



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 9, 2023 – 03:13 AM EDT

PDB ID : 7KT0
Title : DNA Polymerase Mu, dGTP:At Ground State Ternary Complex, 50 mM Mg²⁺ (60min)
Authors : Jamsen, J.A.; Wilson, S.H.
Deposited on : 2020-11-24
Resolution : 1.36 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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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)
Xtrriage (Phenix) : 1.13
EDS : 2.35.1
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) : 2.35.1

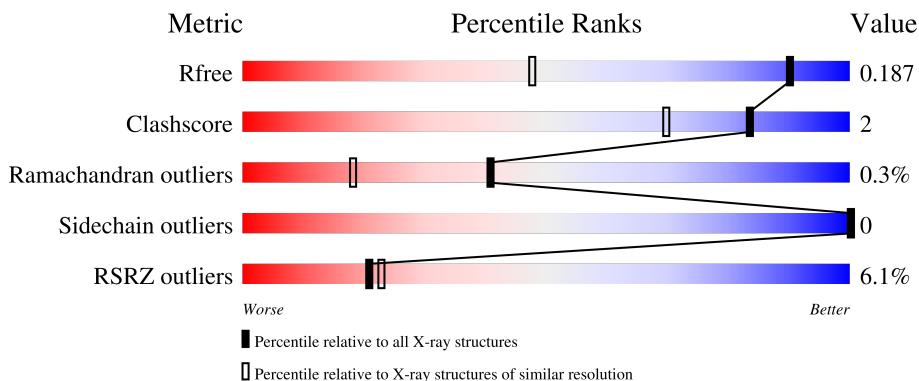
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

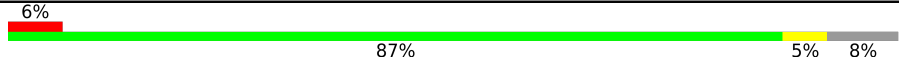



The reported resolution of this entry is 1.36 Å.

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	1509 (1.38-1.34)
Clashscore	141614	1551 (1.38-1.34)
Ramachandran outliers	138981	1530 (1.38-1.34)
Sidechain outliers	138945	1530 (1.38-1.34)
RSRZ outliers	127900	1487 (1.38-1.34)

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 ≥ 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	 6% 87% 5% 8%
2	T	9	 78% 22%
3	P	4	 75% 25%
4	D	4	 50% 50%

2 Entry composition [i](#)

There are 12 unique types of molecules in this entry. The entry contains 3606 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
			Total	C	N	O	S			
1	A	329	2787	1772	507	497	11	0	31	0

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	128	SER	PRO	conflict	UNP Q9NP87
A	129	ALA	ARG	conflict	UNP Q9NP87
A	130	ALA	LYS	conflict	UNP Q9NP87
A	131	ALA	GLY	conflict	UNP Q9NP87
A	?	-	PRO	deletion	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	410	GLY	PRO	conflict	UNP Q9NP87

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	T	9	182	87	36	51	8	0	0	0

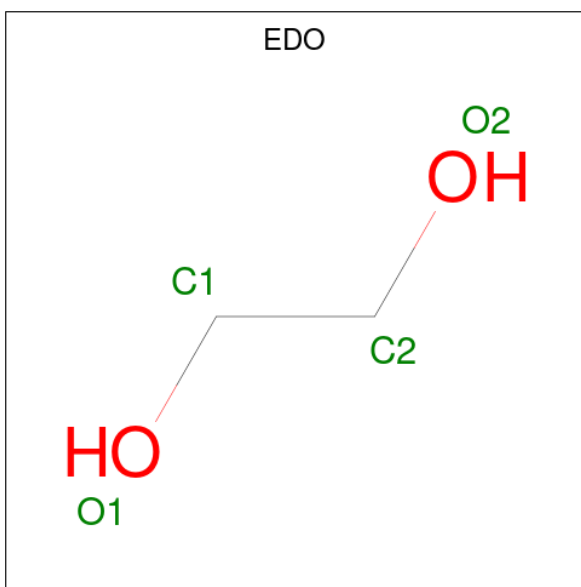
- Molecule 3 is a DNA chain called DNA (5'-D(*CP*GP*TP*A)-3').

