



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 12:49 am BST

PDB ID : 2NVT  
Title : RNA Polymerase II Elongation Complex in 150 mM Mg<sup>2+</sup> with GMPCPP  
Authors : Wang, D.; Bushnell, D.A.; Westover, K.D.; Kaplan, C.D.; Kornberg, R.D.  
Deposited on : 2006-11-13  
Resolution : 3.36 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 : 2.11  
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

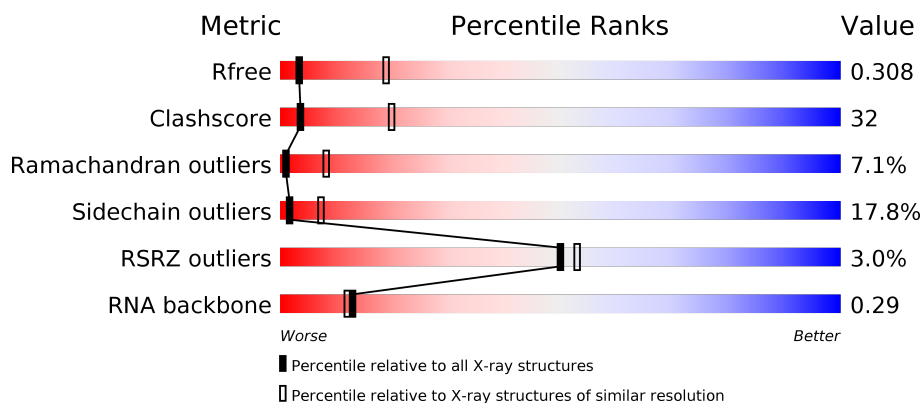
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.36 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	1558 (3.42-3.30)
Clashscore	141614	1627 (3.42-3.30)
Ramachandran outliers	138981	1599 (3.42-3.30)
Sidechain outliers	138945	1598 (3.42-3.30)
RSRZ outliers	127900	1507 (3.42-3.30)
RNA backbone	3102	1023 (3.80-2.92)

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	R	10	
2	T	21	
3	N	7	
4	A	1733	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>2%37%39%13%9%</div></div>
6	C	318	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>%39%39%6%16%</div></div>
7	E	215	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>5%56%34%9%</div></div>
8	F	155	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>36%16%45%</div></div>
9	H	146	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>3%43%38%8%9%</div></div>
10	I	122	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>53%35%8%</div></div>
11	J	70	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>43%33%14%7%</div></div>
12	K	120	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>54%37%5%</div></div>
13	L	70	<div><div><div></div><div></div><div></div><div></div><div></div></div><div>7%23%29%13%34%</div></div>

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 29168 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 RNA chain called 5'-R(\*AP\*AP\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	10	Total	C	N	O	P	0	0	0
			214	97	44	64	9			

- Molecule 2 is a DNA chain called 5'-D(P\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*AP\*CP\*GP\*CP\*CP\*TP\*GP\*GP\*TP\*CP\*TP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	21	Total	C	N	O	P	0	0	0
			426	204	72	129	21			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	7	Total	C	N	O	P	0	0	0
			141	69	24	42	6			

- Molecule 4 is a protein called DNA-directed RNA polymerase II largest subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1411	Total	C	N	O	S	0	0	0
			11090	6993	1942	2094	61			

- Molecule 5 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1114	Total	C	N	O	S	0	0	0
			8861	5610	1549	1647	55			

- Molecule 6 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	267	Total	C	N	O	S	0	0	0
			2101	1320	349	419	13			

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	85	Total	C	N	O	S	0	0	0
			688	439	116	130	3			

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	I	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			

- Molecule 11 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 12 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

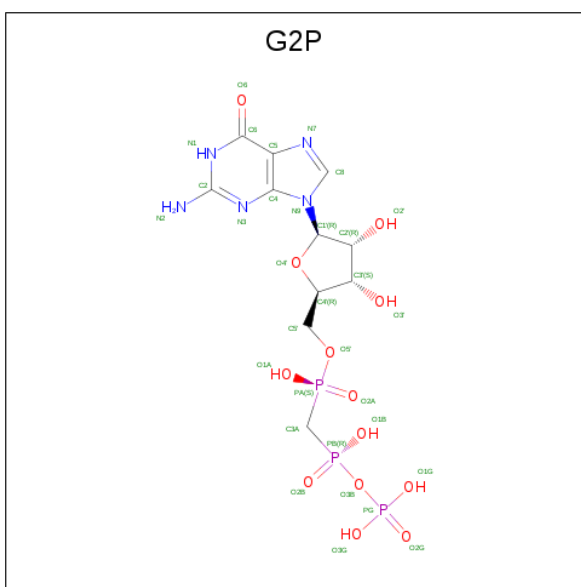
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypep-

tide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

- Molecule 14 is PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER (three-letter code: G2P) (formula:  $C_{11}H_{18}N_5O_{13}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
14	T	1	Total	C	N	O	P	0	0
			32	11	5	13	3		

- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	J	1	Total Zn 1 1	0	0
15	B	1	Total Zn 1 1	0	0
15	I	2	Total Zn 2 2	0	0
15	C	1	Total Zn 1 1	0	0
15	A	2	Total Zn 2 2	0	0
15	L	1	Total Zn 1 1	0	0

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

### 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: 5'-R(\*AP\*AP\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*C)-3'

Chain R: 



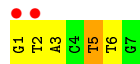
- Molecule 2: 5'-D(P\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*AP\*CP\*GP\*CP\*CP\*TP\*GP\*GP\*TP\*CP\*TP\*T)-3'

Chain T: 



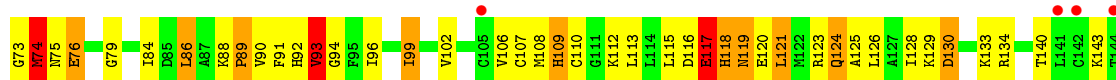
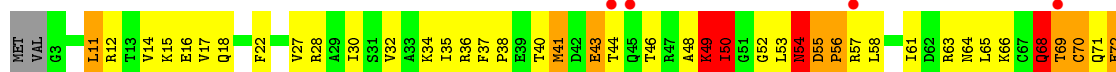
- Molecule 3: 5'-D(\*GP\*TP\*AP\*CP\*TP\*TP\*G)-3'

Chain N: 



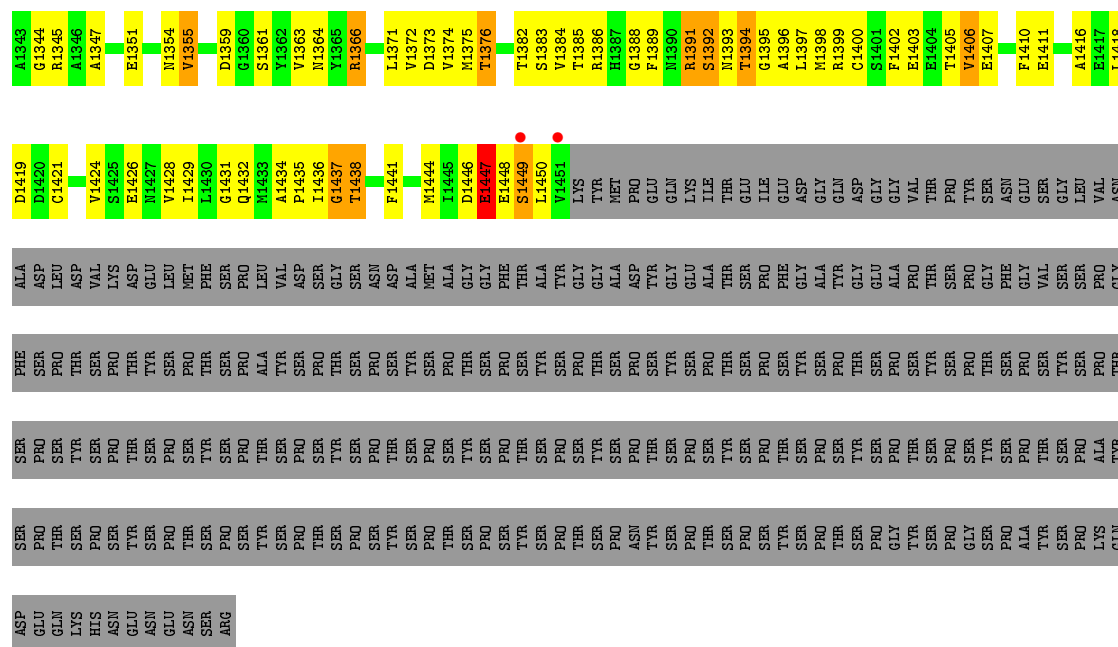
- Molecule 4: DNA-directed RNA polymerase II largest subunit

Chain A: 

