



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2020 – 12:30 PM BST

PDB ID : 4Y7N
Title : The Structure Insight into 5-Carboxycytosine Recognition by RNA Polymerase II during Transcription Elongation.
Authors : Wang, L.; Chong, J.; Wang, D.
Deposited on : 2015-02-15
Resolution : 3.30 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

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 : 2.13.1
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) : 2.13.1

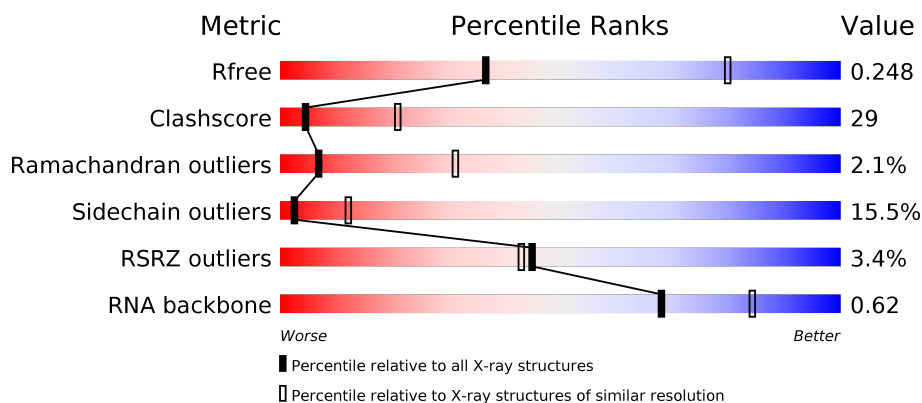
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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.






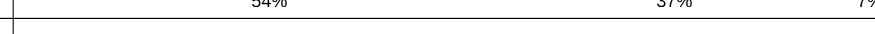




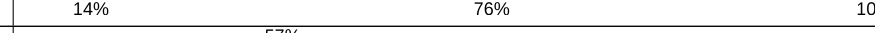
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)
RNA backbone	3102	1117 (3.70-2.90)

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 ≥ 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	1733	<div> <div>3%</div> <div> <div></div> <div>43%</div> <div>30%</div> <div>7%</div> <div>20%</div> </div> </div>
2	B	1224	<div> <div>2%</div> <div> <div></div> <div>48%</div> <div>35%</div> <div>6%</div> <div>10%</div> </div> </div>
3	C	318	<div> <div></div> <div> <div></div> <div>44%</div> <div>34%</div> <div>6%</div> <div>16%</div> </div> </div>
4	E	215	<div> <div>5%</div> <div> <div></div> <div>59%</div> <div>35%</div> <div>6%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	F	155	
6	H	146	
7	I	122	
8	J	70	
9	K	120	
10	L	70	
11	T	29	
12	N	14	
13	R	9	

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 29193 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 RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1393	Total	C	N	O	S	0	0	0
			10953	6908	1921	2063	61			

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2,DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1103	Total	C	N	O	S	0	0	0
			8762	5549	1532	1626	55			

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

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

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	84	Total	C	N	O	S	0	0	0
			679	434	115	127	3			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	L	44	Total	C	N	O	S	0	0	0
			351	217	70	60	4			

- Molecule 11 is a DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	T	29	Total	C	N	O	P	0	0	0
			587	280	108	171	28			

- Molecule 12 is a DNA chain called DNA (5'-D(*CP*TP*GP*CP*TP*TP*AP*TP*CP*GP*GP*TP*AP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	N	14	Total	C	N	O	P	0	0	0
			284	137	49	85	13			

- Molecule 13 is a RNA chain called RNA (5'-D(*AP*UP*GP*GP*AP*GP*AP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	R	9	Total	C	N	O	P	0	0	0
			198	89	42	59	8			

