



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 4, 2025 – 10:13 AM EDT

PDB ID : 9N5D / pdb\_00009n5d  
Title : RNA polymerase II elongation complex with 8-oxoG at +1 site, CMP added  
Authors : Oh, J.; Wang, D.  
Deposited on : 2025-02-04  
Resolution : 3.35 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 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 : 4-5-2 with Phenix2.0rc1  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.006 (Gargrove)  
Density-Fitness : 1.0.12  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.45.1

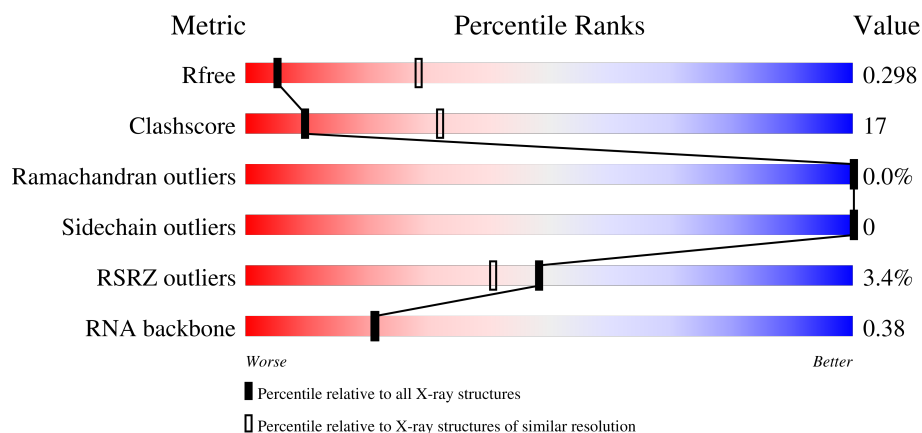
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.35 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))
$R_{free}$	164625	1012 (3.40-3.32)
Clashscore	180529	1035 (3.40-3.32)
Ramachandran outliers	177936	1037 (3.40-3.32)
Sidechain outliers	177891	1037 (3.40-3.32)
RSRZ outliers	164620	1012 (3.40-3.32)
RNA backbone	3690	1010 (3.74-2.98)

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	R	10	<div> <div>10%</div> <div> <div></div> <div>90%</div> <div>10%</div> </div> </div>
2	T	29	<div> <div>31%</div> <div>59%</div> <div>10%</div> </div>
3	N	18	<div> <div>11%</div> <div>72%</div> <div>17%</div> </div>
4	A	1733	<div> <div>4%</div> <div>49%</div> <div>31%</div> <div>20%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	B	1224	<p>2% 60% 32% 8%</p>
6	C	318	<p>% 50% 34% 16%</p>
7	E	215	<p>7% 56% 43% .</p>
8	F	155	<p>% 41% 15% 45%</p>
9	H	146	<p>5% 61% 30% 9%</p>
10	I	122	<p>% 66% 31% .</p>
11	J	70	<p>3% 47% 46% 7%</p>
12	K	120	<p>63% 32% 5%</p>
13	L	70	<p>4% 37% 23% . 39%</p>

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 29066 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 RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	10	Total	C	N	O	P	0	0	0
			219	97	43	69	10			

- Molecule 2 is a DNA chain called Template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	26	Total	C	N	O	P	0	0	0
			521	250	80	165	26			

- Molecule 3 is a DNA chain called Non-template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	15	Total	C	N	O	P	0	0	0
			317	148	71	83	15			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1385	Total	C	N	O	S	0	0	0
			10831	6833	1895	2043	60			

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1121	Total	C	N	O	S	0	0	0
			8849	5601	1550	1645	53			

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPB3.

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 subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	213	Total	C	N	O	S	0	0	0
			1728	1099	301	317	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	86	Total	C	N	O	S	0	0	0
			684	437	115	129	3			

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	133	Total	C	N	O	S	0	0	0
			1064	670	179	211	4			

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	I	118	Total	C	N	O	S	0	0	0
			946	582	170	184	10			

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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 subunit RPB11.