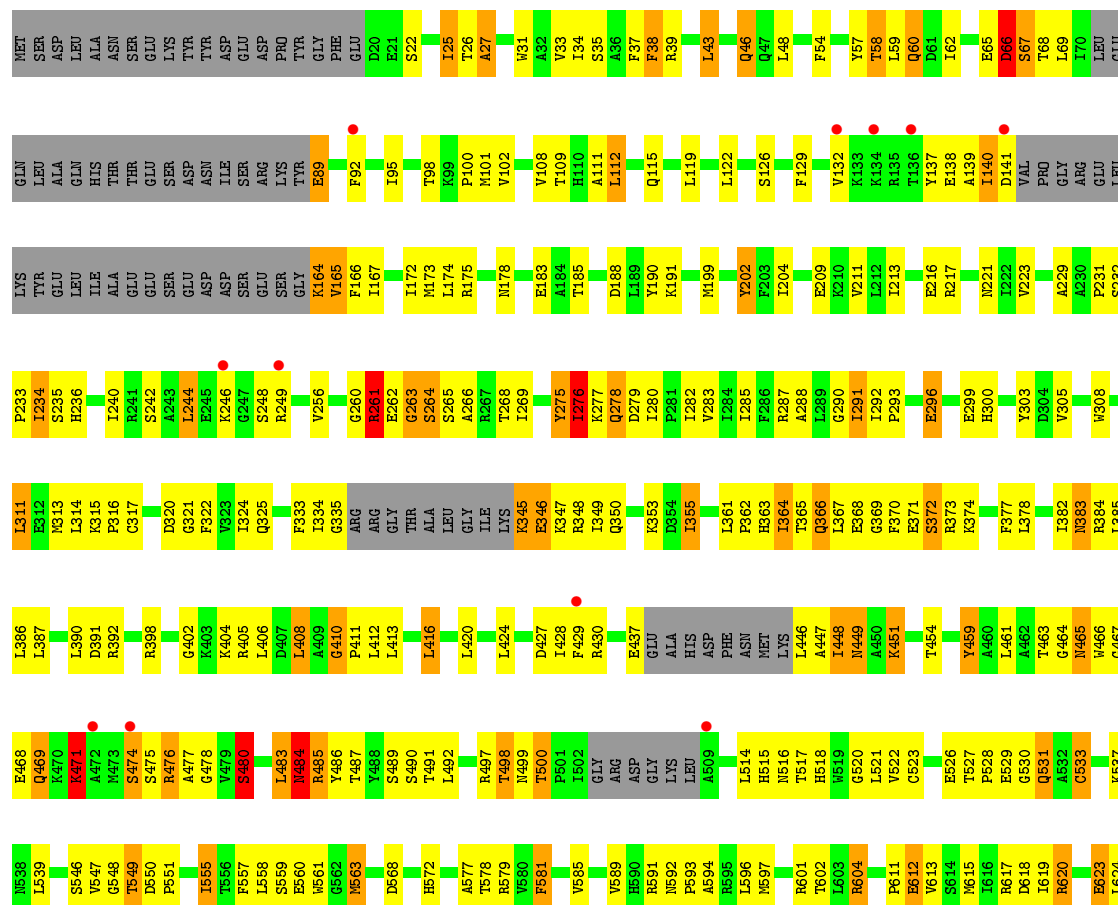






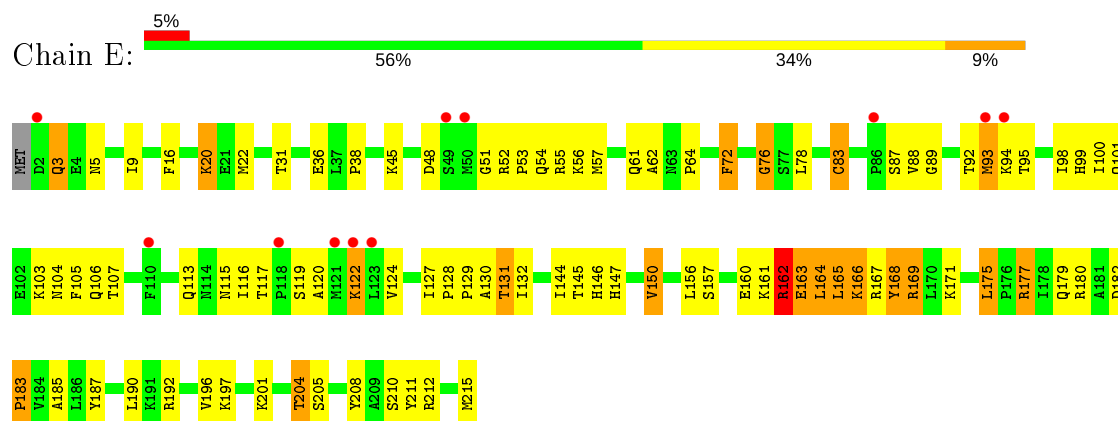


• Molecule 5: DNA-directed RNA polymerase II 140 kDa polypeptide

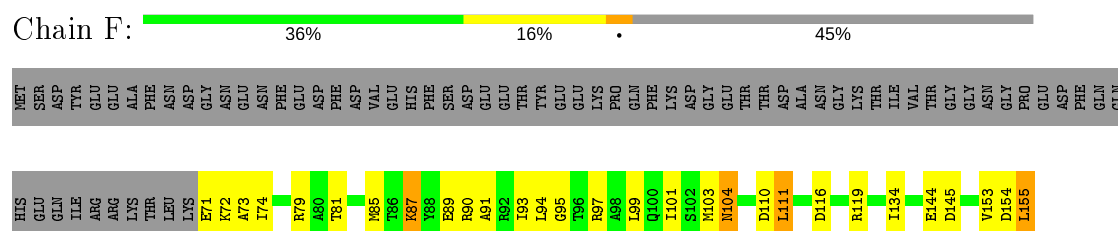




