wwPDB X-ray Structure Validation Summary Report (i)

Apr 28, 2024 – 06:07 pm BST

PDB ID	:	4A3J
Title	:	RNA Polymerase II initial transcribing complex with a 2nt DNA-RNA hybrid and soaked with GMPCPP
Authors	:	Cheung, A.C.M.; Sainsbury, S.; Cramer, P.
Deposited on	:	2011-09-30
Resolution	:	3.70 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 (i) 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 (i)) were used in the production of this report:

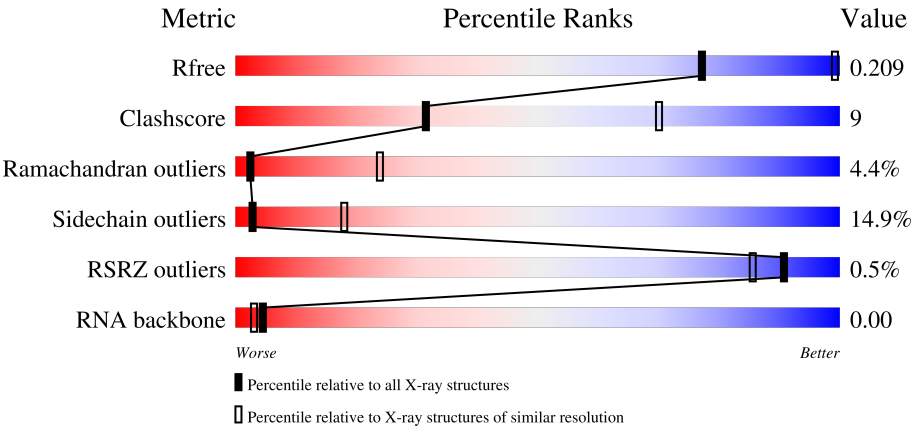
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.70 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.



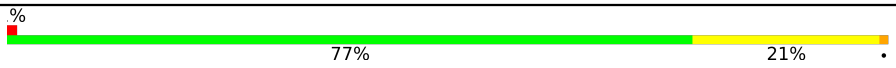
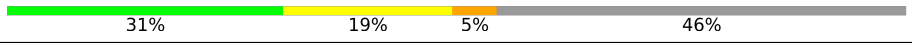
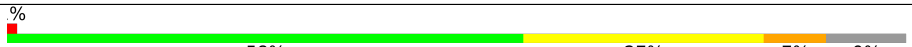
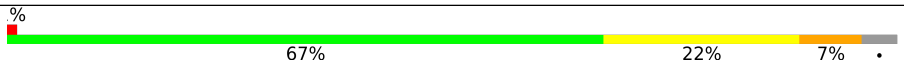
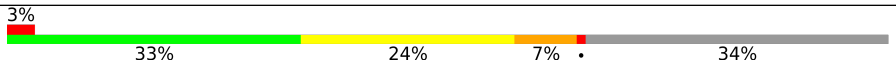
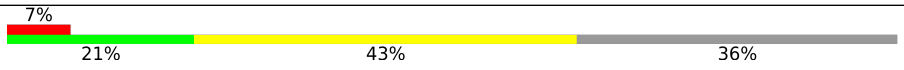
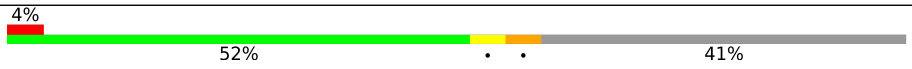
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)
RSRZ outliers	127900	1578 (3.90-3.50)
RNA backbone	3102	1027 (4.40-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	1732	54% 23% • • 18%
2	B	1224	% 63% 23% • • 9%
3	C	318	54% 25% 5% • 16%
4	D	221	47% 29% 5% • 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	215	
6	F	155	
7	G	171	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	
13	N	14	
14	P	2	
15	T	27	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
18	G2P	T	1024	-	-	-	X

2 Entry composition (i)

There are 18 unique types of molecules in this entry. The entry contains 31802 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
			Total	C	N	O	S			
1	A	1425	11197	7051	1958	2126	62	0	0	0

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1115	8859	5609	1554	1641	55	0	0	0

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

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

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	178	1434	887	257	288	2	0	0	0

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1.

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

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2.

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

- Molecule 7 is a protein called RPB7, DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3.

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

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

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

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

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

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	115	Total	C	N	O	S	0	0	1
			920	590	157	171	2			

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4.

