	102	2	0
4	A	30	12	12	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	30	12	12	0	0
5	A	2	0	0	0	0
5	B	2	0	0	0	0
6	B	7	10	10	0	0
7	B	12	16	16	0	0
8	A	48	0	0	2	0
8	B	73	0	0	1	0
8	C	2	0	0	0	0
8	G	6	0	0	0	0
8	H	2	0	0	0	0
All	All	5351	4607	4595	17	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 2.

All (17) 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:191:LYS:NZ	1:B:339:ASP:OD2	2.24	0.70
1:A:168:ARG:NH1	8:A:502:HOH:O	2.29	0.60
1:B:315:ILE:O	1:B:325:GLN:NE2	2.41	0.54
1:A:66:ALA:HB1	1:A:76:ARG:HG2	1.94	0.48
1:A:47:ARG:NH1	2:C:3:DT:H3'	2.29	0.48
1:B:168:ARG:NH1	8:B:504:HOH:O	2.36	0.46
1:B:180:LYS:C	1:B:180:LYS:HD3	2.37	0.45
1:A:325:GLN:NE2	8:A:511:HOH:O	2.49	0.45
1:A:54:PHE:HD2	1:A:80:VAL:HG11	1.82	0.44
1:B:101:GLU:OE1	1:B:333:SER:OG	2.35	0.42
3:H:7:DT:H2''	3:H:8:DG:O4'	2.19	0.42
1:B:67:LEU:HD22	1:B:99:CYS:SG	2.60	0.41
1:B:90:TYR:O	1:B:93:ARG:HD2	2.20	0.41
1:B:83:TYR:CE1	1:B:331:LEU:HD13	2.56	0.41
1:B:116:GLU:HA	1:B:166:PHE:O	2.21	0.41
1:A:108:VAL:HG12	1:A:335:VAL:HG21	2.03	0.40
3:H:8:DG:H2'	3:H:9:DG:C1'	2.51	0.40

There are no symmetry-related clashes.

## 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	262/354 (74%)	256 (98%)	6 (2%)	0	100	100
1	B	246/354 (70%)	236 (96%)	10 (4%)	0	100	100
All	All	508/708 (72%)	492 (97%)	16 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	218/320 (68%)	218 (100%)	0	100	100
1	B	215/320 (67%)	215 (100%)	0	100	100
All	All	433/640 (68%)	433 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA

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
2	8OG	C	4	2	18,25,26	6.30	12 (66%)	21,37,40	1.98	5 (23%)
2	8OG	G	4	2,3	18,25,26	6.19	12 (66%)	21,37,40	1.92	4 (19%)

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	8OG	C	4	2	-	2/3/21/22	0/3/3/3
2	8OG	G	4	2,3	-	0/3/21/22	0/3/3/3

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	4	8OG	C8-N7	17.07	1.52	1.34
2	G	4	8OG	C8-N7	16.70	1.52	1.34
2	C	4	8OG	C4-N3	9.03	1.49	1.35
2	G	4	8OG	C4-N3	8.93	1.49	1.35
2	C	4	8OG	C2-N1	7.97	1.49	1.35
2	G	4	8OG	C2-N1	7.81	1.49	1.35
2	C	4	8OG	C3'-C4'	-7.61	1.32	1.53
2	G	4	8OG	C3'-C4'	-7.34	1.32	1.53
2	C	4	8OG	C6-N1	6.87	1.45	1.33
2	C	4	8OG	O4'-C4'	6.83	1.60	1.45
2	G	4	8OG	C6-N1	6.81	1.44	1.33
2	G	4	8OG	O4'-C4'	6.58	1.59	1.45
2	C	4	8OG	C6-C5	6.21	1.52	1.41
2	G	4	8OG	C6-C5	6.08	1.51	1.41
2	C	4	8OG	C2-N2	6.00	1.45	1.33
2	G	4	8OG	C2-N2	5.90	1.45	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	4	8OG	O4'-C1'	-5.14	1.30	1.42
2	C	4	8OG	O4'-C1'	-5.10	1.30	1.42
2	G	4	8OG	O3'-C3'	3.65	1.51	1.43
2	C	4	8OG	O3'-C3'	3.38	1.50	1.43
2	G	4	8OG	O6-C6	-2.36	1.18	1.24
2	C	4	8OG	O6-C6	-2.36	1.18	1.24
2	G	4	8OG	C2-N3	2.17	1.44	1.34
2	C	4	8OG	C2-N3	2.14	1.44	1.34

