



# wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 8, 2021 – 12:16 PM EST

PDB ID : 7KTE  
Title : DNA Polymerase Mu, 8-oxodGTP:Ct Reaction State Ternary Complex, 50 mM Mg<sup>2+</sup> (90min)  
Authors : Jamsen, J.A.; Wilson, S.H.  
Deposited on : 2020-11-24  
Resolution : 1.48 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](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 : 2.24  
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) : 2.24

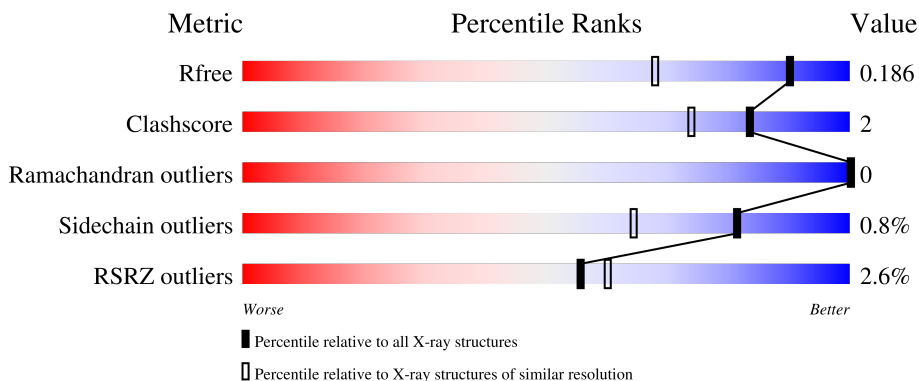
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.48 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	4690 (1.50-1.46)
Clashscore	141614	4955 (1.50-1.46)
Ramachandran outliers	138981	4846 (1.50-1.46)
Sidechain outliers	138945	4844 (1.50-1.46)
RSRZ outliers	127900	4614 (1.50-1.46)

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	356	<div> <div>3%</div> <div>86%</div> <div>6%</div> <div>8%</div> </div>
2	T	9	<div> <div>56%</div> <div>44%</div> </div>
3	P	5	<div> <div>40%</div> <div>40%</div> <div>20%</div> </div>
4	D	4	<div> <div>25%</div> <div>75%</div> </div>

## 2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 3751 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-directed DNA/RNA polymerase mu.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	328	Total	C	N	O	S	0	31	0
			2783	1769	506	497	11			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	127	GLY	-	expression tag	UNP Q9NP87
A	128	SER	-	expression tag	UNP Q9NP87
A	129	ALA	-	expression tag	UNP Q9NP87
A	130	ALA	-	expression tag	UNP Q9NP87
A	131	ALA	-	expression tag	UNP Q9NP87
A	410	GLY	PRO	linker	UNP Q9NP87
A	?	-	GLY	deletion	UNP Q9NP87
A	?	-	ALA	deletion	UNP Q9NP87
A	?	-	ALA	deletion	UNP Q9NP87
A	?	-	VAL	deletion	UNP Q9NP87
A	?	-	GLY	deletion	UNP Q9NP87
A	?	-	GLY	deletion	UNP Q9NP87
A	?	-	SER	deletion	UNP Q9NP87
A	?	-	THR	deletion	UNP Q9NP87
A	?	-	ARG	deletion	UNP Q9NP87
A	?	-	PRO	deletion	UNP Q9NP87
A	?	-	CYS	deletion	UNP Q9NP87
A	?	-	PRO	deletion	UNP Q9NP87

- Molecule 2 is a DNA chain called DNA (5'-D(\*CP\*GP\*GP\*CP\*CP\*TP\*AP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	9	Total	C	N	O	P	0	0	0
			180	86	34	52	8			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	5	Total	C	N	O	P	0	2	0
			123	59	25	34	5			

- Molecule 4 is a DNA chain called DNA (5'-D(P\*GP\*CP\*CP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	4	Total	C	N	O	P	0	2	0
			124	57	24	37	6			