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

- Molecule 13 is a DNA chain called 5'-D(*GP*TP*AP*GP*AP*AP*AP*GP*CP*TP*AP*GP*CP*TP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	N	9	Total	C	N	O	P	0	0	0
			184	89	37	50	8			

- Molecule 14 is a RNA chain called 5'-R(*CP*AP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	P	2	Total	C	N	O	P	0	0	0
			42	19	8	13	2			

- Molecule 15 is a DNA chain called 5'-D(*AP*GP*CP*TP*AP*GP*CP*TP*TP*TP*CP*BRUP*AP*CP*CP *TP*GP*AP*AP*CP*AP*AP*CP*TP*AP*AP*CP)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
15	T	16	Total	Br	C	N	O	P	0	0	0
			325	1	155	57	96	16			

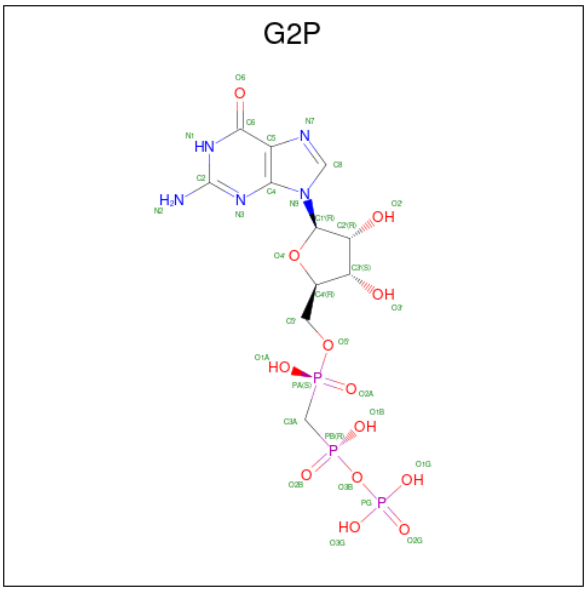
- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

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



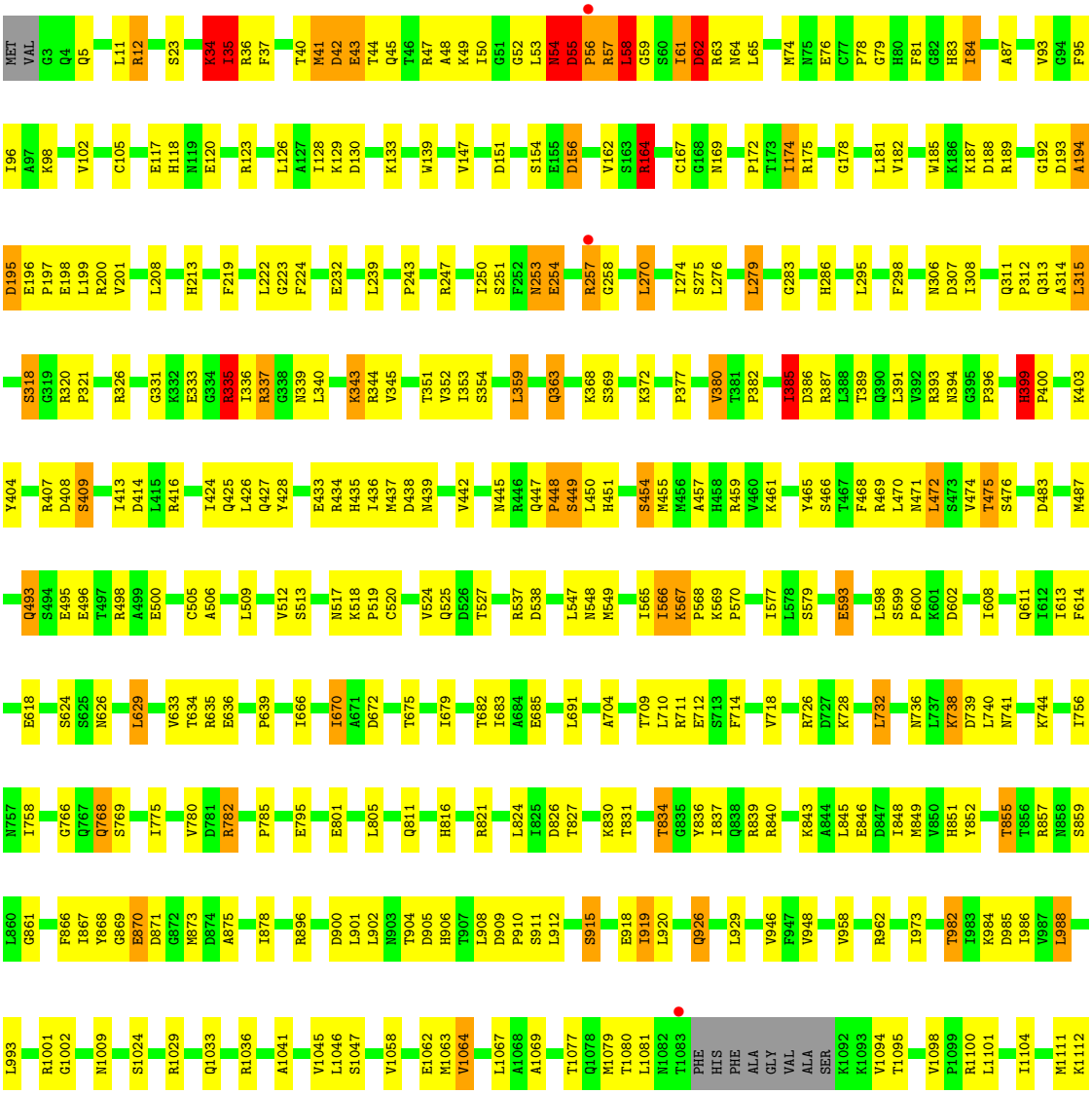
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
18	T	1	Total	C	N	O	P	0	0
			32	11	5	13	3		