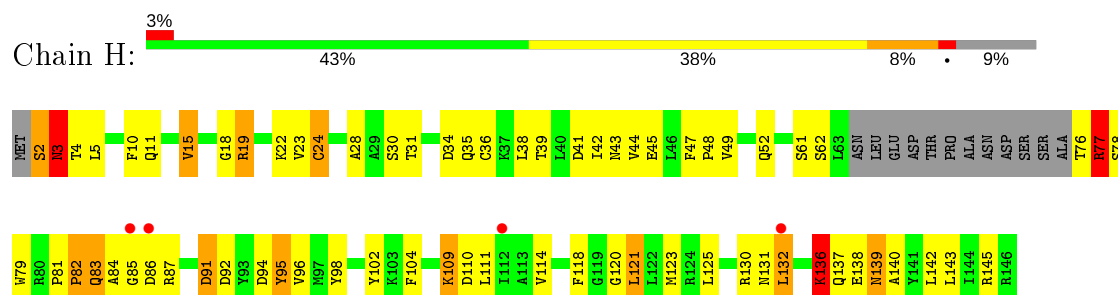
- Molecule 7: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide



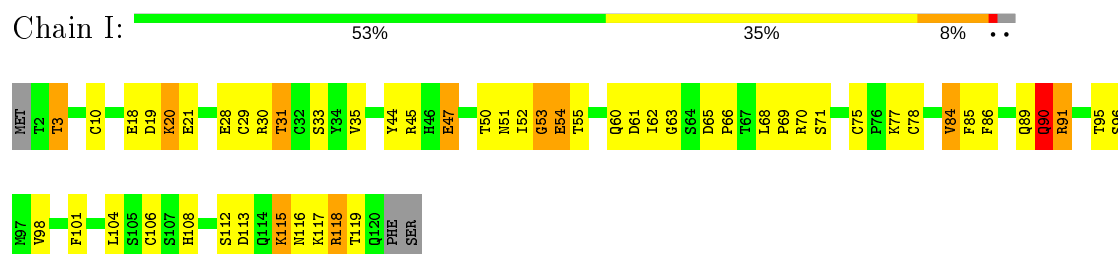
- Molecule 8: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide