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

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

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Na	0	0
			1	1		

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

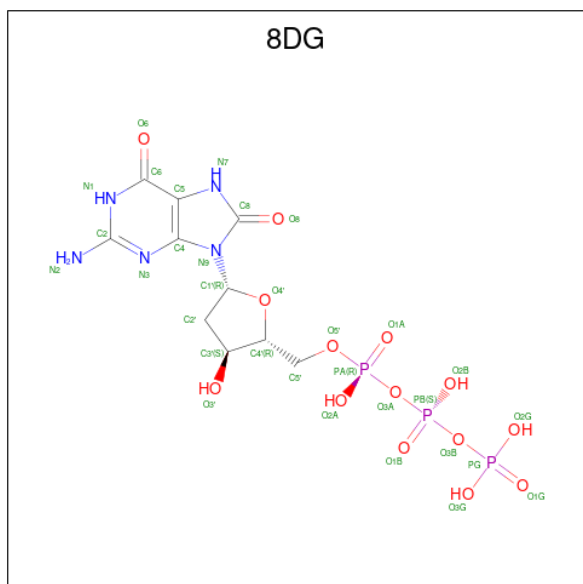
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Cl	0	1
			1	1		

- Molecule 8 is PYROPHOSPHATE (three-letter code: PPV) (formula: H<sub>4</sub>O<sub>7</sub>P<sub>2</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	O	P		0	1
			9	7	2			

- Molecule 9 is 8-OXO-2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: 8DG) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ) (labeled as "Ligand of Interest" by depositor).



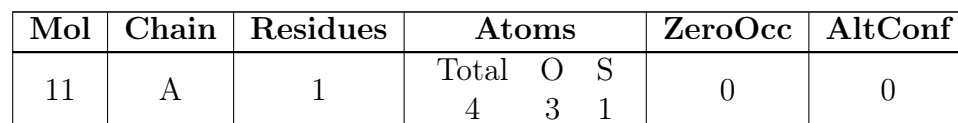
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	A	1	Total	C	N	O	P	
			32	10	5	14	3	1

- Molecule 10 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	A	1	Total	C	O	0	0
			2	1	1		
10	A	1	Total	C	O	0	0
			4	2	2		
10	A	1	Total	C	O	0	0
			3	2	1		
10	A	1	Total	C	O	0	0
			4	2	2		
10	A	1	Total	C	O	0	0
			4	2	2		
10	A	1	Total	C	O	0	0
			4	2	2		
10	P	1	Total	C	O	0	0
			4	2	2		

- Molecule 11 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).



- DTT
- 
- The chemical structure of DTT (1,2-ethanedithiol) is shown. It consists of two carbon atoms, C1 and C2(R), connected by a single bond. C1 is bonded to a hydrogen atom (H1) and a thiol group (S1). C2(R) is bonded to a hydroxyl group (O2) and a thiol group (S2). The thiol groups are represented by yellow sulfur atoms (S1, S2) and white hydrogen atoms (H1, H2). The hydroxyl group is represented by a red oxygen atom (O2) and a white hydrogen atom (H3). The structure is labeled with C1, C2(R), C3(R), and C4, indicating the stereochemistry of the chiral centers.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	1	Total S 1 1	0	0

- 



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
13	T	1	Total	C	O	0	0
			6	3	3		

- Molecule 14 is water.

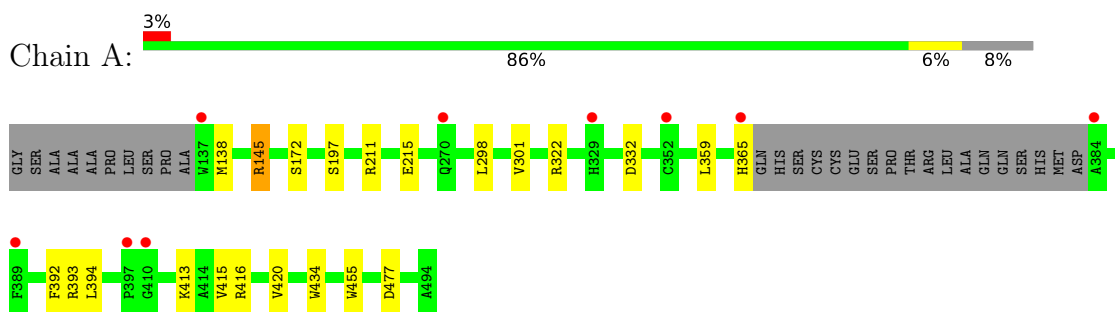
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	366	Total	O	0	28
			374	374		
14	T	43	Total	O	0	2
			45	45		
14	P	19	Total	O	0	0
			19	19		
14	D	22	Total	O	0	2
			22	22		