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 subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	43	Total	C	N	O	S	0	0	0
			337	208	66	59	4			

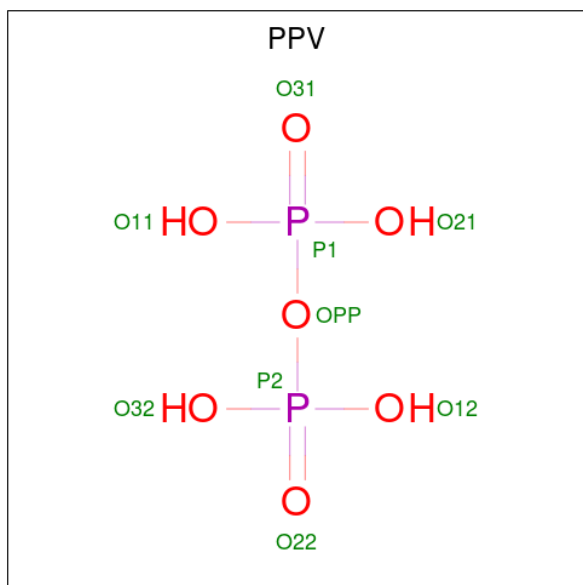
- Molecule 14 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

- Molecule 15 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	1	Total	Mg	0	0
			1	1		

- Molecule 16 is PYROPHOSPHATE (CCD ID: PPV) (formula:  $\text{H}_4\text{O}_7\text{P}_2$ ) (labeled as "Ligand of Interest" by depositor).



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

### 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.

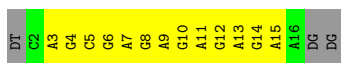
- Molecule 1: RNA



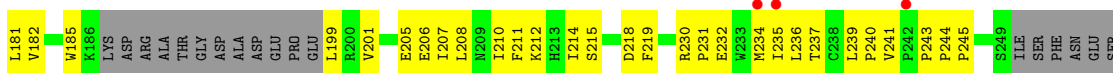
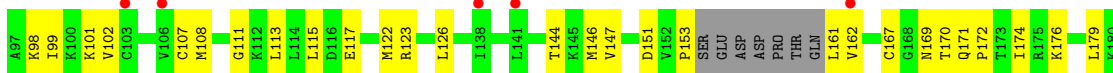
- Molecule 2: Template strand DNA



- Molecule 3: Non-template strand DNA



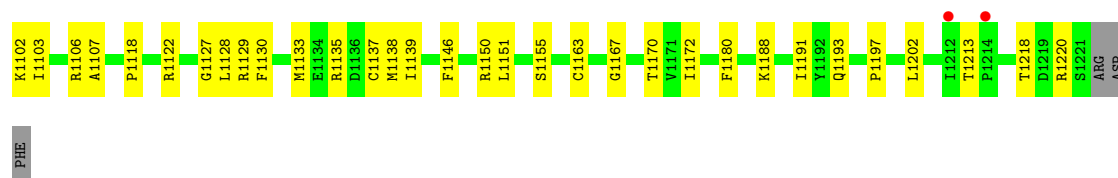
- Molecule 4: DNA-directed RNA polymerase II subunit RPB1