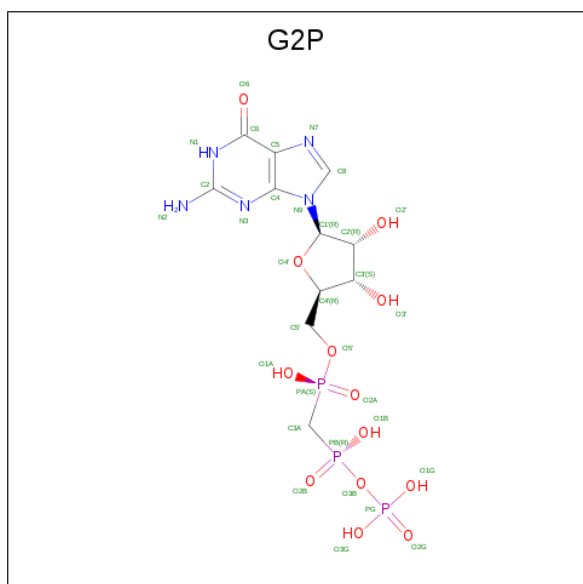
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 16 is PHOSPHOMETHYLPHOSPHONIC ACID GUANYLATE ESTER (three-letter code: G2P) (formula: C₁₁H₁₈N₅O₁₃P₃).

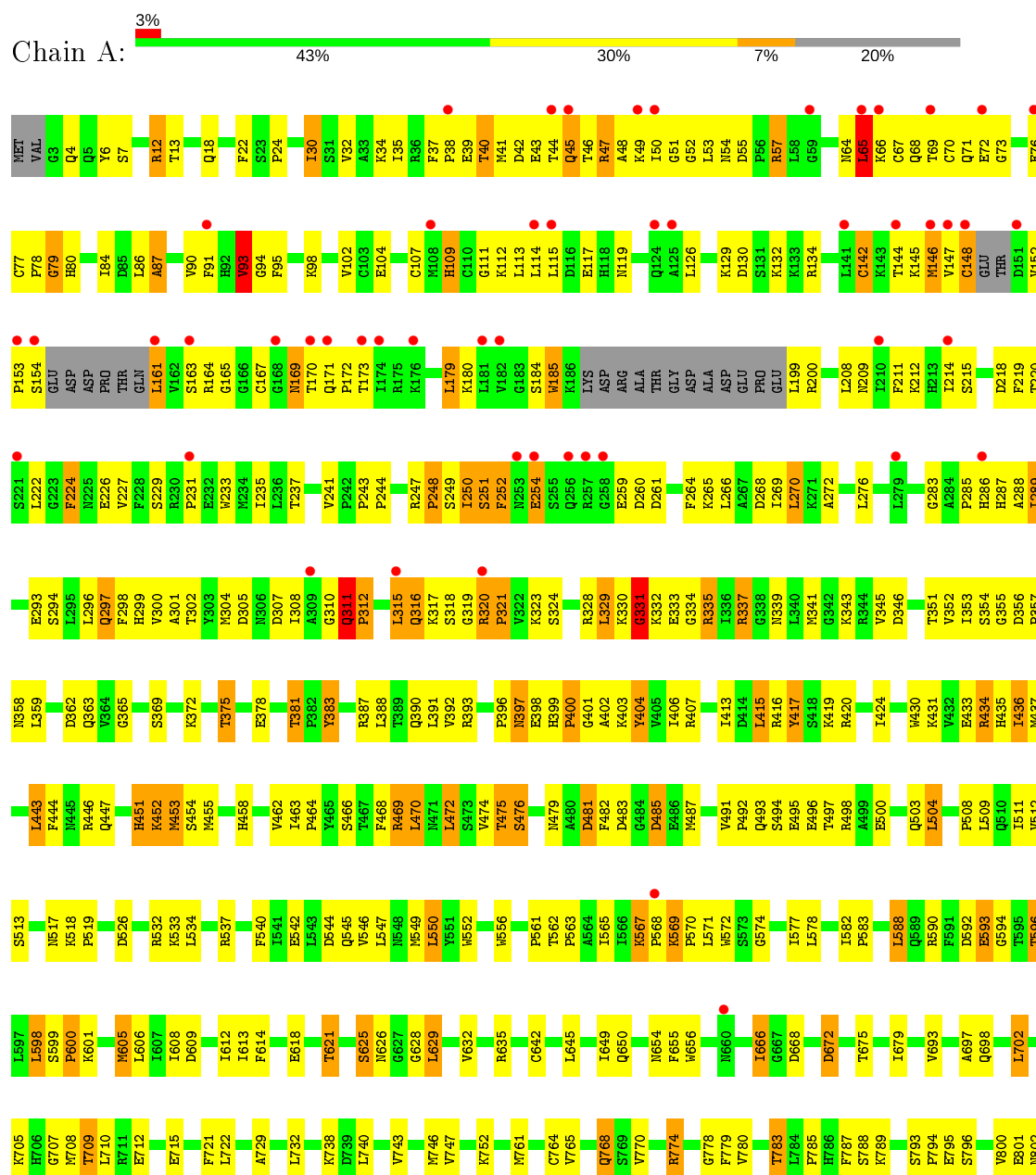


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
16	T	1	Total	C	N	O	P	0	0
			32	11	5	13	3		