### 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.

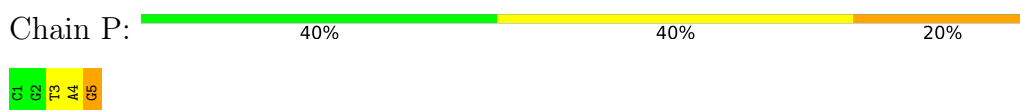
- Molecule 1: DNA-directed DNA/RNA polymerase mu



- Molecule 2: DNA (5'-D(\*CP\*GP\*GP\*CP\*CP\*TP\*AP\*CP\*G)-3')



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



- Molecule 4: DNA (5'-D(P\*GP\*CP\*CP\*G)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	59.94Å 68.52Å 110.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.96 – 1.48 34.96 – 1.48	Depositor EDS
% Data completeness (in resolution range)	99.9 (34.96-1.48) 99.9 (34.96-1.48)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.87 (at 1.48Å)	Xtriage
Refinement program	PHENIX 1.15.2-3472	Depositor
R, $R_{free}$	0.157 , 0.177 0.166 , 0.186	Depositor DCC
$R_{free}$ test set	3831 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	18.1	Xtriage
Anisotropy	0.145	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 50.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3751	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, 8DG, EDO, PPV, DTT, CL, GOL, 8OG, NA, EPE

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	0.64	0/2874	0.83	10/3888 (0.3%)
2	T	1.58	0/201	1.41	2/308 (0.6%)
3	P	1.54	2/111 (1.8%)	1.23	1/167 (0.6%)
4	D	1.32	1/137 (0.7%)	1.15	0/205
All	All	0.81	3/3323 (0.1%)	0.91	13/4568 (0.3%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1	DG	OP3-P	-9.83	1.49	1.61
3	P	3	DT	C5-C7	-6.14	1.46	1.50
3	P	3	DT	C2-O2	-5.12	1.18	1.22

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	322	ARG	NE-CZ-NH2	-10.19	115.21	120.30
1	A	322	ARG	NE-CZ-NH1	9.19	124.90	120.30
1	A	393	ARG	NE-CZ-NH2	-6.73	116.93	120.30
2	T	9	DG	O4'-C1'-N9	-6.72	103.29	108.00
1	A	393	ARG	NE-CZ-NH1	6.64	123.62	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

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	2783	0	2727	12	0
2	T	180	0	102	1	0
3	P	123	0	69	1	0
4	D	124	0	68	1	0
5	A	2	0	0	0	0
6	A	1	0	0	0	0
7	A	1	0	0	1	0
8	A	9	0	0	0	0
9	A	32	0	11	0	0
10	A	21	0	27	0	0
10	P	4	0	6	0	0
11	A	4	0	0	0	0
12	A	1	0	0	0	0
13	T	6	0	8	1	0
14	A	374	0	0	1	0
14	D	22	0	0	0	0
14	P	19	0	0	0	0
14	T	45	0	0	0	0
All	All	3751	0	3018	15	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.

The worst 5 of 15 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:172[B]:SER:HB3	7:A:504[B]:CL:CL	2.35	0.62
1:A:301[A]:VAL:HG11	1:A:359:LEU:HD13	1.86	0.58
1:A:394:LEU:HD22	1:A:415[A]:VAL:HG21	1.94	0.49
2:T:1:DC:H2'	2:T:2:DG:C8	2.49	0.47
1:A:394:LEU:HD22	1:A:415[B]:VAL:HG11	1.97	0.46

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	355/356 (100%)	343 (97%)	12 (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	293/301 (97%)	290 (99%)	3 (1%)	76	54

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	197	SER
1	A	392[A]	PHE
1	A	392[B]	PHE

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 ⓘ

1 non-standard protein/DNA/RNA residue is 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	8OG	P	5[B]	5,3,2	18,25,26	1.34	2 (11%)	21,37,40	2.92	6 (28%)

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	8OG	P	5[B]	5,3,2	-	0/3/21/22	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	P	5[B]	8OG	C6-N1	3.71	1.39	1.33
3	P	5[B]	8OG	C8-N7	3.14	1.37	1.34

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	P	5[B]	8OG	C5-C6-N1	-8.77	111.43	123.43
3	P	5[B]	8OG	C2'-C1'-N9	-6.02	109.43	116.01
3	P	5[B]	8OG	C2-N1-C6	5.55	124.75	115.93
3	P	5[B]	8OG	C2-N3-C4	-3.38	111.50	115.36
3	P	5[B]	8OG	C1'-N9-C4	-2.62	122.11	126.99