- Molecule 9: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide



- Molecule 10: DNA-directed RNA polymerase II subunit 9

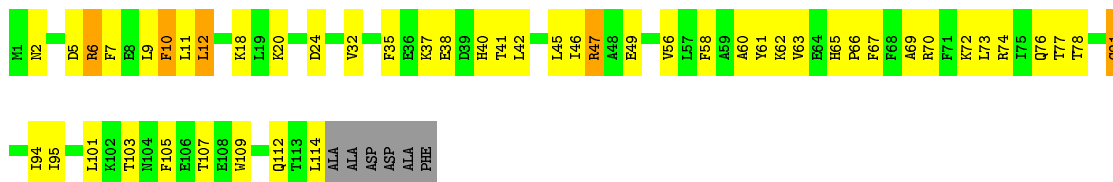


- Molecule 11: DNA-directed RNA polymerases I/II/III subunit 10

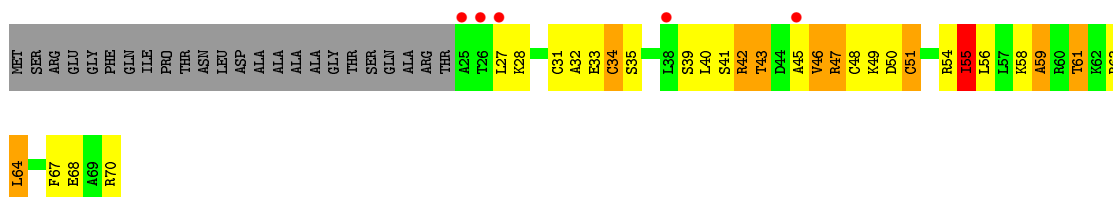
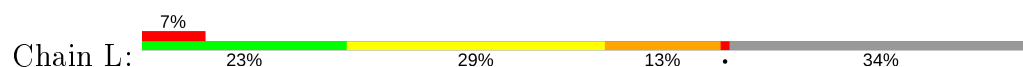




- Molecule 12: DNA-directed RNA polymerase II 13.6 kDa polypeptide



- Molecule 13: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	170.87Å 222.82Å 195.80Å 90.00° 102.39° 90.00°	Depositor
Resolution (Å)	40.00 – 3.36 39.86 – 3.36	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.36) 97.1 (39.86-3.36)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.12	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.62 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.230 , 0.283 0.257 , 0.308	Depositor DCC
$R_{free}$ test set	4911 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	97.0	Xtriage
Anisotropy	0.181	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 45.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	29168	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	123.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% 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: ZN, MG, G2P

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	R	1.40	3/240 (1.2%)	2.56	27/373 (7.2%)
2	T	1.19	0/475	2.56	38/730 (5.2%)
3	N	0.91	0/157	1.68	6/241 (2.5%)
4	A	0.68	2/11288 (0.0%)	0.82	6/15263 (0.0%)
5	B	0.82	7/9033 (0.1%)	0.91	16/12181 (0.1%)
6	C	0.76	1/2139 (0.0%)	0.89	1/2899 (0.0%)
7	E	0.51	0/1788	0.69	1/2406 (0.0%)
8	F	0.54	0/700	0.75	0/945
9	H	0.56	0/1086	0.77	0/1470
10	I	0.59	0/989	0.73	1/1331 (0.1%)
11	J	0.90	0/541	0.91	0/727
12	K	0.65	0/937	0.76	0/1265
13	L	0.66	0/365	0.93	0/485
All	All	0.74	13/29738 (0.0%)	0.94	96/40316 (0.2%)

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
4	A	0	5
5	B	0	6
9	H	0	1
All	All	0	12

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1222	ARG	CZ-NH1	7.73	1.43	1.33
1	R	10	C	C2-N3	7.26	1.41	1.35

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	R	10	C	C2'-C1'	6.83	1.60	1.53
1	R	10	C	N3-C4	6.79	1.38	1.33
4	A	1020	CYS	CB-SG	-6.59	1.71	1.82

The worst 5 of 96 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	22	DT	C4-C5-C7	-16.64	109.02	119.00
2	T	27	DT	C6-C5-C7	-12.28	115.53	122.90
2	T	13	DA	O4'-C1'-N9	11.91	116.34	108.00
2	T	20	DC	O4'-C4'-C3'	-11.64	99.01	106.00
2	T	22	DT	C5-C4-O4	-11.59	116.79	124.90

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	115	LEU	Peptide
4	A	117	GLU	Peptide
4	A	342	GLY	Peptide
4	A	482	PHE	Peptide
4	A	484	GLY	Peptide

## 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	R	214	0	111	18	0
2	T	426	0	239	23	0
3	N	141	0	82	2	0
4	A	11090	0	11173	784	1
5	B	8861	0	8884	722	0
6	C	2101	0	2056	149	1
7	E	1752	0	1776	74	0
8	F	688	0	707	26	0
9	H	1068	0	1040	54	0
10	I	971	0	927	36	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	J	532	0	542	71	0
12	K	919	0	929	65	0
13	L	363	0	386	27	0
14	T	32	0	14	0	0
15	A	2	0	0	0	0
15	B	1	0	0	0	0
15	C	1	0	0	0	0
15	I	2	0	0	0	0
15	J	1	0	0	0	0
15	L	1	0	0	0	0
16	A	2	0	0	0	0
All	All	29168	0	28866	1832	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 32.

