WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 5, 2023 – 07:50 AM EDT

PDB ID	:	6VFA
Title	:	DNA Polymerase Mu, 8-oxorGTP:Ct Ground State Ternary Complex, 50 mM Mn2+ (15 min)
Authors	:	Jamsen, J.A.; Wilson, S.H.
Deposited on	:	2020-01-03
Resolution	:	1.76 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

MolProbity	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
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 1.76 Å.

There are no overall percentile quality scores available for this entry.

MolProbity and EDS failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition ⓘ

There are 11 unique types of molecules in this entry. The entry contains 3301 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	327	Total	C	N	O	S	0	11	0
			2599	1660	461	469	9			

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	?	-	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*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*A)-3').

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

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

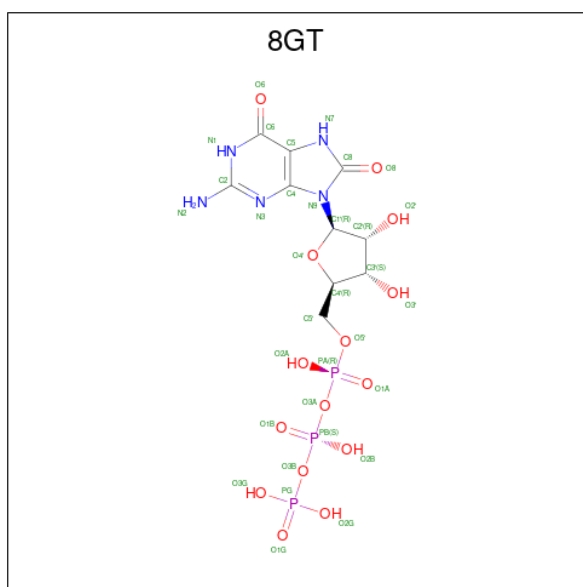| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	4	Total 83	C 38	N 16	O 25	P 4	0	0	0

- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	6	Total 6	Mn 6	0	0
5	T	1	Total 1	Mn 1	0	0
5	D	1	Total 1	Mn 1	0	0

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

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

- Molecule 7 is 8-OXO-GUANOSINE-5'-TRIPHOSPHATE (three-letter code: 8GT) (formula: $C_{10}H_{16}N_5O_{15}P_3$) (labeled as "Ligand of Interest" by depositor).

8GT

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	N	O	P	0	0
			33	10	5	15	3		

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

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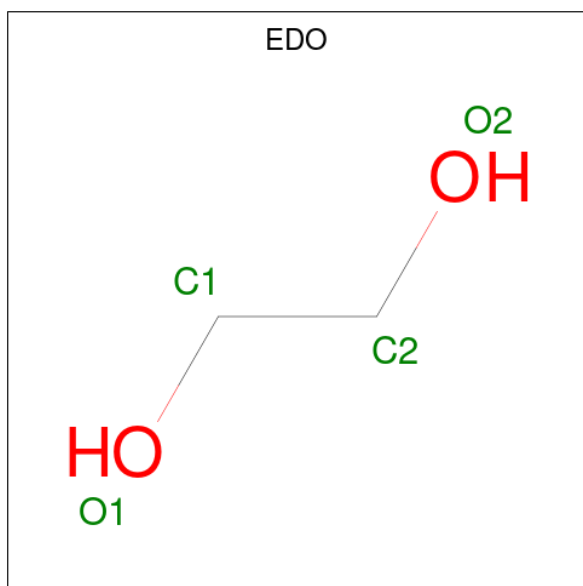
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			4	2	2		
8	A	1	Total	C	O	0	0
			4	2	2		

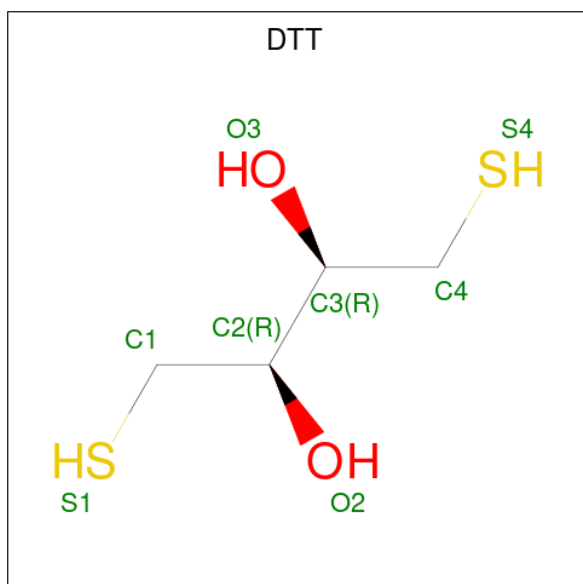- Molecule 9 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID