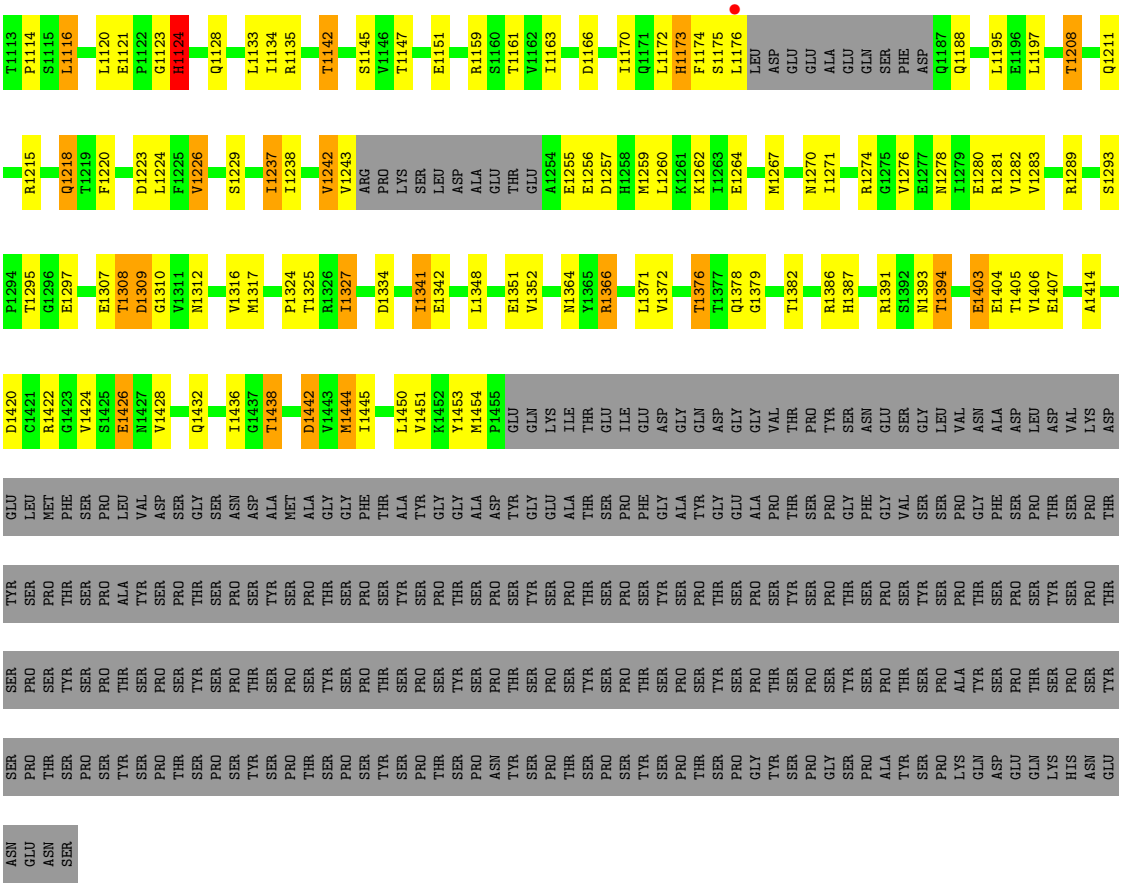
3 Residue-property plots (i)

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: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB1