The worst 5 of 1832 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)
4:A:336:ILE:CD1	4:A:1405:THR:CG2	1.78	1.57
5:B:976:ILE:HD11	5:B:991:GLY:C	1.14	1.48
4:A:341:MET:SD	4:A:1428:VAL:HG12	1.62	1.39
4:A:336:ILE:CD1	4:A:1405:THR:HG21	0.93	1.39
5:B:757:PRO:CG	5:B:757:PRO:CB	1.75	1.38

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)
4:A:418:SER:OG	6:C:87:PHE:O[2_555]	1.88	0.32

## 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	
4	A	1401/1733 (81%)	1073 (77%)	223 (16%)	105 (8%)	1	7
5	B	1096/1224 (90%)	823 (75%)	190 (17%)	83 (8%)	1	7
6	C	265/318 (83%)	211 (80%)	38 (14%)	16 (6%)	1	11
7	E	212/215 (99%)	178 (84%)	25 (12%)	9 (4%)	3	19
8	F	83/155 (54%)	65 (78%)	14 (17%)	4 (5%)	2	16
9	H	129/146 (88%)	90 (70%)	25 (19%)	14 (11%)	0	3
10	I	117/122 (96%)	86 (74%)	24 (20%)	7 (6%)	1	11
11	J	63/70 (90%)	53 (84%)	6 (10%)	4 (6%)	1	10
12	K	112/120 (93%)	94 (84%)	16 (14%)	2 (2%)	8	35
13	L	44/70 (63%)	26 (59%)	12 (27%)	6 (14%)	0	1
All	All	3522/4173 (84%)	2699 (77%)	573 (16%)	250 (7%)	1	8

5 of 250 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	50	ILE
4	A	56	PRO
4	A	68	GLN
4	A	69	THR
4	A	72	GLU

### 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	
4	A	1231/1520 (81%)	1008 (82%)	223 (18%)	1	7
5	B	967/1061 (91%)	774 (80%)	193 (20%)	1	4
6	C	235/274 (86%)	205 (87%)	30 (13%)	4	18
7	E	196/197 (100%)	164 (84%)	32 (16%)	2	10
8	F	75/137 (55%)	70 (93%)	5 (7%)	16	47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
9	H	117/128 (91%)	93 (80%)	24 (20%)	1	4
10	I	113/116 (97%)	97 (86%)	16 (14%)	3	14
11	J	60/65 (92%)	46 (77%)	14 (23%)	1	2
12	K	99/102 (97%)	88 (89%)	11 (11%)	6	24
13	L	40/57 (70%)	30 (75%)	10 (25%)	0	2
All	All	3133/3657 (86%)	2575 (82%)	558 (18%)	2	7

5 of 558 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	296	GLU
5	B	668	ASP
10	I	75	CYS
5	B	355	ILE
5	B	489	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 87 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	215	GLN
5	B	657	HIS
9	H	139	ASN
5	B	255	GLN
5	B	484	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	10/10 (100%)	3 (30%)	1 (10%)

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	2	A
1	R	3	G
1	R	8	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	R	1	A

## 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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 10 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	$\# Z  > 2$
14	G2P	T	3000	-	26,34,34	2.60	8 (30%)	30,54,54	1.56	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
14	G2P	T	3000	-	-	3/18/38/38	0/3/3/3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	T	3000	G2P	C5-C6	-8.02	1.39	1.52
14	T	3000	G2P	C4-N9	-6.71	1.38	1.47
14	T	3000	G2P	PB-O3B	4.10	1.62	1.58
14	T	3000	G2P	PA-O5'	3.03	1.61	1.57

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	T	3000	G2P	C8-N9	-2.43	1.37	1.45

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
14	T	3000	G2P	C4-C5-N7	4.27	108.12	102.46
14	T	3000	G2P	PG-O3B-PB	-3.48	120.37	132.62
14	T	3000	G2P	O4'-C4'-C5'	3.12	119.65	109.37
14	T	3000	G2P	O5'-C5'-C4'	2.42	117.32	108.99
14	T	3000	G2P	O6-C6-N1	-2.41	119.45	122.69

There are no chirality outliers.