(three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).
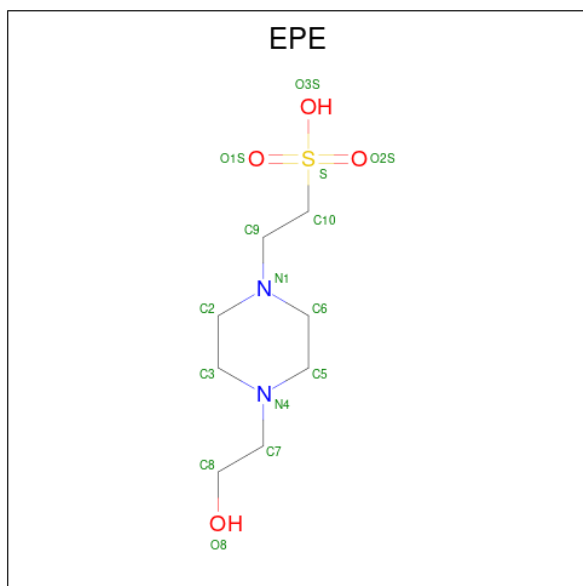
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
9	A	1	5	1	3	1	0	0

- Molecule 10 is 2,3-DIHYDROXY-1,4-DITHIOBUTANE (three-letter code: DTT) (formula: $C_4H_{10}O_2S_2$).

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

- Molecule 11 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	237	Total 242	O 242	0	8
11	T	27	Total 27	O 27	0	0
11	P	15	Total 15	O 15	0	0
11	D	20	Total 20	O 20	0	0

MolProbity and EDS failed to run properly - this section is therefore empty.

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	60.08Å 62.22Å 118.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.96 – 1.76	Depositor
% Data completeness (in resolution range)	99.0 (34.96-1.76)	Depositor
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.91 (at 1.76Å)	Xtriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, R_{free}	0.175 , 0.203	Depositor
Wilson B-factor (Å^2)	23.5	Xtriage
Anisotropy	0.114	Xtriage
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.015 for k,h,-l	Xtriage
Total number of atoms	3301	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.47% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

4 Model quality ⓘ

4.1 Standard geometry ⓘ

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts ⓘ

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles ⓘ

4.3.1 Protein backbone ⓘ

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains ⓘ

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA ⓘ

MolProbity failed to run properly - this section is therefore empty.

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

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

4.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

4.6 Ligand geometry ⓘ

Of 14 ligands modelled in this entry, 9 are monoatomic and 1 is modelled with single atom - leaving 4 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	EDO	A	510	-	3,3,3	0.46	0	2,2,2	0.47	0
7	8GT	A	508	5	31,35,35	1.11	4 (12%)	45,56,56	1.75	7 (15%)
8	EDO	A	509	-	3,3,3	0.45	0	2,2,2	0.33	0
9	EPE	A	511	-	4,4,15	0.98	0	5,6,20	1.57	1 (20%)

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	EDO	A	510	-	-	1/1/1/1	-
7	8GT	A	508	5	-	7/22/38/38	0/3/3/3
8	EDO	A	509	-	-	0/1/1/1	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	508	8GT	C8-N7	-3.75	1.31	1.38
7	A	508	8GT	C4-N9	-2.36	1.34	1.39
7	A	508	8GT	C8-N9	-2.35	1.36	1.40
7	A	508	8GT	O8-C8	-2.32	1.18	1.23

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	A	508	8GT	C2'-C1'-N9	-7.83	105.53	115.94
7	A	508	8GT	N7-C8-N9	4.25	111.62	106.58
9	A	511	EPE	O2S-S-O1S	-3.38	108.90	118.02
7	A	508	8GT	C5-N7-C8	-3.31	104.71	109.47
7	A	508	8GT	C2'-C3'-C4'	2.89	108.25	102.64
7	A	508	8GT	O4'-C1'-C2'	2.44	111.96	106.64
7	A	508	8GT	C4-C5-N7	2.14	110.21	106.08
7	A	508	8GT	PA-O3A-PB	-2.03	125.86	132.83