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	P	5[B]	8OG	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 16 ligands modelled in this entry, 4 are monoatomic and 1 is modelled with single atom - leaving 11 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
8	PPV	A	505[B]	5	6,8,8	0.67	0	13,13,13	1.51	2 (15%)
10	EDO	A	508	-	3,3,3	0.65	0	2,2,2	0.87	0
10	EDO	A	510	-	3,3,3	0.41	0	2,2,2	0.79	0
10	EDO	A	511	-	3,3,3	0.47	0	2,2,2	0.23	0
9	8DG	A	506[A]	5	26,34,34	1.14	2 (7%)	34,54,54	2.30	8 (23%)
10	EDO	P	101	-	3,3,3	0.56	0	2,2,2	0.41	0
13	GOL	T	101	-	5,5,5	0.86	0	5,5,5	1.12	1 (20%)
11	EPE	A	513	-	0,3,15	-	-	0,3,20	-	-
10	EDO	A	507	-	1,1,3	0.03	0	-	-	-
10	EDO	A	512	-	3,3,3	0.52	0	2,2,2	0.26	0
10	EDO	A	509	-	2,2,3	0.51	0	1,1,2	0.30	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	PPV	A	505[B]	5	-	3/6/6/6	-
10	EDO	A	508	-	-	0/1/1/1	-
10	EDO	A	511	-	-	1/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	8DG	A	506[A]	5	-	4/18/34/34	0/3/3/3
10	EDO	P	101	-	-	0/1/1/1	-
13	GOL	T	101	-	-	0/4/4/4	-
10	EDO	A	510	-	-	1/1/1/1	-
10	EDO	A	512	-	-	0/1/1/1	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	506[A]	8DG	C6-N1	3.75	1.39	1.33
9	A	506[A]	8DG	C8-N7	3.11	1.37	1.34

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	506[A]	8DG	C5-C6-N1	-8.70	111.53	123.43
9	A	506[A]	8DG	C2-N1-C6	5.62	124.85	115.93
9	A	506[A]	8DG	C2'-C1'-N9	-4.31	111.31	116.01
9	A	506[A]	8DG	C2-N3-C4	-3.39	111.48	115.36
8	A	505[B]	PPV	P2-OPP-P1	-3.26	121.65	132.83

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	A	505[B]	PPV	P1-OPP-P2-O12
9	A	506[A]	8DG	PB-O3B-PG-O3G
10	A	511	EDO	O1-C1-C2-O2
10	A	510	EDO	O1-C1-C2-O2
8	A	505[B]	PPV	P1-OPP-P2-O32

There are no ring outliers.

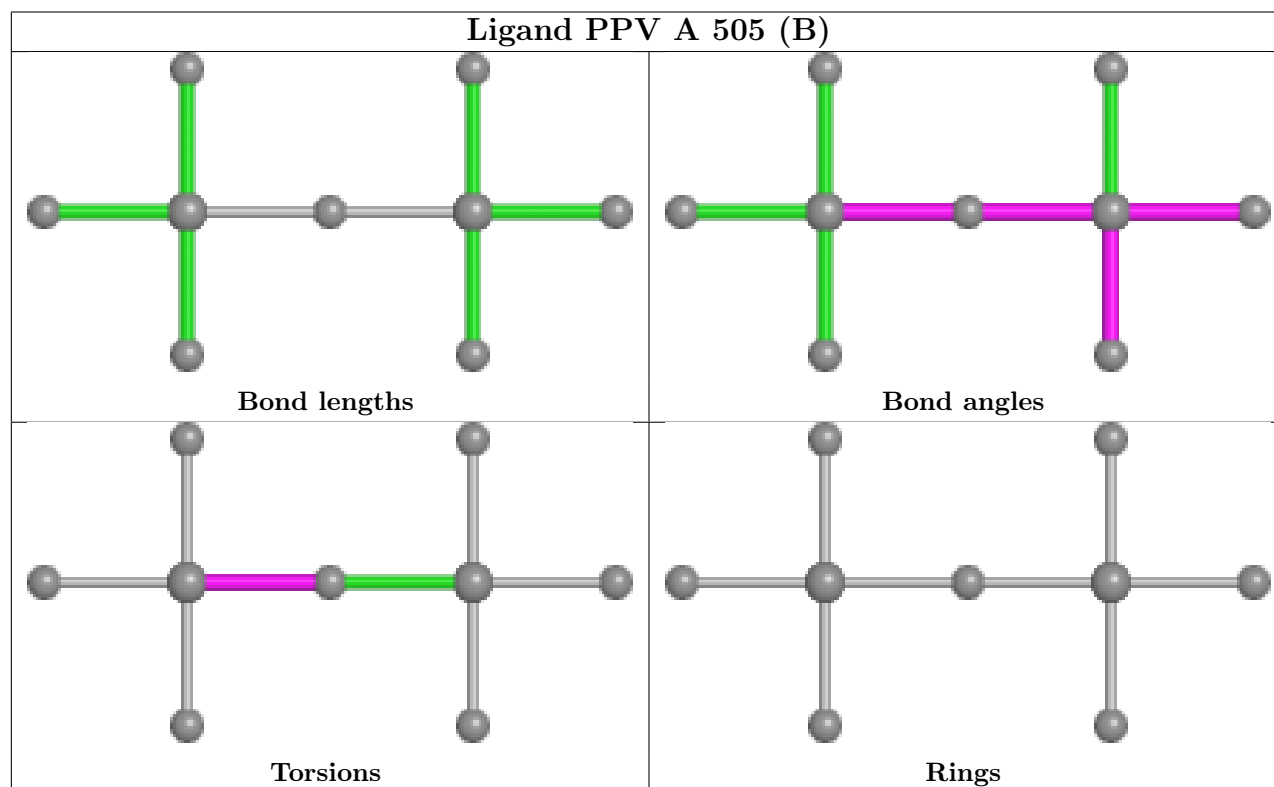
1 monomer is involved in 1 short contact:

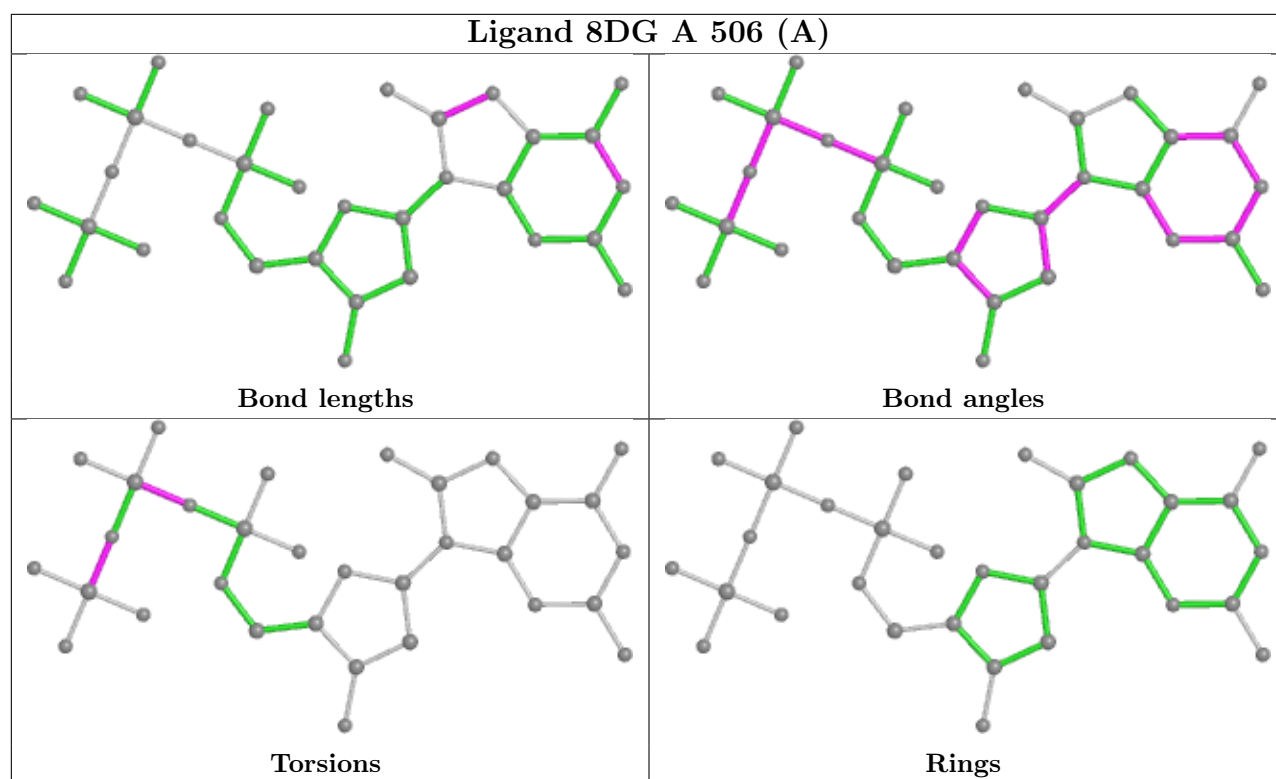
Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	T	101	GOL	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is