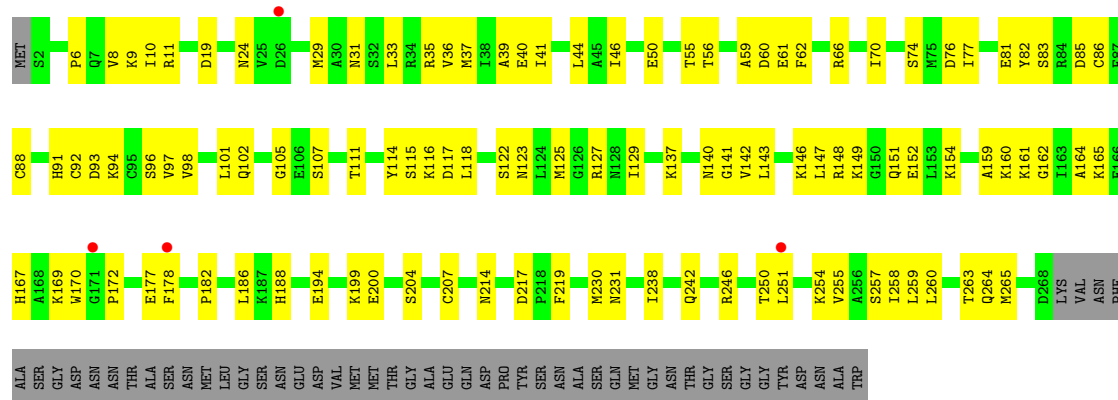




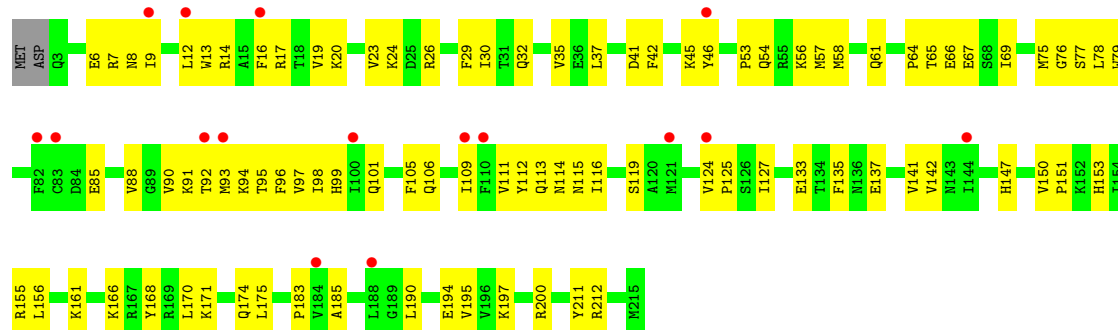




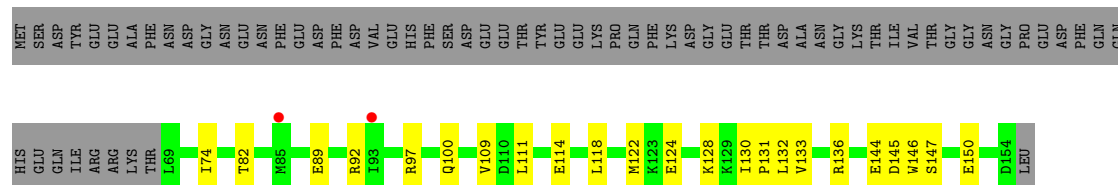
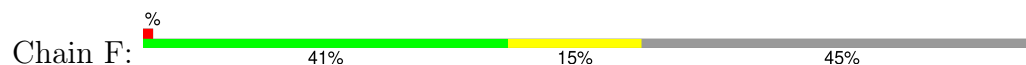
• Molecule 6: DNA-directed RNA polymerase II subunit RPB3