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	4	8OG	N3-C2-N1	-5.30	120.15	127.22
2	G	4	8OG	N3-C2-N1	-5.18	120.31	127.22
2	C	4	8OG	C2-N3-C4	4.57	120.58	115.36
2	G	4	8OG	C2-N3-C4	4.26	120.22	115.36
2	C	4	8OG	C2'-C1'-N9	-3.25	112.46	116.01
2	G	4	8OG	C5-C6-N1	-2.79	119.62	123.43
2	C	4	8OG	C5-C6-N1	-2.68	119.77	123.43
2	G	4	8OG	C6-N1-C2	2.32	119.61	115.93
2	C	4	8OG	C6-N1-C2	2.27	119.54	115.93

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	4	8OG	O4'-C4'-C5'-O5'
2	C	4	8OG	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 9 ligands modelled in this entry, 4 are monoatomic - leaving 5 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$
6	PEG	B	404	-	6,6,6	0.49	0	5,5,5	0.29	0
7	GOL	B	405	-	5,5,5	0.89	0	5,5,5	0.98	0
4	DTP	A	401	5	26,32,32	0.85	1 (3%)	30,50,50	1.53	5 (16%)
7	GOL	B	406	-	5,5,5	0.88	0	5,5,5	1.01	0
4	DTP	B	401	5	26,32,32	0.83	1 (3%)	30,50,50	1.54	5 (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
6	PEG	B	404	-	-	1/4/4/4	-
7	GOL	B	405	-	-	0/4/4/4	-
4	DTP	A	401	5	-	2/18/34/34	0/3/3/3
7	GOL	B	406	-	-	0/4/4/4	-
4	DTP	B	401	5	-	5/18/34/34	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	401	DTP	C5-C4	2.38	1.47	1.40
4	B	401	DTP	C5-C4	2.27	1.46	1.40

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	401	DTP	C2'-C1'-N9	-3.84	105.41	114.27
4	A	401	DTP	N3-C2-N1	-3.67	122.94	128.68
4	B	401	DTP	N3-C2-N1	-3.67	122.94	128.68
4	A	401	DTP	PB-O3B-PG	-3.36	121.29	132.83
4	B	401	DTP	PB-O3B-PG	-3.26	121.65	132.83
4	A	401	DTP	C2'-C1'-N9	-3.20	106.88	114.27

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	401	DTP	PA-O3A-PB	-2.71	123.53	132.83
4	B	401	DTP	C4-C5-N7	-2.40	106.89	109.40
4	A	401	DTP	C4-C5-N7	-2.35	106.95	109.40
4	B	401	DTP	PA-O3A-PB	-2.34	124.80	132.83

There are no chirality outliers.