3 Residue-property plots

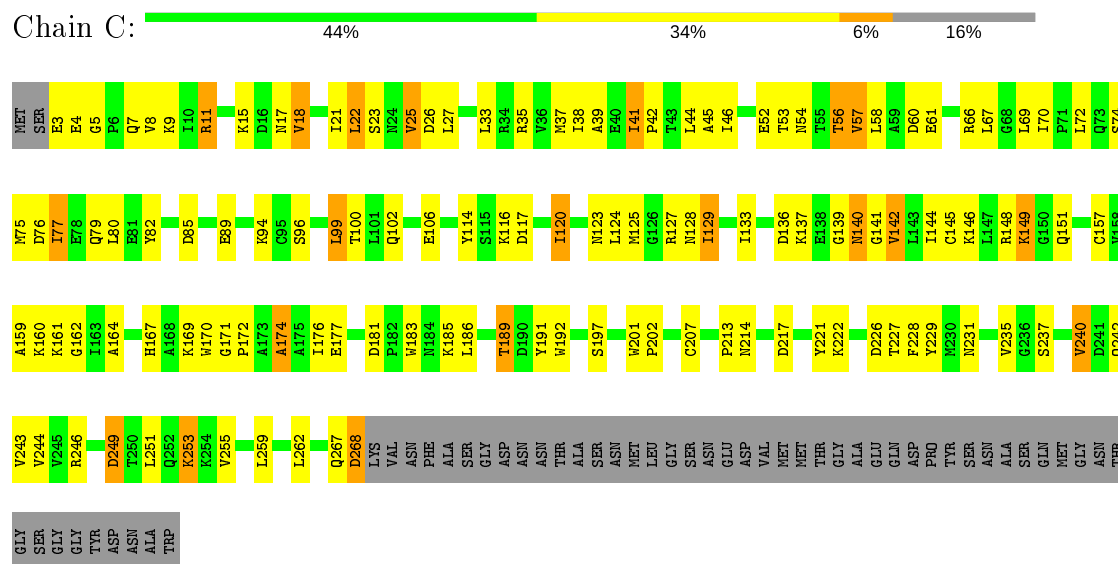
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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 RNA polymerase II subunit RPB1