within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	328/356 (92%)	-0.18	9 (2%) 54 58	11, 20, 37, 62	0
2	T	9/9 (100%)	-0.68	0 100 100	17, 19, 25, 28	0
3	P	4/5 (80%)	-0.26	0 100 100	17, 18, 20, 21	0
4	D	4/4 (100%)	-0.59	0 100 100	19, 19, 21, 25	0
All	All	345/374 (92%)	-0.20	9 (2%) 56 60	11, 20, 37, 62	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	137	TRP	7.0
1	A	397	PRO	3.5
1	A	270	GLN	2.8
1	A	384	ALA	2.7
1	A	410	GLY	2.7

### 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	8OG	P	5[B]	23/24	0.94	0.09	15,21,23,27	23

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands ⓘ

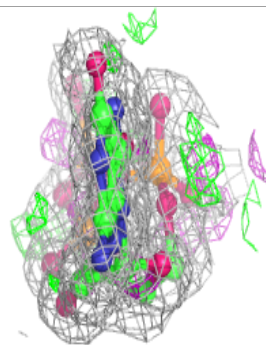
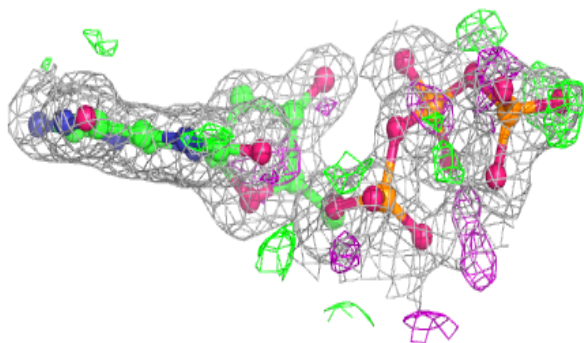
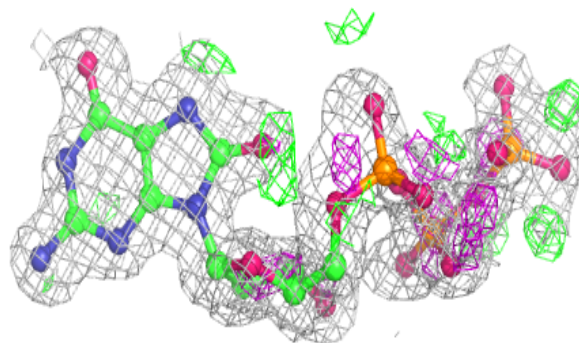
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
13	GOL	T	101	6/6	0.69	0.18	55,59,65,67	0
10	EDO	A	511	4/4	0.71	0.20	48,57,62,62	4
10	EDO	A	509	3/4	0.78	0.14	35,35,40,45	3
10	EDO	A	510	4/4	0.89	0.14	29,31,37,47	4
10	EDO	A	507	2/4	0.90	0.15	14,14,14,17	0
9	8DG	A	506[A]	32/32	0.94	0.09	19,24,28,31	32
8	PPV	A	505[B]	9/9	0.95	0.10	17,21,25,27	9
10	EDO	A	508	4/4	0.95	0.17	20,22,28,39	4
10	EDO	A	512	4/4	0.95	0.08	25,29,34,39	4
10	EDO	P	101	4/4	0.95	0.14	22,27,31,34	0
12	DTT	A	514	1/8	0.95	0.10	54,54,54,54	1
5	MG	A	502	1/1	0.95	0.08	23,23,23,23	0
7	CL	A	504[B]	1/1	0.96	0.11	42,42,42,42	1
11	EPE	A	513	4/15	0.96	0.09	19,25,32,32	4
5	MG	A	501	1/1	0.97	0.06	20,20,20,20	0
6	NA	A	503	1/1	0.99	0.04	18,18,18,18	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.