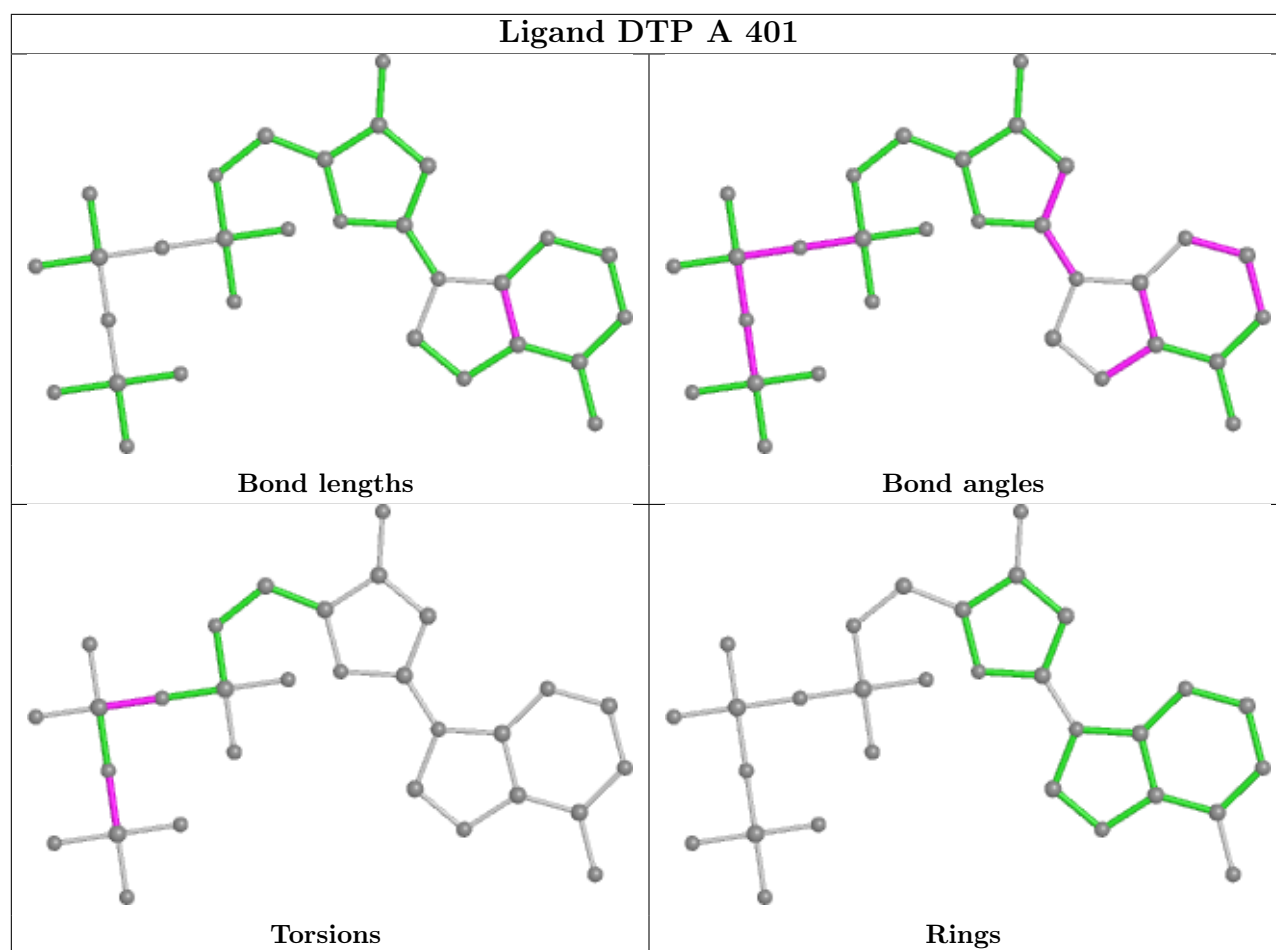
All (8) torsion outliers are listed below:

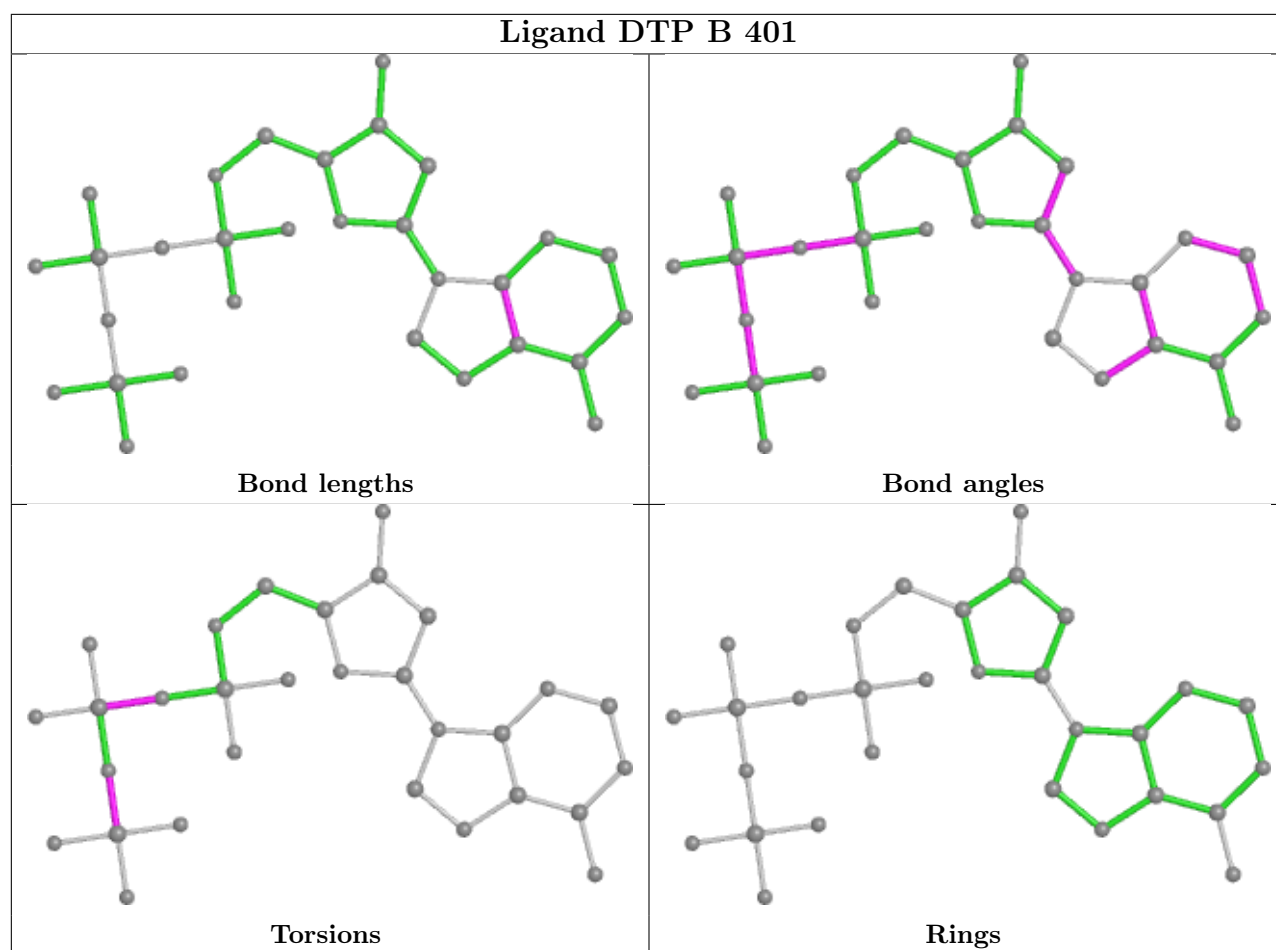
Mol	Chain	Res	Type	Atoms
4	B	401	DTP	PB-O3B-PG-O2G
4	B	401	DTP	PB-O3B-PG-O3G
6	B	404	PEG	O2-C3-C4-O4
4	B	401	DTP	PA-O3A-PB-O2B
4	A	401	DTP	PA-O3A-PB-O2B
4	A	401	DTP	PB-O3B-PG-O3G
4	B	401	DTP	PA-O3A-PB-O1B
4	B	401	DTP	PB-O3B-PG-O1G

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	268/354 (75%)	1.08	53 (19%) <b>1</b> <b>0</b>	41, 63, 116, 158	0
1	B	250/354 (70%)	0.86	38 (15%) <b>2</b> <b>1</b>	37, 51, 86, 107	0
2	C	13/17 (76%)	1.99	7 (53%) <b>0</b> <b>0</b>	78, 114, 170, 171	0
2	G	15/17 (88%)	1.48	6 (40%) <b>0</b> <b>0</b>	56, 102, 176, 187	0
3	D	11/13 (84%)	2.60	7 (63%) <b>0</b> <b>0</b>	120, 139, 169, 178	0
3	H	9/13 (69%)	1.61	3 (33%) <b>0</b> <b>0</b>	89, 119, 156, 158	0
All	All	566/768 (73%)	1.05	114 (20%) <b>1</b> <b>0</b>	37, 57, 138, 187	0