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	508	8GT	C5'-O5'-PA-O3A
7	A	508	8GT	C5'-O5'-PA-O1A
7	A	508	8GT	O4'-C4'-C5'-O5'
7	A	508	8GT	C3'-C4'-C5'-O5'
8	A	510	EDO	O1-C1-C2-O2
7	A	508	8GT	PB-O3A-PA-O2A
7	A	508	8GT	C5'-O5'-PA-O2A
7	A	508	8GT	PB-O3A-PA-O1A

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 then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand 8GT A 508

Bond lengths

Bond angles

Torsions

Rings
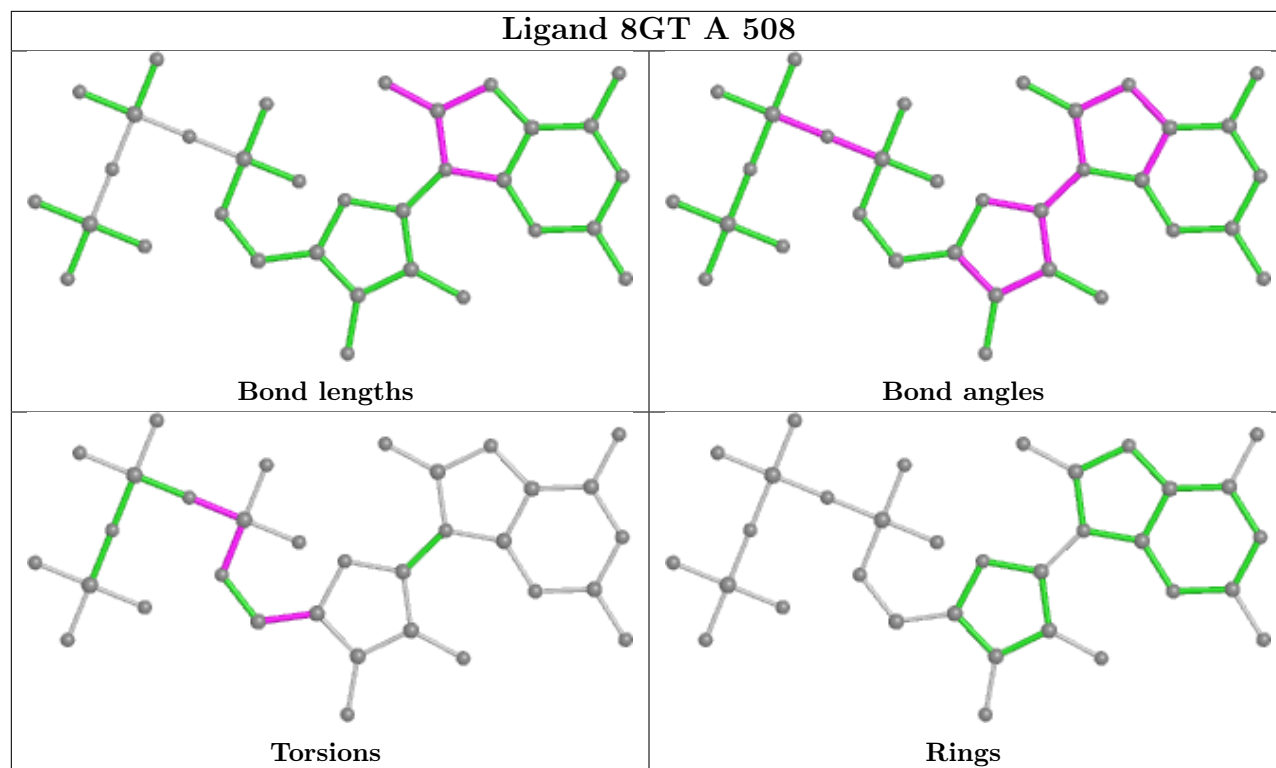
4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data ⓘ

5.1 Protein, DNA and RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates ⓘ

EDS failed to run properly - this section is therefore empty.

5.4 Ligands ⓘ

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers ⓘ

EDS failed to run properly - this section is therefore empty.