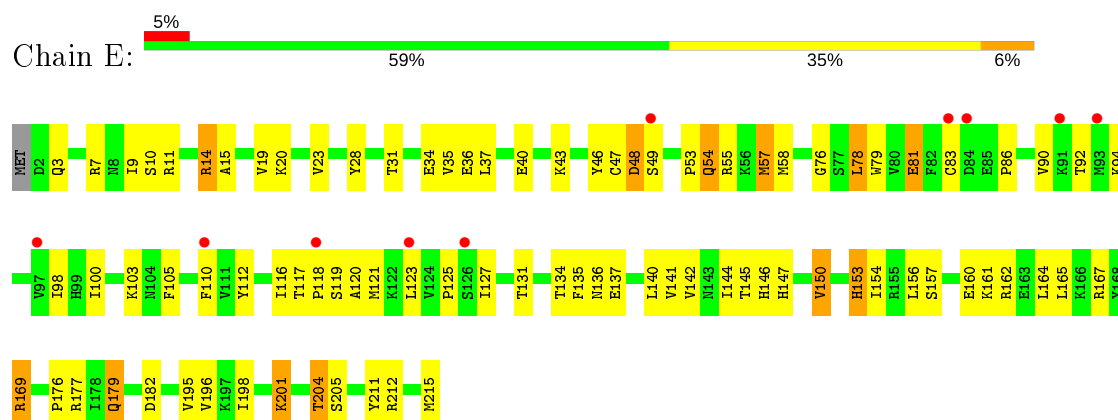




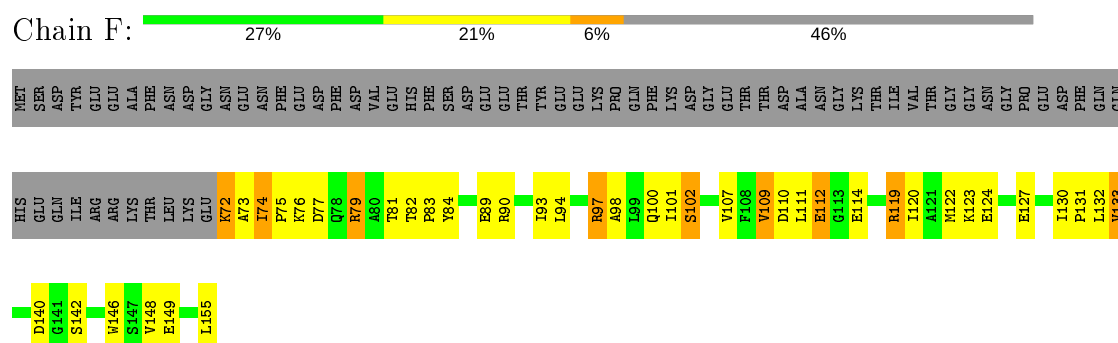
- Molecule 3: DNA-directed RNA polymerase II subunit RPB3



- Molecule 4: DNA-directed RNA polymerases I, II, and III subunit RPABC1

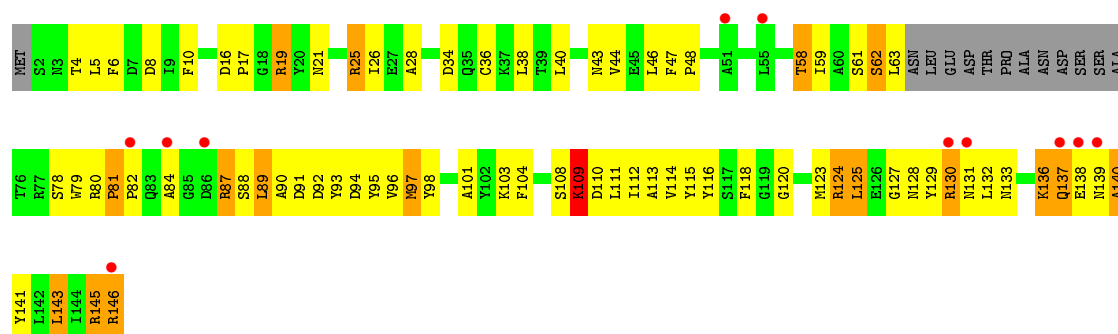


- Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC2

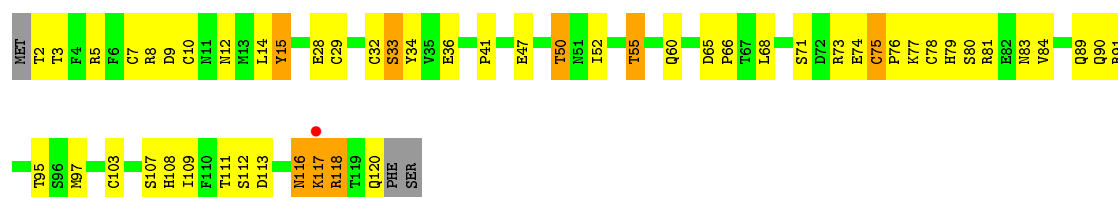


- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC3

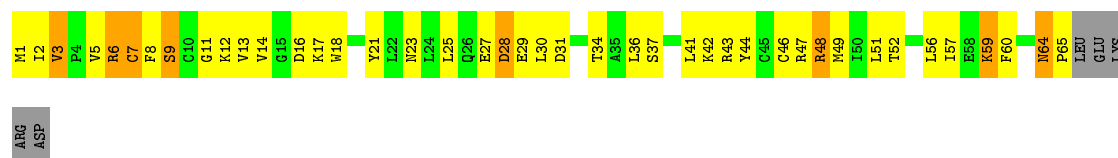




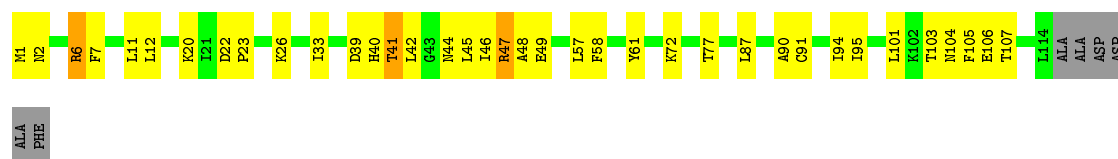
• Molecule 7: DNA-directed RNA polymerase II subunit RPB9