All (114) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	14	GLU	6.4
1	A	337	PHE	6.1
1	A	290	PHE	5.3
2	G	14	DC	5.2
1	A	2	ASN	5.0
3	D	6	DG	4.9
3	H	6	DG	4.7
1	A	113	PHE	4.7
1	A	112	TYR	4.7
1	B	112	TYR	4.7
1	A	1	MET	4.6
2	C	2	DA	4.5
1	B	39	PRO	4.5
1	B	290	PHE	4.4
1	A	13	GLU	4.3
1	B	337	PHE	4.3
1	A	170	LEU	4.3
1	B	331	LEU	4.3
1	A	70	LYS	4.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	71	VAL	4.2
1	A	3	ARG	4.1
1	A	115	LEU	4.0
1	A	284	TYR	4.0
1	A	5	TRP	4.0
2	G	15	DC	3.9
2	C	14	DC	3.9
1	A	100	TYR	3.9
1	B	113	PHE	3.9
1	A	261	PRO	3.8
3	D	4	DG	3.8
3	D	5	DT	3.7
1	A	199	LEU	3.7
1	B	38	PRO	3.7
1	A	11	GLN	3.7
2	C	15	DC	3.7
1	A	111	LEU	3.7
3	D	7	DT	3.6
1	A	331	LEU	3.6
1	A	8	LYS	3.5
1	A	73	ASP	3.5
1	B	115	LEU	3.5
1	A	12	ILE	3.4
1	B	138	VAL	3.4
1	A	114	ASP	3.4
1	A	103	ILE	3.4
1	A	348	PRO	3.4
1	B	170	LEU	3.4
1	B	279	VAL	3.3
1	A	163	ASP	3.3
1	B	66	ALA	3.3
2	C	11	DA	3.3
1	B	97	LEU	3.2
3	D	11	DA	3.2
1	B	315	ILE	3.2
1	B	327	PHE	3.1
1	A	171	ILE	3.1
1	A	332	VAL	3.1
1	A	6	GLU	3.1
1	B	65	PHE	3.1
1	B	348	PRO	3.1
1	B	111	LEU	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	86	PHE	3.1
2	G	13	DA	3.0
1	B	100	TYR	3.0
1	B	292	LEU	3.0
1	A	72	GLY	3.0
1	A	169	HIS	2.9
1	B	70	LYS	2.9
1	B	71	VAL	2.9
1	A	102	VAL	2.9
1	B	316	GLN	2.9
1	A	123	PRO	2.9
1	A	75	GLN	2.9
1	A	292	LEU	2.8
1	B	99	CYS	2.8
1	B	278	PHE	2.8
1	A	279	VAL	2.7
1	A	64	VAL	2.7
2	C	10	DC	2.7
3	D	12	DG	2.7
1	B	332	VAL	2.7
1	B	171	ILE	2.7
1	A	301	ARG	2.7
1	A	315	ILE	2.7
1	A	179	PHE	2.7
1	A	99	CYS	2.7
1	A	200	LEU	2.7
1	A	120	PRO	2.6
1	B	134	LEU	2.6
1	A	283	VAL	2.6
1	B	87	TRP	2.6
2	C	12	DC	2.5
1	B	148	VAL	2.5
1	B	169	HIS	2.5
1	A	66	ALA	2.4
2	G	16	DC	2.4
1	A	59	LYS	2.4
1	A	198	ASP	2.4
3	H	7	DT	2.3
2	G	11	DA	2.3
1	B	90	TYR	2.3
1	A	95	ASN	2.2
2	G	9	DC	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	142	LEU	2.2
1	A	65	PHE	2.2
1	B	262	ASP	2.1
1	B	275	LYS	2.1
3	H	8	DG	2.1
1	A	39	PRO	2.1
3	D	10	DT	2.0
1	A	36	GLU	2.0
1	B	293	TYR	2.0
2	C	9	DC	2.0
1	B	172	PHE	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
2	8OG	C	4	23/24	0.94	0.14	68,83,96,105	0
2	8OG	G	4	23/24	0.96	0.15	46,56,69,76	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, 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.