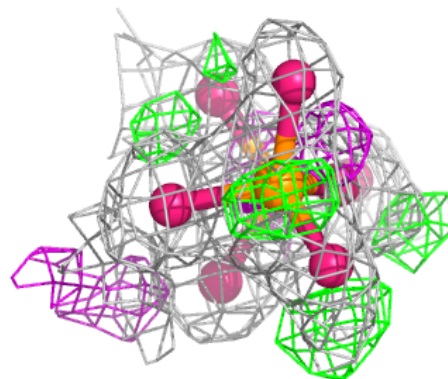
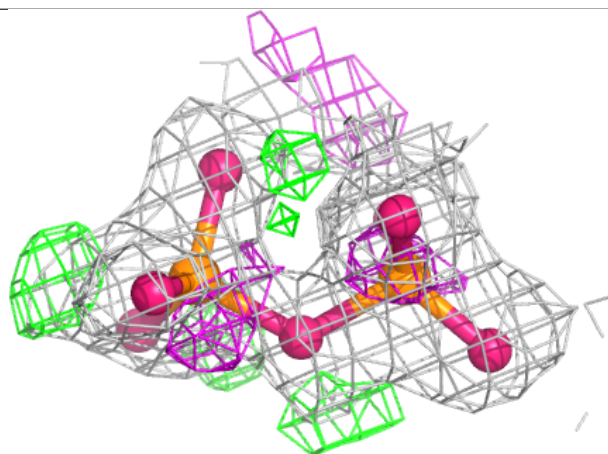
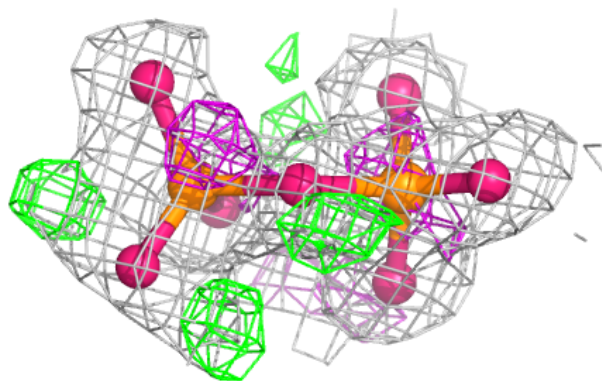
**Electron density around 8DG A 506 (A):**

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



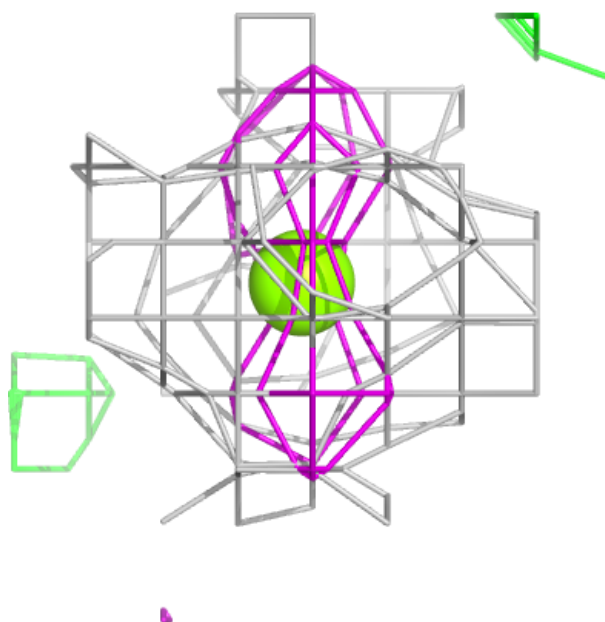
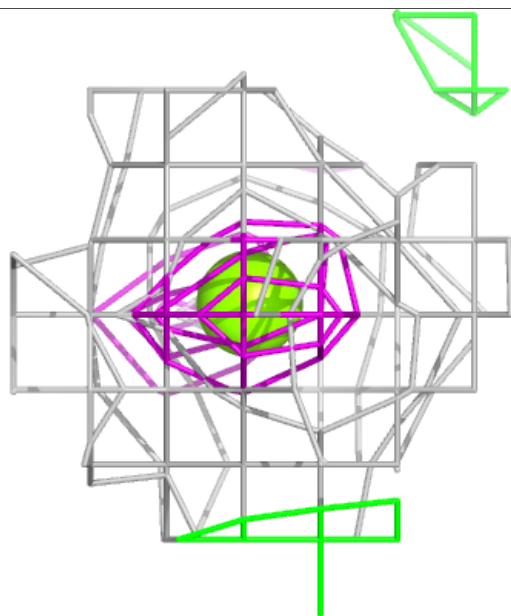
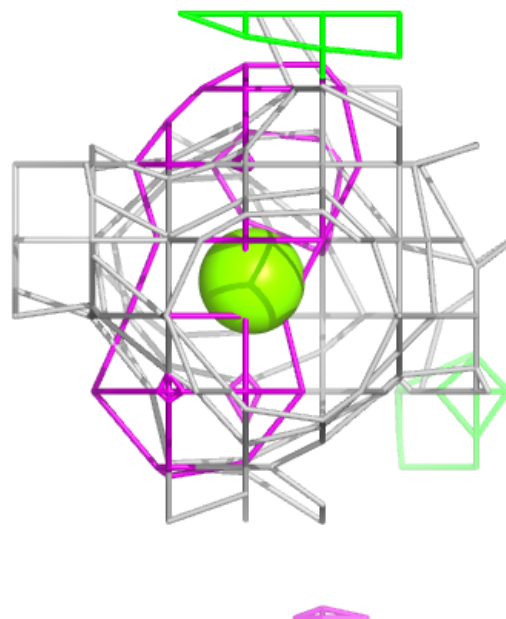
**Electron density around PPV A 505 (B):**