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				P
3	P	4	79	39	15	22	3	0	0	0

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

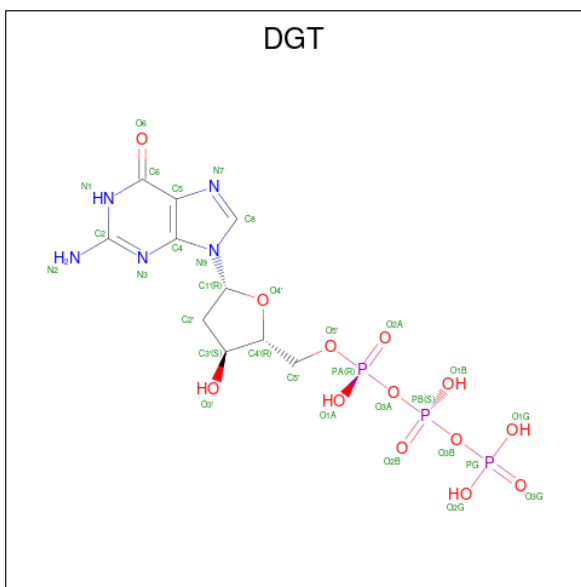
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				P
4	D	4	83	38	16	25	4	0	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

- Molecule 6 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DGT) (formula: C₁₀H₁₆N₅O₁₃P₃) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	O P	0	0
			13	10 3		

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg) (labeled as "Ligand of Interest" by depositor).

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

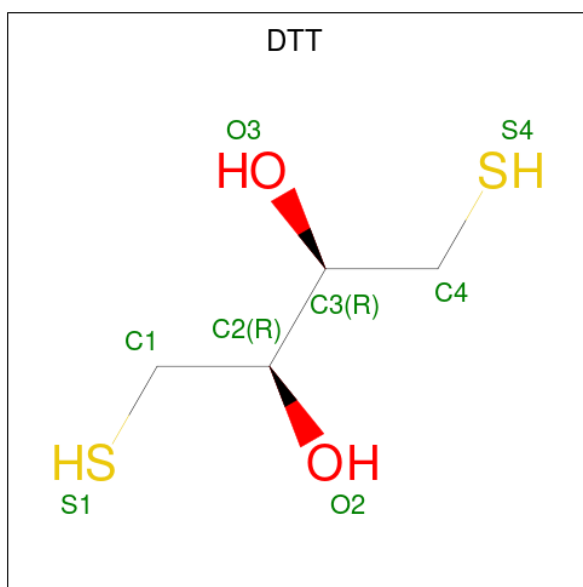
- Molecule 8 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

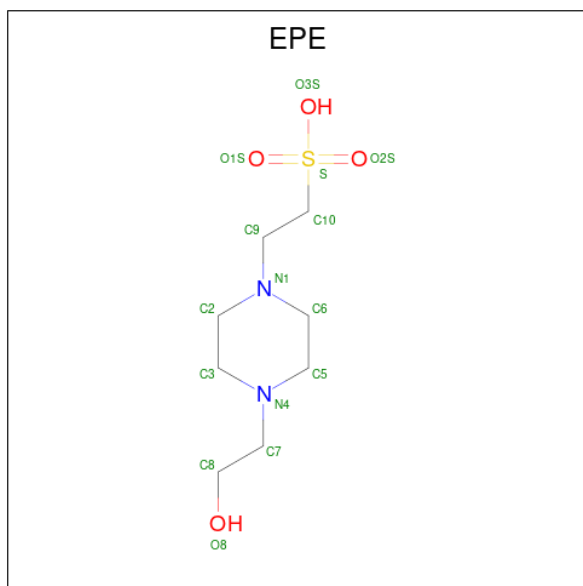
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	Cl	0	1
			1	1		

- Molecule 10 is 2,3-DIHYDROXY-1,4-DITHIOBUTANE (three-letter code: DTT) (formula: C₄H₁₀O₂S₂).