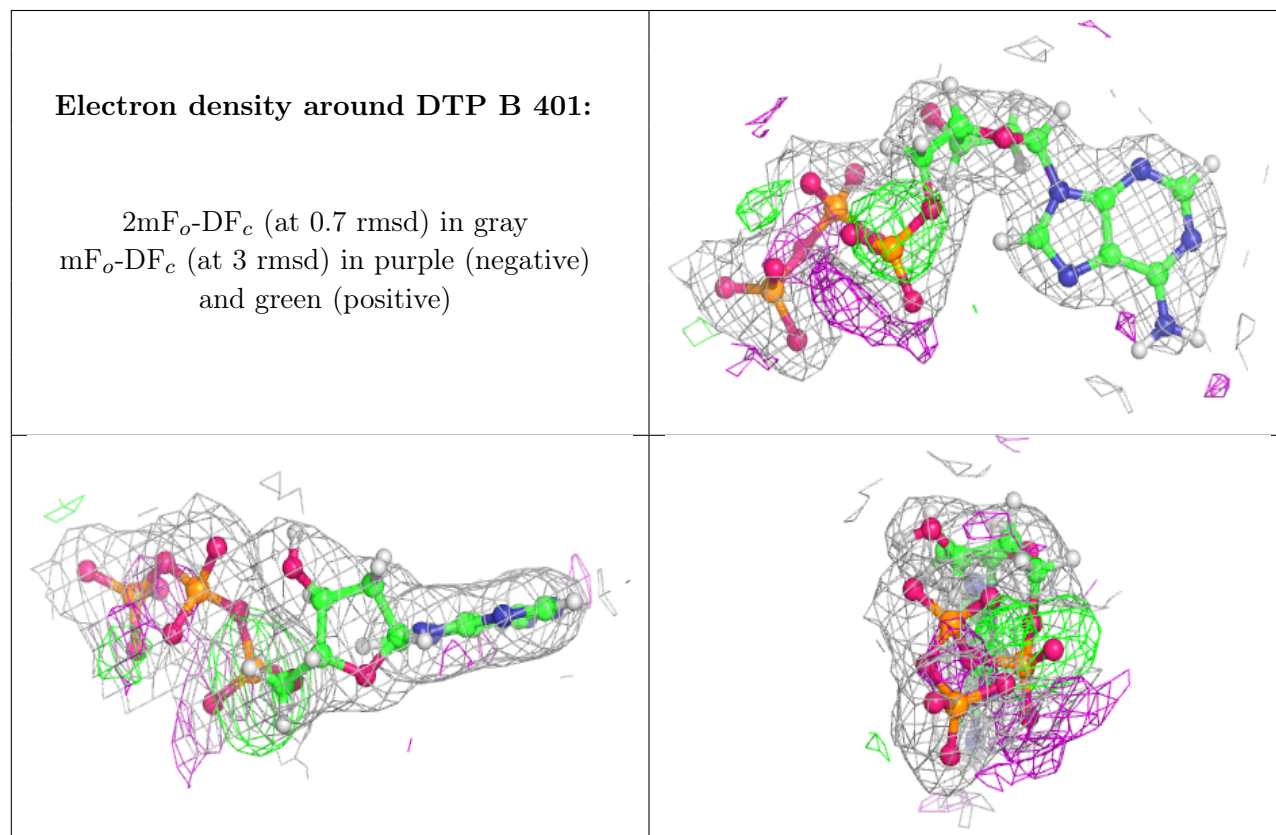
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	CA	A	402	1/1	0.57	0.47	101,101,101,101	0
6	PEG	B	404	7/7	0.72	0.16	73,88,104,104	0
5	CA	A	403	1/1	0.85	0.32	87,87,87,87	0
7	GOL	B	405	6/6	0.89	0.21	68,82,95,99	0
5	CA	B	402	1/1	0.90	0.19	42,42,42,42	0