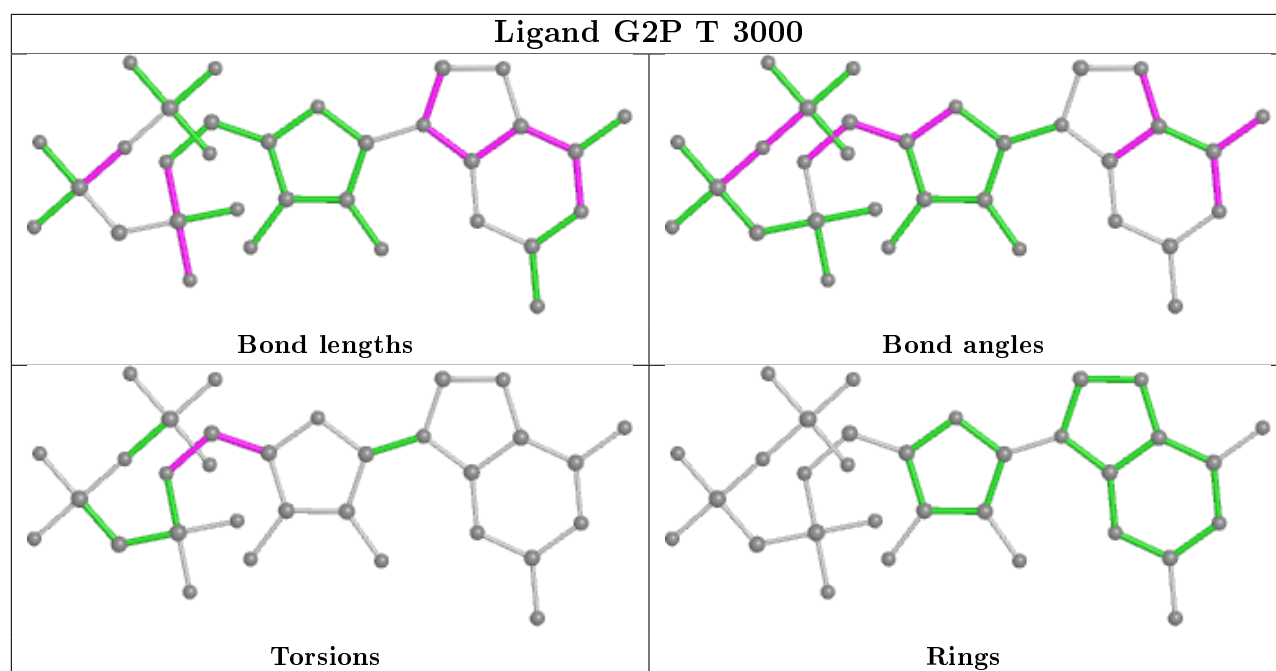
All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
14	T	3000	G2P	O4'-C4'-C5'-O5'
14	T	3000	G2P	C3'-C4'-C5'-O5'
14	T	3000	G2P	C4'-C5'-O5'-PA

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 [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	R	10/10 (100%)	-0.60	0 <span>100</span> <span>100</span>	76, 118, 173, 179	0
2	T	21/21 (100%)	0.38	3 (14%) <span>2</span> <span>3</span>	100, 128, 234, 236	0
3	N	7/7 (100%)	1.52	2 (28%) <span>0</span> <span>0</span>	225, 230, 244, 248	0
4	A	1411/1733 (81%)	0.06	49 (3%) <span>44</span> <span>46</span>	83, 119, 142, 195	0
5	B	1114/1224 (91%)	0.18	30 (2%) <span>54</span> <span>57</span>	76, 121, 146, 177	0
6	C	267/318 (83%)	-0.08	3 (1%) <span>80</span> <span>84</span>	104, 119, 141, 159	0
7	E	214/215 (99%)	0.15	11 (5%) <span>28</span> <span>30</span>	100, 124, 146, 151	0
8	F	85/155 (54%)	-0.16	0 <span>100</span> <span>100</span>	107, 126, 146, 152	0
9	H	133/146 (91%)	0.14	4 (3%) <span>50</span> <span>53</span>	115, 134, 153, 156	0
10	I	119/122 (97%)	-0.01	0 <span>100</span> <span>100</span>	101, 122, 146, 158	0
11	J	65/70 (92%)	-0.03	0 <span>100</span> <span>100</span>	94, 118, 138, 145	0
12	K	114/120 (95%)	-0.09	0 <span>100</span> <span>100</span>	114, 123, 134, 140	0
13	L	46/70 (65%)	0.52	5 (10%) <span>5</span> <span>6</span>	131, 170, 177, 179	0
All	All	3606/4211 (85%)	0.09	107 (2%) <span>50</span> <span>53</span>	76, 121, 147, 248	0

The worst 5 of 107 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	N	1	DG	6.0
5	B	474	SER	5.5
4	A	1085	HIS	5.4
4	A	1089	VAL	5.2
4	A	175	ARG	5.1

### 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 ⓘ

There are no carbohydrates in this entry.