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

- Molecule 11 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
11	A	1	Total O S 4 3 1	0	0

- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	337	Total 351	O 351	0	30
12	T	44	Total 44	O 44	0	0
12	P	19	Total 19	O 19	0	0
12	D	16	Total 17	O 17	0	2

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	59.92Å 68.57Å 110.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.28 – 1.36 34.28 – 1.36	Depositor EDS
% Data completeness (in resolution range)	99.7 (34.28-1.36) 99.7 (34.28-1.36)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.97 (at 1.36Å)	Xtrriage
Refinement program	PHENIX 1.15.2-3472	Depositor
R, R_{free}	0.161 , 0.179 0.170 , 0.187	Depositor DCC
R_{free} test set	4883 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	18.7	Xtrriage
Anisotropy	0.177	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 49.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3606	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.07% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, EPE, DGT, DTT, MG, CL, NA

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	0.40	0/2877	0.67	3/3890 (0.1%)
2	T	0.94	0/204	1.05	0/313
3	P	1.10	0/88	1.03	0/134
4	D	1.36	1/92 (1.1%)	1.05	1/138 (0.7%)
All	All	0.53	1/3261 (0.0%)	0.73	4/4475 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1	DG	OP3-P	-9.64	1.49	1.61

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	393	ARG	NE-CZ-NH2	-5.38	117.61	120.30
1	A	387	ARG	NE-CZ-NH1	5.25	122.92	120.30
1	A	393	ARG	NE-CZ-NH1	5.11	122.86	120.30
4	D	2	DC	C1'-O4'-C4'	-5.09	105.01	110.10

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	2787	0	2731	13	0
2	T	182	0	102	1	0
3	P	79	0	47	1	0
4	D	83	0	45	0	0
5	A	22	0	30	0	0
6	A	13	0	0	0	0
7	A	2	0	0	0	0
8	A	1	0	0	0	0
9	A	1	0	0	1	0
10	A	1	0	0	0	0
11	A	4	0	0	0	0
12	A	351	0	0	1	0
12	D	17	0	0	0	0
12	P	19	0	0	0	0
12	T	44	0	0	0	0
All	All	3606	0	2955	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.

All (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	9:A:511[B]:CL:CL	2.38	0.60
1:A:266:ARG:HG3	1:A:280:LEU:HD21	1.88	0.54
1:A:301[B]:VAL:HG11	1:A:359:LEU:HD13	1.89	0.53
1:A:360:ILE:HG21	1:A:363:HIS:CD2	2.43	0.53
12:A:626:HOH:O	3:P:4:DA:H5''	2.09	0.52
1:A:420[A]:VAL:HG21	1:A:434:TRP:CE2	2.45	0.51
1:A:394:LEU:HD23	1:A:413[A]:LYS:HE3	1.94	0.48
1:A:390[B]:CYS:SG	1:A:392[B]:PHE:CD1	3.07	0.47
2:T:1:DC:H2'	2:T:2:DG:C8	2.50	0.47
1:A:390[B]:CYS:HG	1:A:392[B]:PHE:HD1	1.64	0.44
1:A:420[A]:VAL:HG21	1:A:434:TRP:CZ2	2.53	0.44
1:A:390[B]:CYS:SG	1:A:392[B]:PHE:HD1	2.42	0.43
1:A:354:LEU:HD13	1:A:392[B]:PHE:HZ	1.85	0.42
1:A:298:LEU:O	1:A:301[A]:VAL:HG22	2.21	0.41
1:A:330:ASP:OD1	1:A:330:ASP:C	2.59	0.41

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	356/356 (100%)	347 (98%)	8 (2%)	1 (0%)	41 18

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	411	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	291/301 (97%)	291 (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. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	242	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 13 ligands modelled in this entry, 4 are monoatomic and 1 is modelled with single atom - leaving 8 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
5	EDO	A	501	-	3,3,3	0.44	0	2,2,2	0.42	0
5	EDO	A	503	-	3,3,3	0.58	0	2,2,2	0.59	0
5	EDO	A	504	-	1,1,3	0.04	0	-		
11	EPE	A	513	-	0,3,15	-	-	0,3,20	-	-
6	DGT	A	507	7	8,12,33	1.44	1 (12%)	15,20,52	1.73	5 (33%)
5	EDO	A	505	-	3,3,3	0.46	0	2,2,2	0.35	0
5	EDO	A	502	-	3,3,3	0.43	0	2,2,2	0.41	0
5	EDO	A	506	-	3,3,3	0.48	0	2,2,2	0.25	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
5	EDO	A	501	-	-	0/1/1/1	-
5	EDO	A	503	-	-	0/1/1/1	-
6	DGT	A	507	7	-	4/12/12/34	-
5	EDO	A	505	-	-	1/1/1/1	-
5	EDO	A	502	-	-	0/1/1/1	-
5	EDO	A	506	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	507	DGT	PA-O2A	3.80	1.62	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	507	DGT	O5'-PA-O1A	2.73	118.09	107.64
6	A	507	DGT	PB-O3B-PG	-2.67	123.67	132.83
6	A	507	DGT	O1B-PB-O2B	2.40	124.10	112.24
6	A	507	DGT	O1G-PG-O3B	2.14	111.83	104.64
6	A	507	DGT	O2G-PG-O3G	2.05	118.70	110.68

There are no chirality outliers.