• Molecule 7: DNA-directed RNA polymerases I, II, and III subunit RPABC1

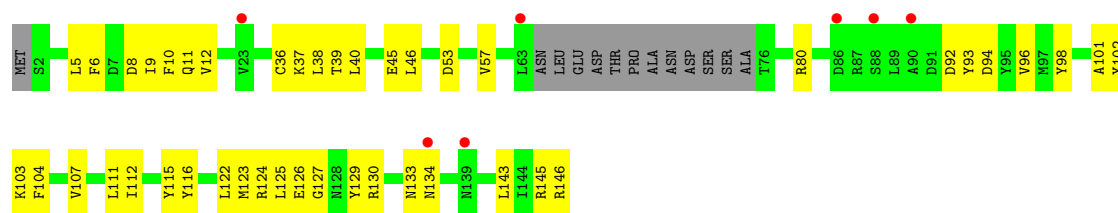


• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC2

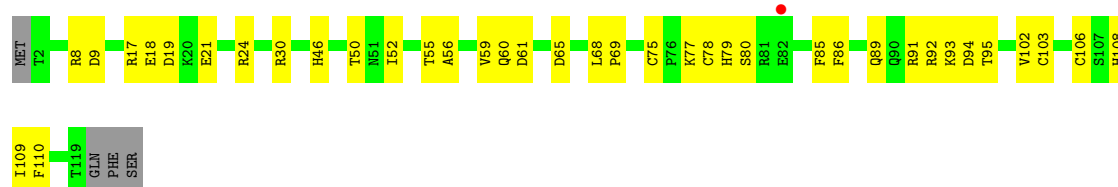


• Molecule 9: DNA-directed RNA polymerases I, II, and III subunit RPABC3





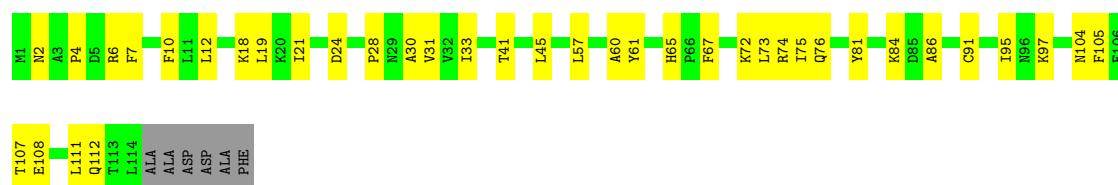
• Molecule 10: DNA-directed RNA polymerase II subunit RPB9



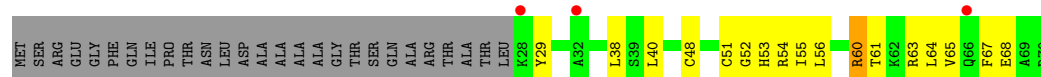
• Molecule 11: DNA-directed RNA polymerases I, II, and III subunit RPABC5



• Molecule 12: DNA-directed RNA polymerase II subunit RPB11



• Molecule 13: DNA-directed RNA polymerases I, II, and III subunit RPABC4



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	165.90Å 222.27Å 193.15Å 90.00° 100.27° 90.00°	Depositor
Resolution (Å)	49.31 – 3.35 49.31 – 3.35	Depositor EDS
% Data completeness (in resolution range)	98.9 (49.31-3.35) 98.8 (49.31-3.35)	Depositor EDS
$R_{merge}$	0.51	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.13 (at 3.33Å)	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
R, $R_{free}$	0.265 , 0.300 0.264 , 0.298	Depositor DCC
$R_{free}$ test set	2000 reflections (1.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	91.1	Xtriage
Anisotropy	0.445	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 75.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.23$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	29066	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.03% 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: PPV, 8OG, MG, ZN

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	0.14	0/245	0.32	0/379
2	T	0.23	0/551	0.49	0/843
3	N	0.21	0/359	0.39	0/553
4	A	0.15	0/11023	0.39	0/14912
5	B	0.14	0/9020	0.35	0/12172
6	C	0.14	0/2139	0.34	0/2899
7	E	0.15	0/1764	0.39	0/2376
8	F	0.13	0/696	0.36	0/943
9	H	0.14	0/1082	0.42	0/1466
10	I	0.16	0/964	0.40	0/1301
11	J	0.17	0/541	0.40	0/727
12	K	0.14	0/937	0.36	0/1265
13	L	0.16	0/339	0.41	0/450
All	All	0.15	0/29660	0.38	0/40286

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	1
5	B	0	2
13	L	0	1
All	All	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	311	GLN	Peptide
5	B	267	ARG	Sidechain
5	B	635	ARG	Sidechain
13	L	60	ARG	Sidechain

## 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	219	0	110	1	0
2	T	521	0	297	19	0
3	N	317	0	166	14	0
4	A	10831	0	10873	419	1
5	B	8849	0	8808	311	0
6	C	2101	0	2058	85	0
7	E	1728	0	1744	79	0
8	F	684	0	692	18	0
9	H	1064	0	1029	40	0
10	I	946	0	888	36	1
11	J	532	0	544	38	0
12	K	919	0	929	35	0
13	L	337	0	354	15	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	A	1	0	0	0	0
16	A	9	0	0	0	0
All	All	29066	0	28492	991	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 17.