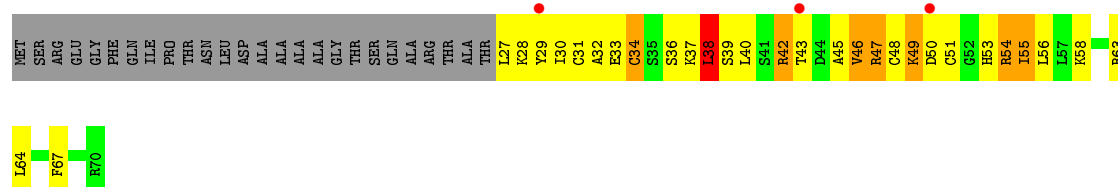
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC5



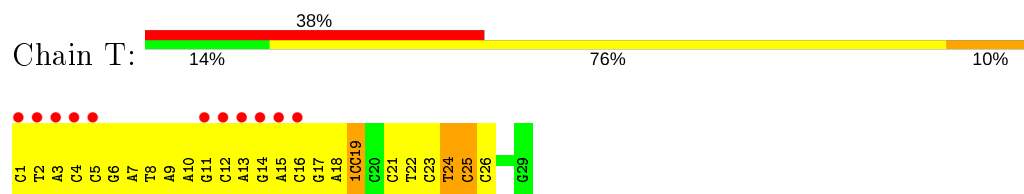
• Molecule 9: DNA-directed RNA polymerase II subunit RPB11



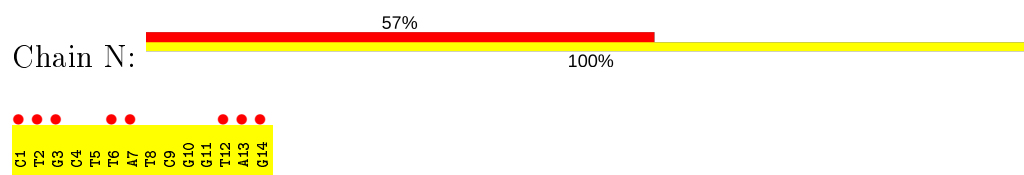
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC4



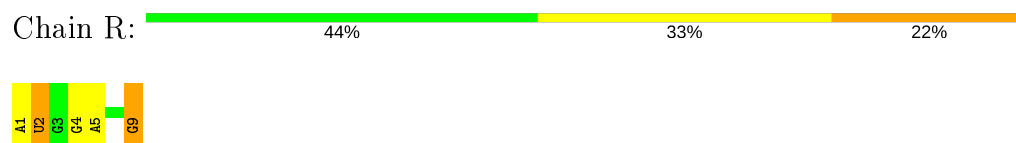
- Molecule 11: DNA (29-MER)



- Molecule 12: DNA (5'-D(*CP*TP*GP*CP*TP*TP*AP*TP*CP*GP*GP*TP*AP*G)-3')



- Molecule 13: RNA (5'-D(*AP*UP*GP*GP*AP*GP*AP*GP*G)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	168.18Å 222.61Å 192.83Å 90.00° 101.59° 90.00°	Depositor
Resolution (Å)	48.88 – 3.30 48.88 – 3.26	Depositor EDS
% Data completeness (in resolution range)	98.0 (48.88-3.30) 98.1 (48.88-3.26)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.64 (at 3.25Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.207 , 0.250 0.205 , 0.248	Depositor DCC
R_{free} test set	5268 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	83.9	Xtriage
Anisotropy	0.481	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 52.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	29193	wwPDB-VP
Average B, all atoms (Å ²)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.70% 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: 1CC, 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	A	0.43	1/11146 (0.0%)	0.65	5/15066 (0.0%)
2	B	0.47	1/8932 (0.0%)	0.67	3/12045 (0.0%)
3	C	0.47	0/2133	0.71	1/2891 (0.0%)
4	E	0.38	0/1788	0.55	0/2406
5	F	0.40	0/691	0.61	0/933
6	H	0.38	0/1086	0.56	0/1470
7	I	0.40	0/989	0.64	1/1331 (0.1%)
8	J	0.49	0/541	0.69	0/727
9	K	0.46	0/937	0.61	0/1265
10	L	0.32	0/353	0.63	0/468
11	T	0.61	0/632	1.17	7/969 (0.7%)
12	N	0.18	0/317	0.56	0/488
13	R	0.56	0/223	1.01	0/348
All	All	0.44	2/29768 (0.0%)	0.67	17/40407 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	707	PRO	N-CD	5.28	1.55	1.47
1	A	400	PRO	N-CD	5.19	1.55	1.47