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



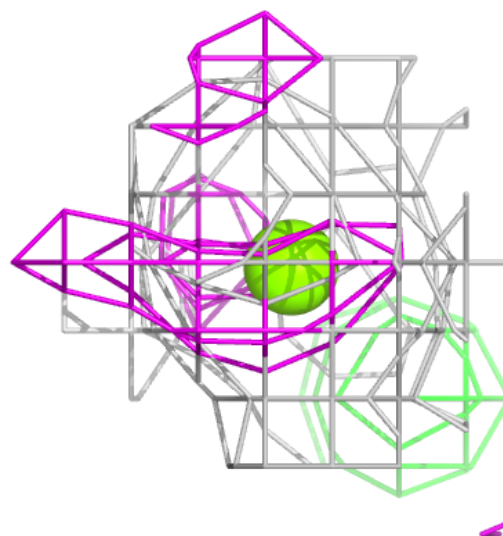
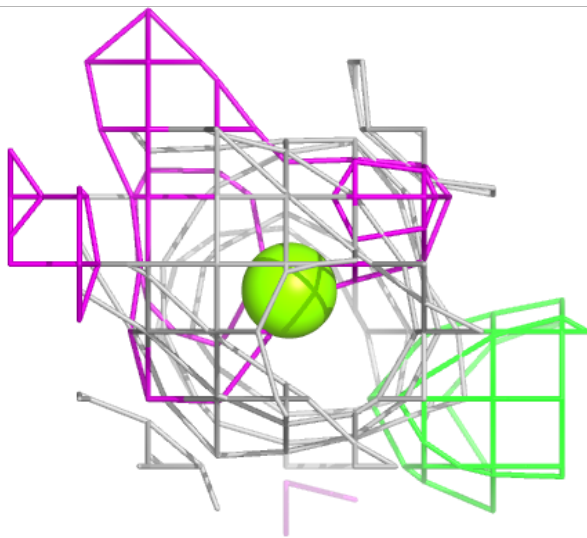
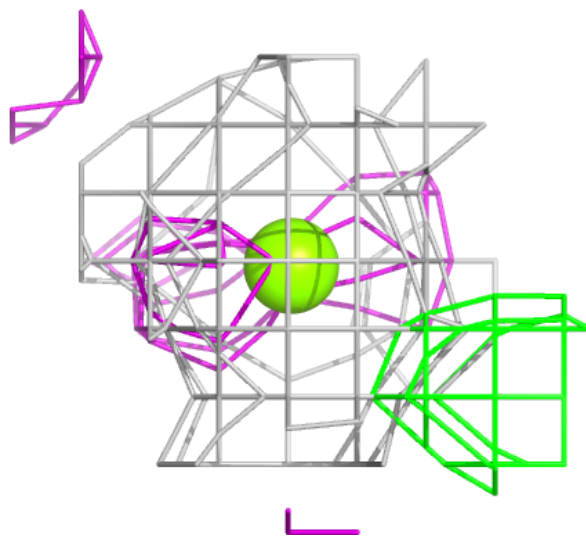
**Electron density around MG A 502:**

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



**Electron density around MG A 501:**

$2mF_o-DF_c$  (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.