The worst 5 of 991 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)
7:E:96:PHE:HA	7:E:99:HIS:HD2	1.26	1.00
6:C:66:ARG:NH2	11:J:3:VAL:O	2.03	0.91
7:E:96:PHE:HA	7:E:99:HIS:CD2	2.08	0.88
4:A:567:LYS:HB2	9:H:96:VAL:HB	1.56	0.88
11:J:36:LEU:HD13	11:J:47:ARG:HD3	1.55	0.86

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:917:SER:OG	10:I:21:GLU:OE1[4_546]	2.18	0.02

## 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	
4	A	1371/1733 (79%)	1334 (97%)	36 (3%)	1 (0%)	48	77
5	B	1101/1224 (90%)	1078 (98%)	23 (2%)	0	100	100
6	C	265/318 (83%)	260 (98%)	5 (2%)	0	100	100
7	E	211/215 (98%)	205 (97%)	6 (3%)	0	100	100
8	F	84/155 (54%)	82 (98%)	2 (2%)	0	100	100
9	H	129/146 (88%)	124 (96%)	5 (4%)	0	100	100
10	I	116/122 (95%)	112 (97%)	4 (3%)	0	100	100
11	J	63/70 (90%)	62 (98%)	1 (2%)	0	100	100
12	K	112/120 (93%)	111 (99%)	1 (1%)	0	100	100
13	L	41/70 (59%)	39 (95%)	2 (5%)	0	100	100
All	All	3493/4173 (84%)	3407 (98%)	85 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	958	VAL

### 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	1194/1520 (79%)	1194 (100%)	0	100	100
5	B	955/1061 (90%)	955 (100%)	0	100	100
6	C	235/274 (86%)	235 (100%)	0	100	100
7	E	192/197 (98%)	192 (100%)	0	100	100
8	F	73/137 (53%)	73 (100%)	0	100	100
9	H	116/128 (91%)	116 (100%)	0	100	100
10	I	109/116 (94%)	109 (100%)	0	100	100
11	J	60/65 (92%)	60 (100%)	0	100	100
12	K	99/102 (97%)	99 (100%)	0	100	100
13	L	37/57 (65%)	37 (100%)	0	100	100
All	All	3070/3657 (84%)	3070 (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. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	1093	GLN
5	B	1141	HIS
7	E	153	HIS
4	A	881	GLN
4	A	877	HIS

### 5.3.3 RNA ⓘ



Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	9/10 (90%)	1 (11%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	9	G

There are no RNA pucker outliers to report.

## 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	$\# Z  > 2$
2	8OG	T	19	1,2	22,25,26	4.09	18 (81%)	26,37,40	1.68	7 (26%)

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	T	19	1,2	-	3/7/21/22	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	19	8OG	C8-N7	7.45	1.52	1.38
2	T	19	8OG	C8-N9	6.09	1.51	1.40
2	T	19	8OG	C2-N3	6.01	1.47	1.33
2	T	19	8OG	C4-N3	5.81	1.47	1.34
2	T	19	8OG	C2-N2	4.96	1.45	1.34

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	19	8OG	C2-N3-C4	4.50	120.05	112.30
2	T	19	8OG	O6-C6-C5	-3.11	119.77	127.26
2	T	19	8OG	C2-N1-C6	-2.71	120.19	125.11
2	T	19	8OG	C5-C6-N1	2.68	119.49	112.13
2	T	19	8OG	C5-N7-C8	-2.50	106.02	109.47

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	T	19	8OG	O4'-C4'-C5'-O5'
2	T	19	8OG	C3'-C4'-C5'-O5'
2	T	19	8OG	C4'-C5'-O5'-P

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 9 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	$\# Z  > 2$
16	PPV	A	1804	-	6,8,8	0.79	0	12,13,13	0.80	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
16	PPV	A	1804	-	-	1/6/6/6	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