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	331	GLY	N-CA-C	7.95	132.97	113.10
11	T	24	DT	N3-C4-O4	7.17	124.20	119.90
11	T	24	DT	P-O5'-C5'	-6.69	110.20	120.90
1	A	919	ILE	CB-CA-C	-6.33	98.94	111.60
11	T	25	DC	O4'-C4'-C3'	-6.09	102.06	104.50

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	10953	0	11053	681	0
2	B	8762	0	8797	459	0
3	C	2095	0	2051	104	0
4	E	1752	0	1776	60	0
5	F	679	0	701	31	0
6	H	1068	0	1040	88	0
7	I	971	0	927	36	0
8	J	532	0	542	66	0
9	K	919	0	929	29	0
10	L	351	0	374	65	0
11	T	587	0	326	123	0
12	N	284	0	161	49	0
13	R	198	0	99	9	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	2	0	0	0	0
16	T	32	0	14	2	0
All	All	29193	0	28790	1651	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 29.

The worst 5 of 1651 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:317:LYS:CG	1:A:318:SER:HB2	1.25	1.56
2:B:714:GLU:HA	2:B:715:ALA:CB	1.40	1.45
1:A:317:LYS:HG3	1:A:318:SER:CB	1.46	1.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:317:LYS:CG	1:A:318:SER:CB	1.98	1.42
10:L:38:LEU:HD12	10:L:40:LEU:CG	1.52	1.36

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	1379/1733 (80%)	1169 (85%)	177 (13%)	33 (2%)	6	28
2	B	1085/1224 (89%)	927 (85%)	135 (12%)	23 (2%)	7	31
3	C	264/318 (83%)	228 (86%)	35 (13%)	1 (0%)	34	66
4	E	212/215 (99%)	183 (86%)	25 (12%)	4 (2%)	8	34
5	F	82/155 (53%)	72 (88%)	10 (12%)	0	100	100
6	H	129/146 (88%)	106 (82%)	16 (12%)	7 (5%)	2	12
7	I	117/122 (96%)	98 (84%)	17 (14%)	2 (2%)	9	35
8	J	63/70 (90%)	56 (89%)	6 (10%)	1 (2%)	9	36
9	K	112/120 (93%)	106 (95%)	6 (5%)	0	100	100
10	L	42/70 (60%)	28 (67%)	11 (26%)	3 (7%)	1	7
All	All	3485/4173 (84%)	2973 (85%)	438 (13%)	74 (2%)	7	31

5 of 74 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	286	HIS
1	A	321	PRO
1	A	593	GLU
1	A	1393	ASN
2	B	230	ALA

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	1216/1520 (80%)	1030 (85%)	186 (15%)	2	12
2	B	957/1061 (90%)	813 (85%)	144 (15%)	3	13
3	C	234/274 (85%)	198 (85%)	36 (15%)	2	12
4	E	196/197 (100%)	167 (85%)	29 (15%)	3	13
5	F	74/137 (54%)	59 (80%)	15 (20%)	1	5
6	H	117/128 (91%)	99 (85%)	18 (15%)	2	12
7	I	113/116 (97%)	91 (80%)	22 (20%)	1	5
8	J	60/65 (92%)	50 (83%)	10 (17%)	2	10
9	K	99/102 (97%)	89 (90%)	10 (10%)	7	27
10	L	39/57 (68%)	29 (74%)	10 (26%)	0	2
All	All	3105/3657 (85%)	2625 (84%)	480 (16%)	2	12

5 of 480 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	345	LYS
2	B	755	ILE
7	I	90	GLN
2	B	387	LEU
2	B	481	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	734	HIS
2	B	878	GLN
6	H	137	GLN
2	B	762	ASN
2	B	842	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
13	R	8/9 (88%)	2 (25%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
13	R	2	U
13	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$
11	1CC	T	19	11	15,23,24	1.23	2 (13%)	15,33,36	1.62	3 (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
11	1CC	T	19	11	-	0/4/25/26	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	T	19	1CC	O5'-C5'	-2.51	1.38	1.44
11	T	19	1CC	C2-N3	-2.26	1.33	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	T	19	1CC	C2-N3-C4	4.14	121.02	116.02
11	T	19	1CC	C2'-C1'-N1	-2.86	107.67	114.27
11	T	19	1CC	N4-C4-N3	2.53	120.61	117.03