*Continued on next page...*

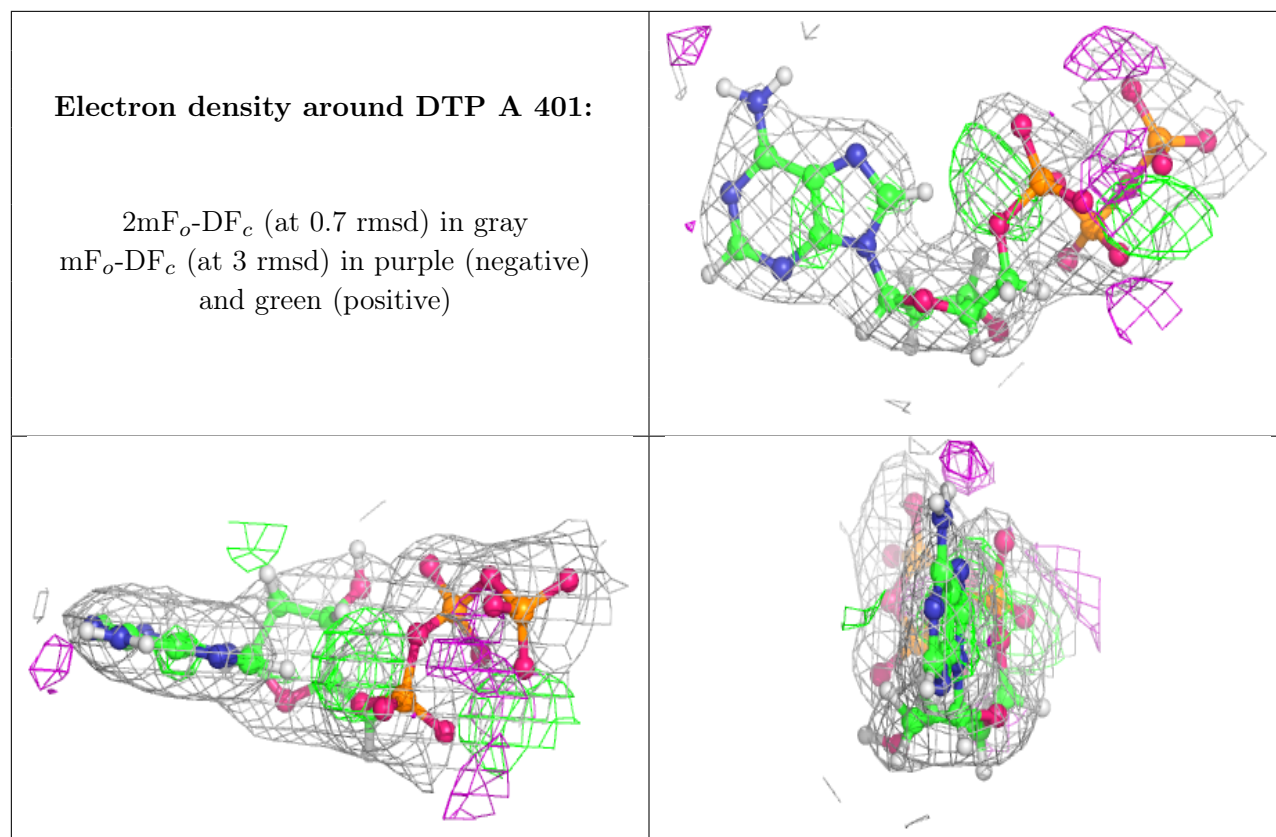
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	CA	B	403	1/1	0.91	0.07	62,62,62,62	0
4	DTP	B	401	30/30	0.92	0.21	33,50,66,75	0
7	GOL	B	406	6/6	0.92	0.27	61,74,79,80	0
4	DTP	A	401	30/30	0.94	0.21	50,66,88,106	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.