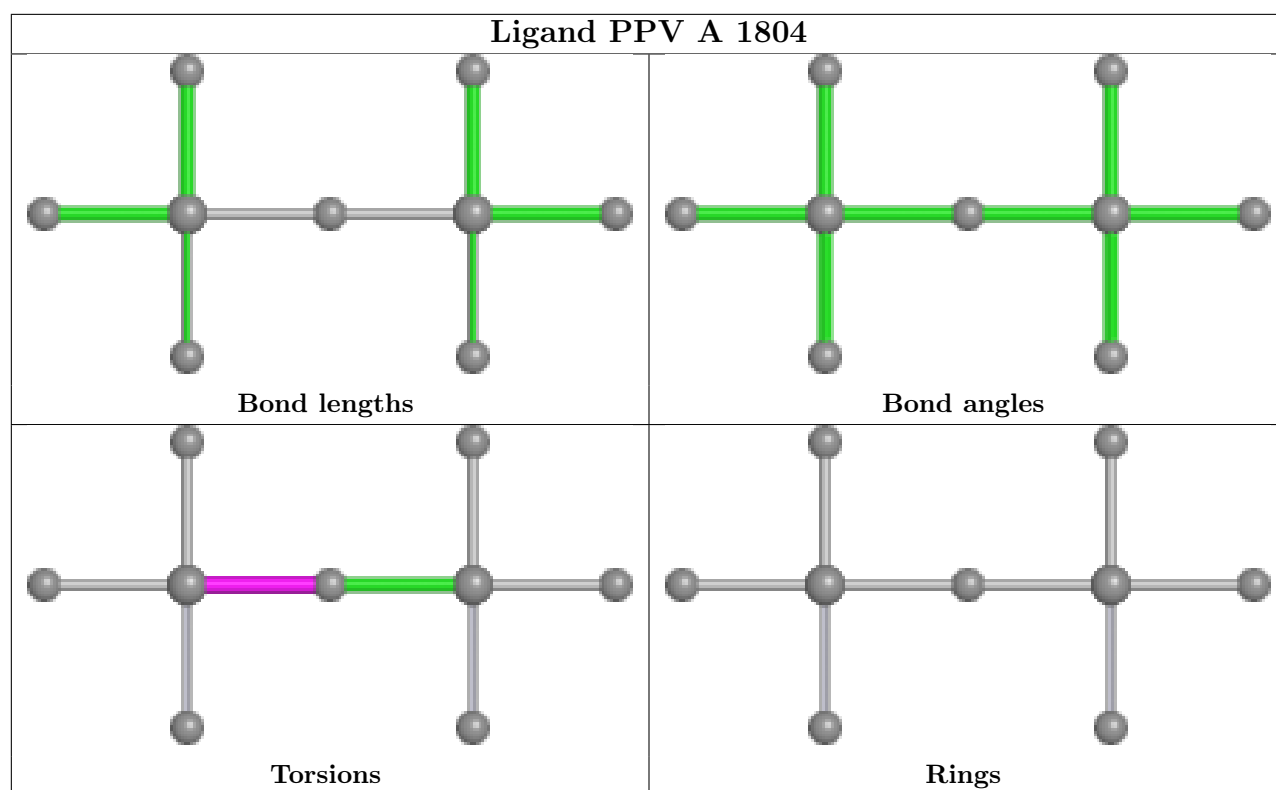
All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
16	A	1804	PPV	P1-OPP-P2-O32

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%)	0.12	1 (10%) 14 16	81, 94, 140, 153	0
2	T	25/29 (86%)	0.20	0 100 100	76, 155, 214, 221	0
3	N	15/18 (83%)	-0.04	0 100 100	150, 173, 221, 227	0
4	A	1385/1733 (79%)	0.41	61 (4%) 39 34	45, 87, 157, 209	0
5	B	1121/1224 (91%)	0.31	25 (2%) 62 53	48, 84, 142, 186	0
6	C	267/318 (83%)	0.16	4 (1%) 71 64	53, 86, 119, 147	0
7	E	213/215 (99%)	0.61	16 (7%) 22 21	61, 119, 178, 195	0
8	F	86/155 (55%)	0.12	2 (2%) 61 51	68, 93, 127, 156	0
9	H	133/146 (91%)	0.41	7 (5%) 33 29	83, 107, 147, 177	0
10	I	118/122 (96%)	0.27	1 (0%) 82 76	68, 100, 123, 135	0
11	J	65/70 (92%)	0.18	2 (3%) 51 44	59, 81, 108, 128	0
12	K	114/120 (95%)	0.17	0 100 100	52, 85, 111, 130	0
13	L	43/70 (61%)	0.67	3 (6%) 24 23	68, 120, 165, 186	0
All	All	3595/4230 (84%)	0.35	122 (3%) 48 41	45, 89, 156, 227	0