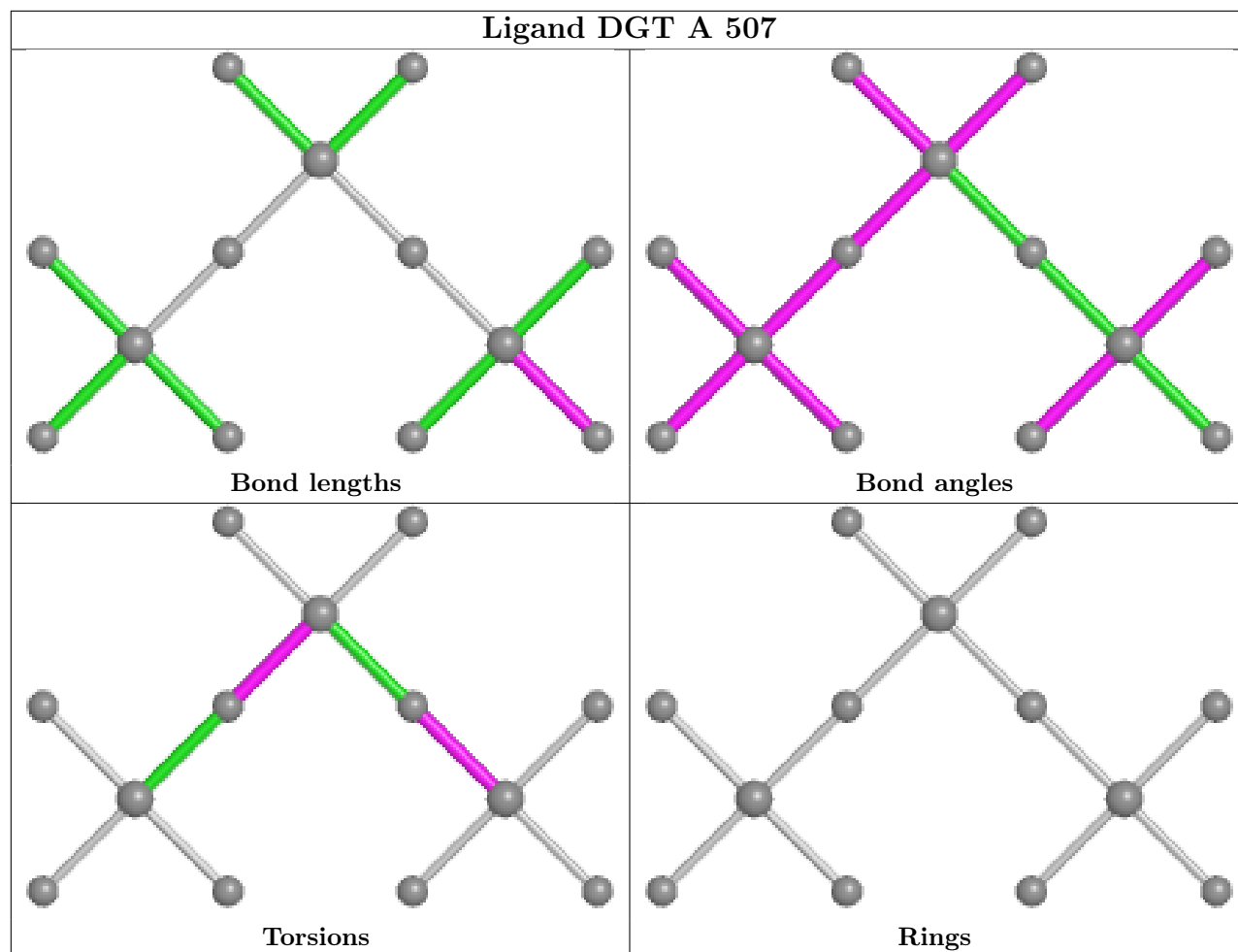
All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	507	DGT	PB-O3A-PA-O1A
6	A	507	DGT	PB-O3A-PA-O5'
5	A	505	EDO	O1-C1-C2-O2
6	A	507	DGT	PG-O3B-PB-O2B
6	A	507	DGT	PB-O3A-PA-O2A