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	T	19	1CC	3	0

5.5 Carbohydrates [i](#)

There are no monosaccharides 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	$\# Z > 2$
16	G2P	T	101	15	26,34,34	2.81	12 (46%)	30,54,54	1.94	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
16	G2P	T	101	15	-	4/18/38/38	0/3/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	T	101	G2P	C4-N9	-8.27	1.36	1.47
16	T	101	G2P	C6-N1	5.12	1.41	1.33
16	T	101	G2P	PB-O2B	4.31	1.61	1.51
16	T	101	G2P	PA-O2A	4.15	1.61	1.51
16	T	101	G2P	C5-C6	-4.14	1.45	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	T	101	G2P	C4-C5-N7	6.43	110.99	102.46
16	T	101	G2P	C5-C6-N1	-4.09	113.15	118.19
16	T	101	G2P	O4'-C1'-N9	3.50	114.25	109.04
16	T	101	G2P	O1G-PG-O3B	3.03	114.81	104.64
16	T	101	G2P	O2B-PB-C3A	-2.13	103.44	109.07

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
16	T	101	G2P	O4'-C4'-C5'-O5'
16	T	101	G2P	C3'-C4'-C5'-O5'
16	T	101	G2P	C5'-O5'-PA-O1A
16	T	101	G2P	C4'-C5'-O5'-PA

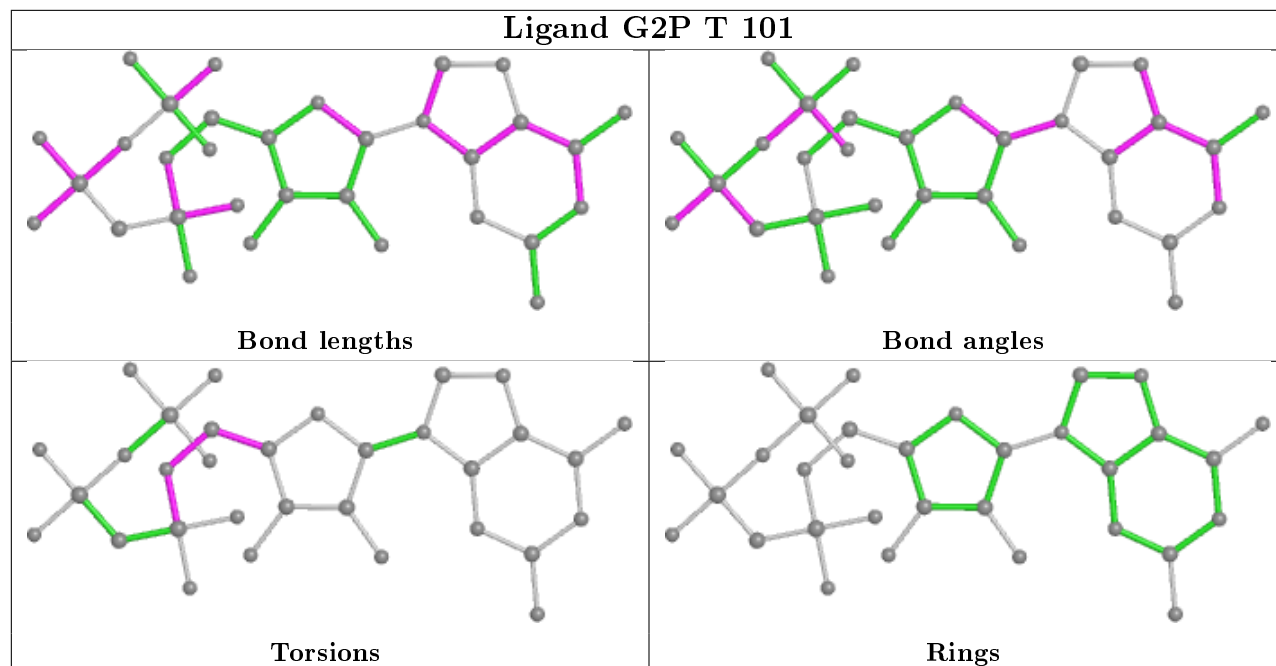
There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	T	101	G2P	2	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 ⓘ