The worst 5 of 122 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
7	E	83	CYS	4.6
7	E	121	MET	4.5
4	A	1332	PHE	4.5
7	E	110	PHE	4.2
7	E	93	MET	4.2

## 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	T	19	23/24	0.71	0.12	86,107,128,146	0

## 6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

## 6.4 Ligands [i](#)

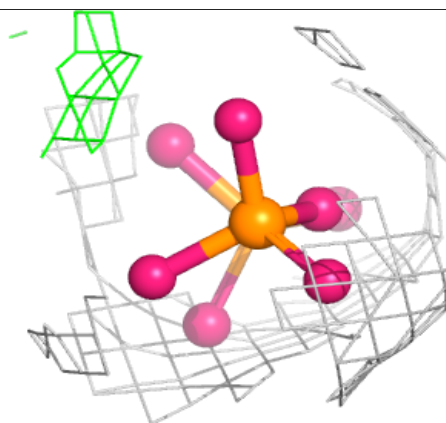
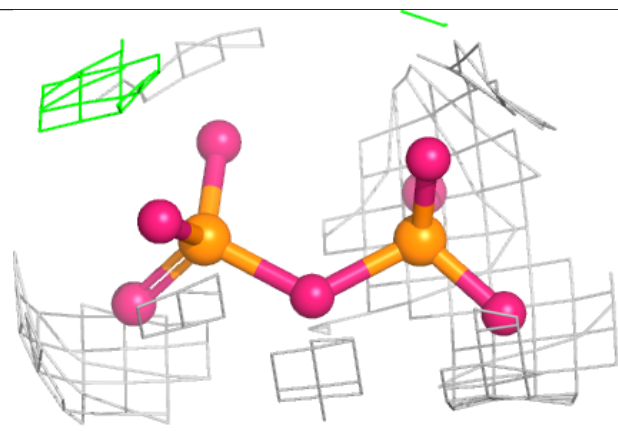
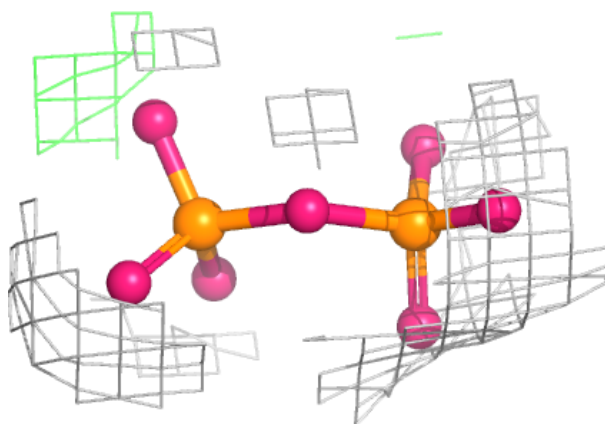
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
16	PPV	A	1804	9/9	0.42	0.12	134,142,165,171	0
14	ZN	L	101	1/1	0.75	0.14	174,174,174,174	0
15	MG	A	1803	1/1	0.83	0.08	88,88,88,88	0
14	ZN	I	202	1/1	0.86	0.10	157,157,157,157	0
14	ZN	B	1301	1/1	0.95	0.06	163,163,163,163	0
14	ZN	J	101	1/1	0.96	0.11	88,88,88,88	0
14	ZN	A	1801	1/1	0.96	0.05	233,233,233,233	0
14	ZN	I	201	1/1	0.97	0.04	89,89,89,89	0
14	ZN	A	1802	1/1	0.97	0.04	161,161,161,161	0
14	ZN	C	401	1/1	1.00	0.09	100,100,100,100	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 PPV A 1804:**

$2mF_o - DF_c$  (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.