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	329/356 (92%)	0.19	21 (6%) 19 21	13, 21, 39, 63	0
2	T	9/9 (100%)	-0.36	0 100 100	19, 21, 26, 29	0
3	P	4/4 (100%)	0.16	0 100 100	17, 19, 22, 24	0
4	D	4/4 (100%)	-0.47	0 100 100	21, 21, 23, 29	0
All	All	346/373 (92%)	0.17	21 (6%) 21 23	13, 21, 39, 63	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	137	TRP	8.0
1	A	270	GLN	4.0
1	A	267[A]	GLU	4.0
1	A	397	PRO	3.9
1	A	410	GLY	3.8
1	A	365	HIS	3.3
1	A	293	SER	3.2
1	A	366	GLN	3.1
1	A	352[A]	CYS	2.9
1	A	201	GLY	2.6
1	A	335	ILE	2.5
1	A	384	ALA	2.4
1	A	439	LEU	2.4
1	A	468	THR	2.3
1	A	431	LEU	2.3
1	A	389	PHE	2.3
1	A	420[A]	VAL	2.2
1	A	464	PRO	2.2
1	A	334	LEU	2.1
1	A	197	SER	2.0
1	A	395	PRO	2.0

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

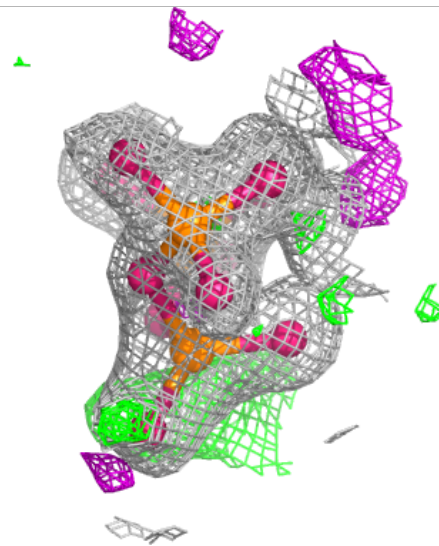
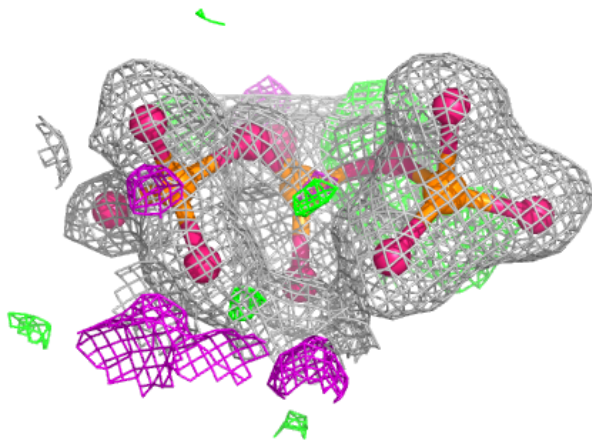
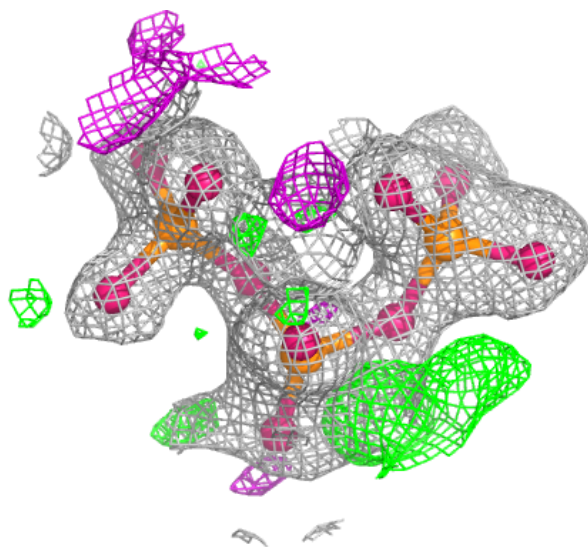
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
5	EDO	A	505	4/4	0.79	0.13	60,62,65,69	0
5	EDO	A	504	2/4	0.82	0.16	14,14,14,20	0
5	EDO	A	502	4/4	0.87	0.28	29,35,41,50	4
5	EDO	A	501	4/4	0.90	0.20	30,34,43,50	0
5	EDO	A	503	4/4	0.92	0.23	19,21,25,35	4
9	CL	A	511[B]	1/1	0.93	0.09	37,37,37,37	1
6	DGT	A	507	13/31	0.94	0.11	18,23,30,30	13
7	MG	A	509	1/1	0.94	0.06	18,18,18,18	1
5	EDO	A	506	4/4	0.94	0.06	29,31,35,40	4
10	DTT	A	512[A]	1/8	0.94	0.09	51,51,51,51	1
11	EPE	A	513	4/15	0.95	0.12	23,26,26,33	4
7	MG	A	508	1/1	0.98	0.08	19,19,19,19	0
8	NA	A	510	1/1	0.99	0.04	19,19,19,19	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.