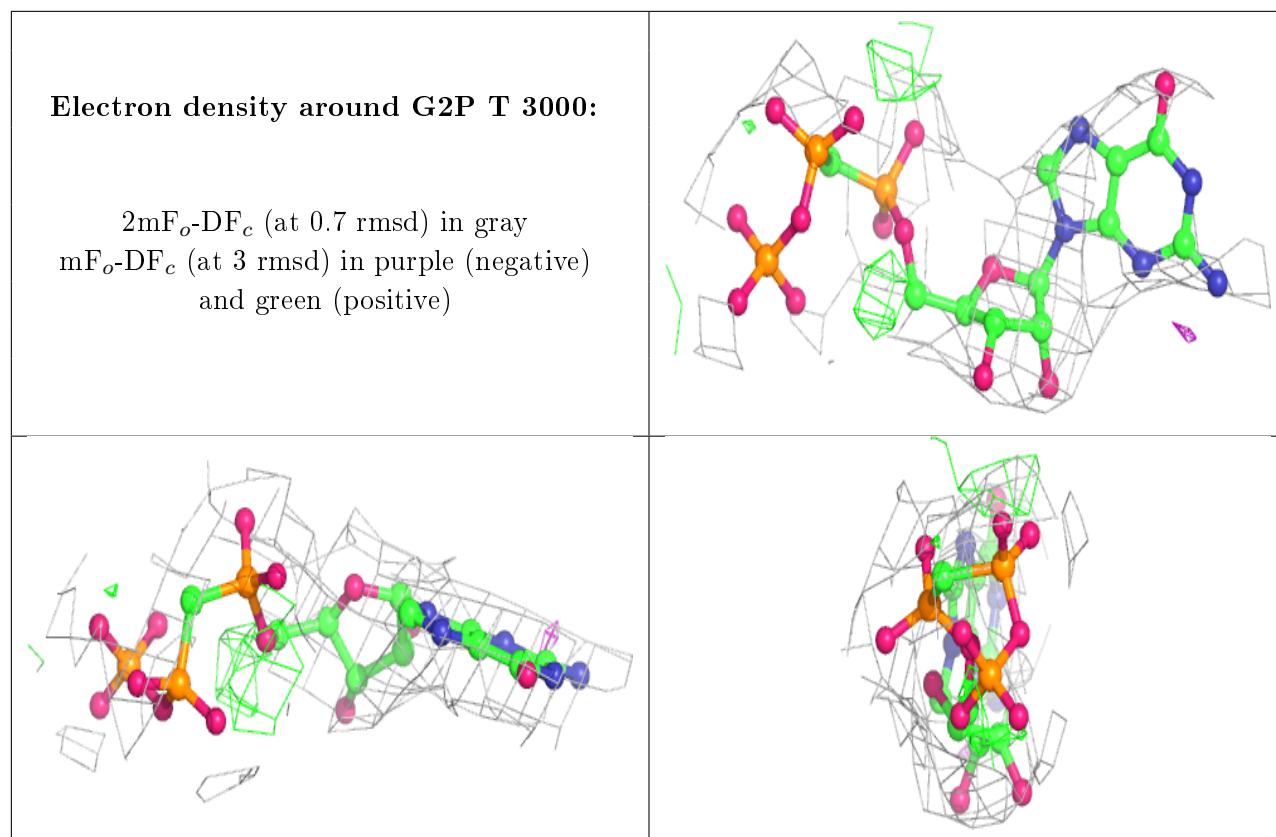
### 6.4 Ligands ⓘ

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( $\text{\AA}^2$ )	Q<0.9
14	G2P	T	3000	32/32	0.85	0.27	120,126,149,149	0
15	ZN	A	1734	1/1	0.86	0.07	141,141,141,141	0
16	MG	A	2002	1/1	0.93	0.61	79,79,79,79	0
15	ZN	A	1735	1/1	0.94	0.09	141,141,141,141	0
15	ZN	B	1307	1/1	0.94	0.07	134,134,134,134	0
15	ZN	L	105	1/1	0.94	0.04	170,170,170,170	0
15	ZN	I	204	1/1	0.96	0.04	128,128,128,128	0
15	ZN	I	203	1/1	0.97	0.11	114,114,114,114	0
16	MG	A	2003	1/1	0.98	0.15	50,50,50,50	0
15	ZN	J	101	1/1	0.99	0.18	107,107,107,107	0
15	ZN	C	319	1/1	0.99	0.07	120,120,120,120	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.