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	1393/1733 (80%)	0.05	58 (4%) 36 34	45, 88, 191, 305	0
2	B	1103/1224 (90%)	-0.08	21 (1%) 66 65	42, 79, 153, 277	0
3	C	266/318 (83%)	-0.31	0 100 100	50, 79, 117, 170	0
4	E	214/215 (99%)	0.04	10 (4%) 31 29	63, 114, 201, 236	0
5	F	84/155 (54%)	-0.24	0 100 100	65, 93, 138, 182	0
6	H	133/146 (91%)	0.46	11 (8%) 11 11	75, 116, 207, 288	0
7	I	119/122 (97%)	-0.18	1 (0%) 86 86	58, 97, 139, 186	0
8	J	65/70 (92%)	-0.23	0 100 100	56, 74, 113, 131	0
9	K	114/120 (95%)	-0.13	0 100 100	54, 86, 115, 164	0
10	L	44/70 (62%)	0.69	3 (6%) 17 17	73, 158, 247, 292	0
11	T	28/29 (96%)	1.43	11 (39%) 0 0	60, 289, 368, 372	0
12	N	14/14 (100%)	1.82	8 (57%) 0 0	249, 290, 330, 348	0
13	R	9/9 (100%)	-0.47	0 100 100	54, 66, 111, 116	0
All	All	3586/4225 (84%)	-0.00	123 (3%) 45 43	42, 87, 187, 372	0

The worst 5 of 123 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	257	ARG	6.9
6	H	86	ASP	6.5
2	B	1221	SER	6.1
2	B	250	PHE	5.8
10	L	50	ASP	5.7

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, 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
11	1CC	T	19	22/23	0.95	0.13	71,87,132,144	0

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

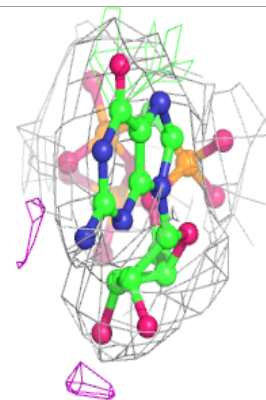
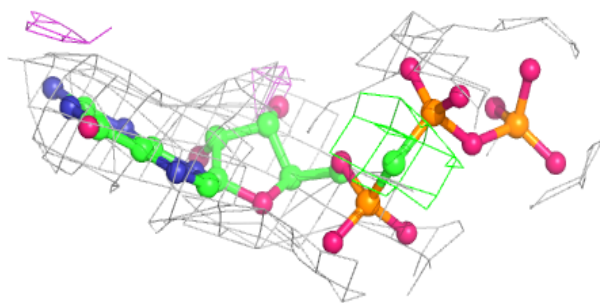
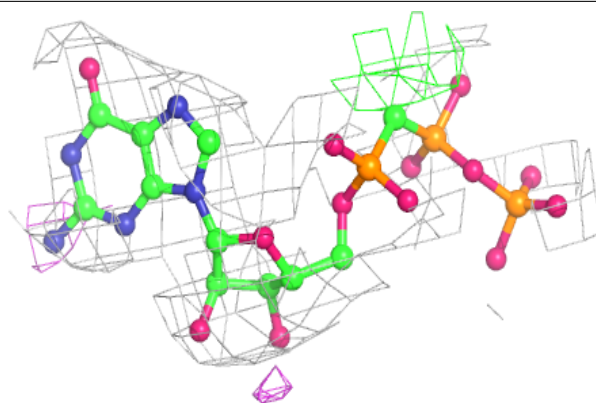
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
14	ZN	A	1801	1/1	0.70	0.08	266,266,266,266	0
15	MG	A	1804	1/1	0.83	0.20	87,87,87,87	0
14	ZN	L	101	1/1	0.87	0.06	172,172,172,172	0
16	G2P	T	101	32/32	0.91	0.18	59,73,127,131	0
14	ZN	A	1802	1/1	0.94	0.08	136,136,136,136	0
14	ZN	B	1301	1/1	0.97	0.12	123,123,123,123	0
14	ZN	J	101	1/1	0.97	0.19	92,92,92,92	0
14	ZN	I	202	1/1	0.97	0.13	93,93,93,93	0
15	MG	A	1803	1/1	0.97	0.17	42,42,42,42	0
14	ZN	I	201	1/1	0.98	0.09	104,104,104,104	0
14	ZN	C	401	1/1	0.99	0.12	107,107,107,107	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 G2P T 101:

$2mF_o-DF_c$ (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.