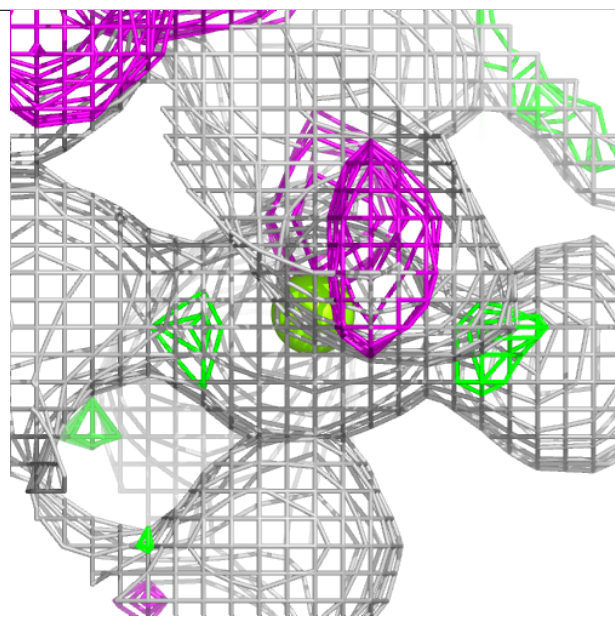
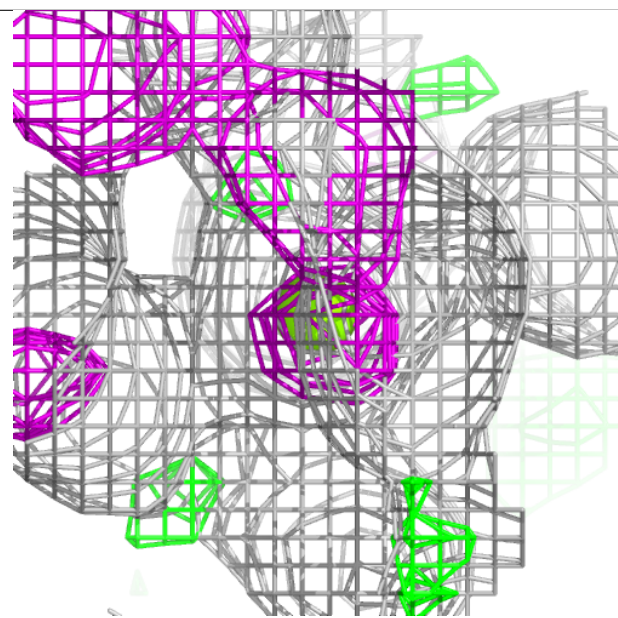
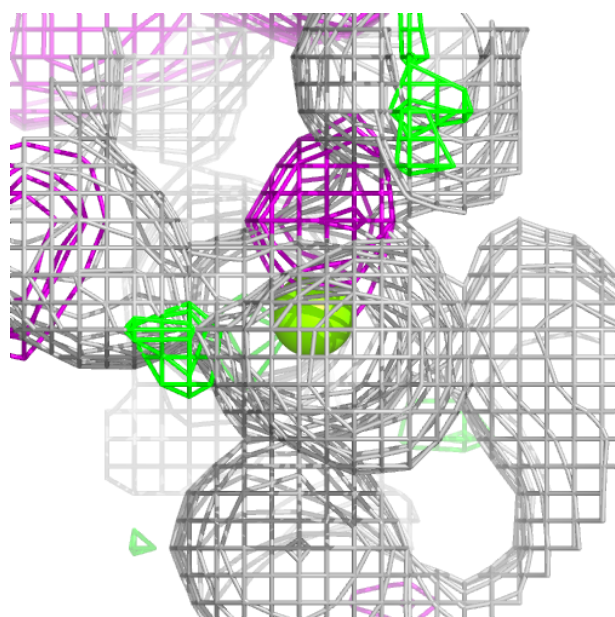
Electron density around DGT A 507:

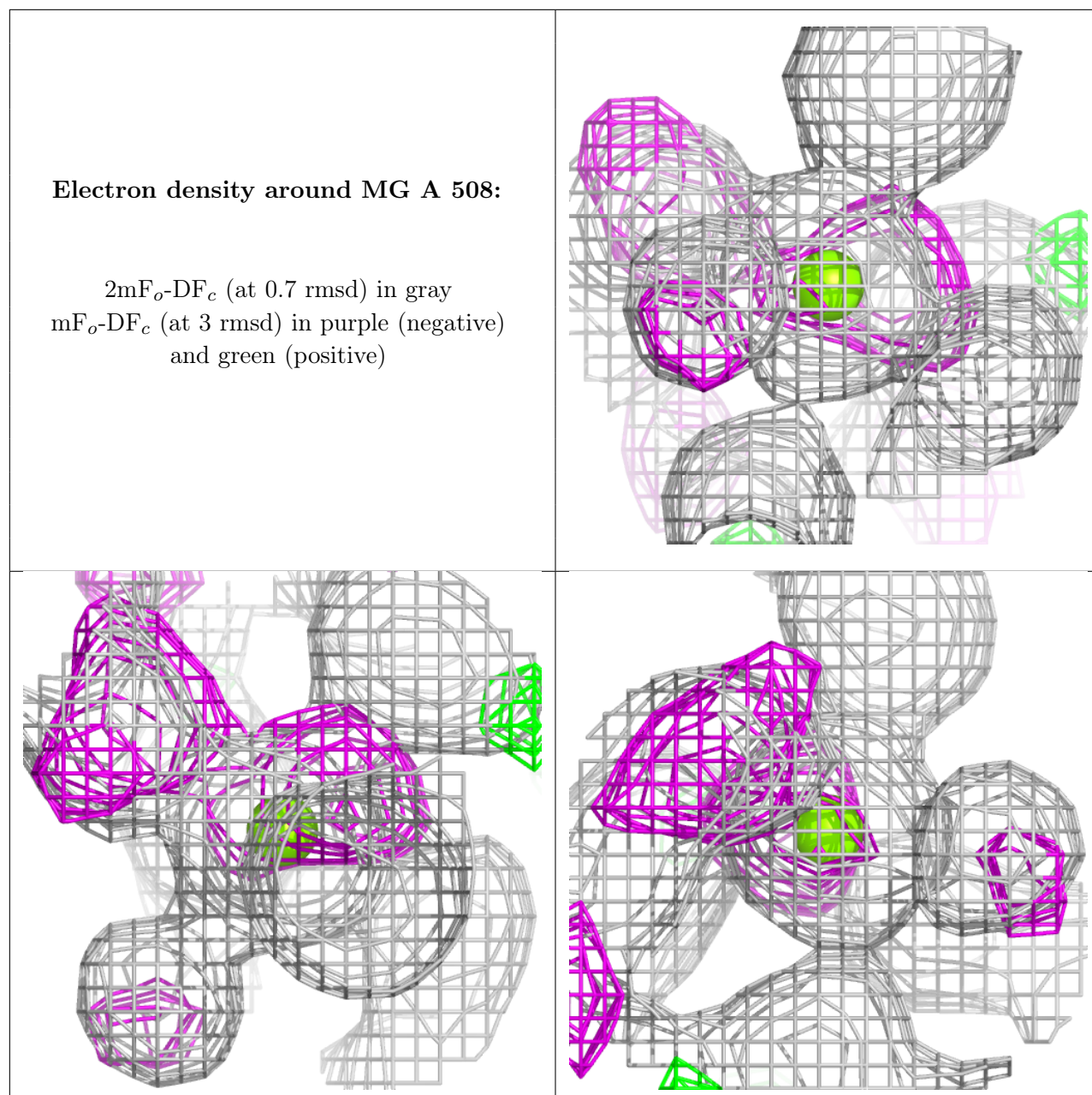
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around MG A 509:

$2mF_o-DF_c$ (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.