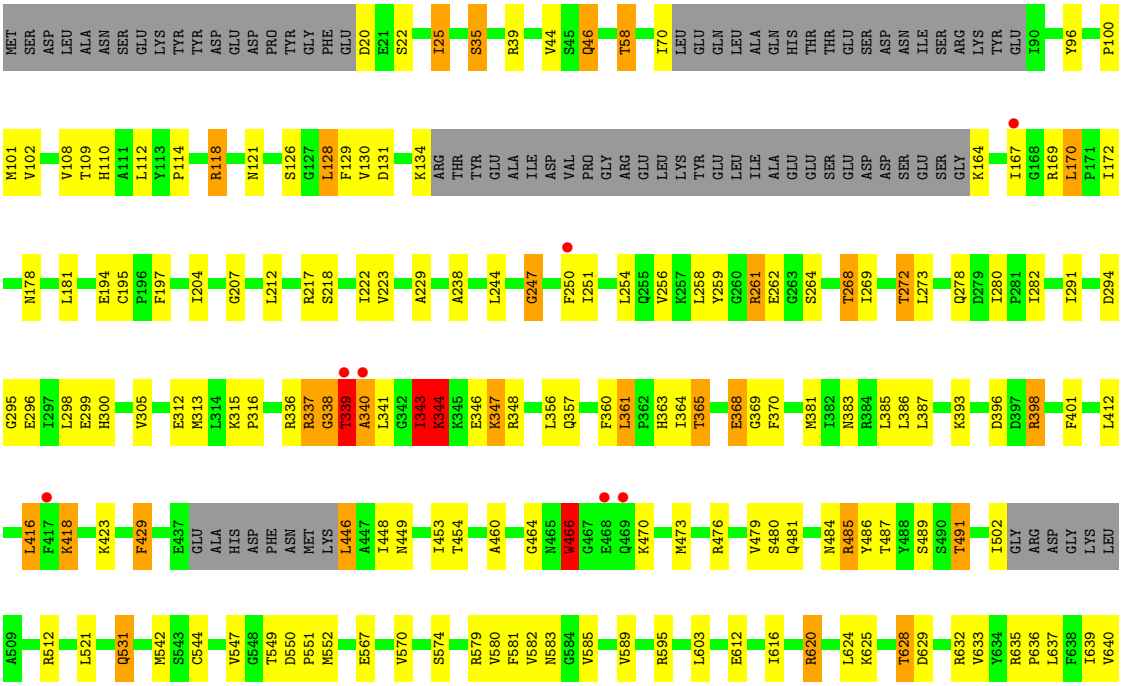
Chain A:

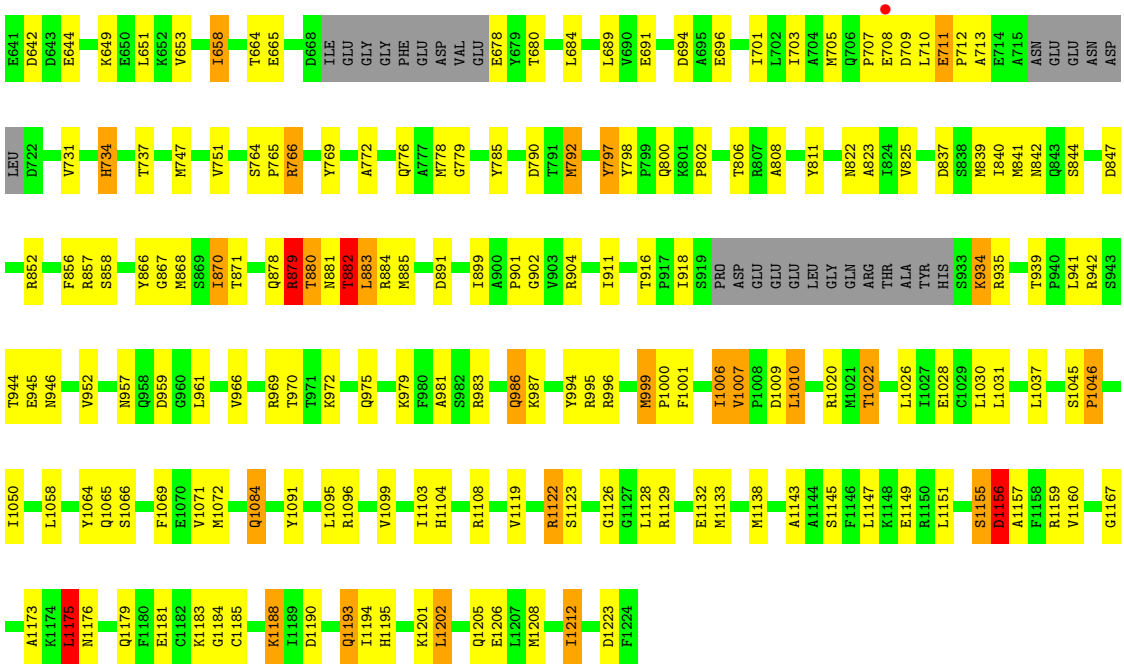




• Molecule 2: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2

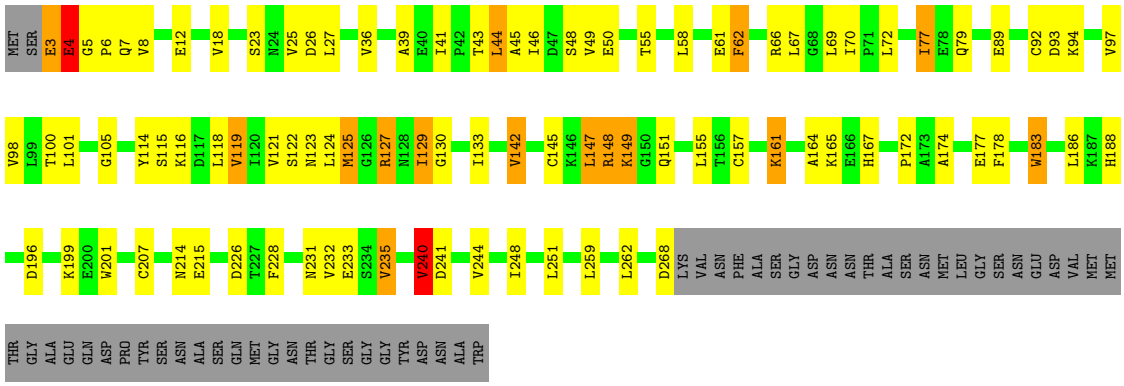
Chain B:





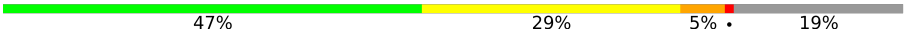
• Molecule 3: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3

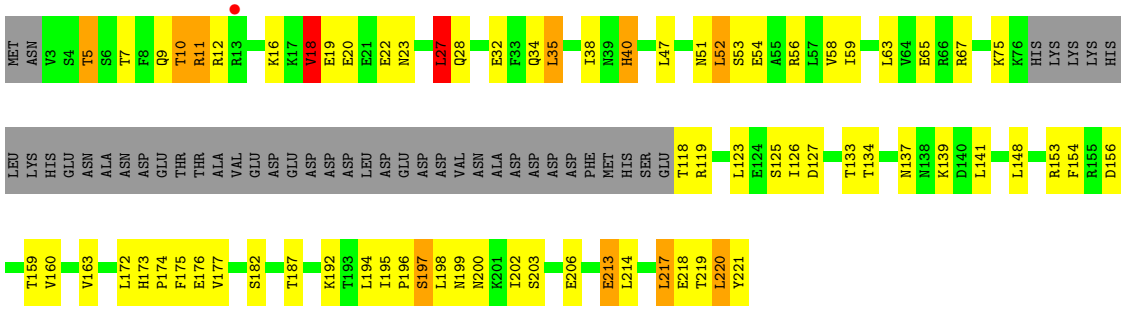
Chain C:



• Molecule 4: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4

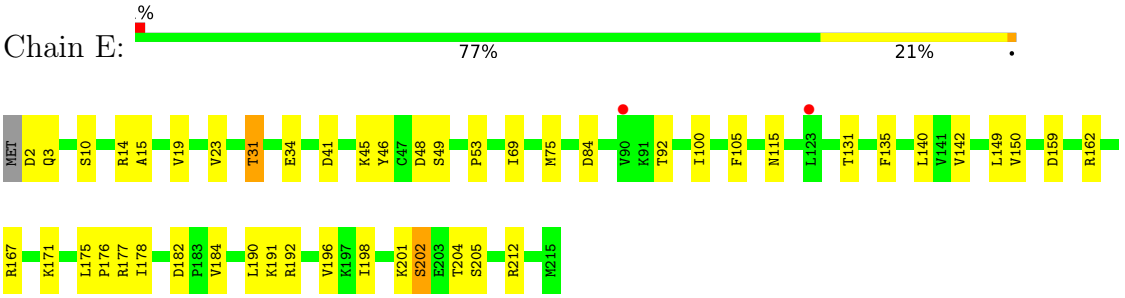
Chain D:





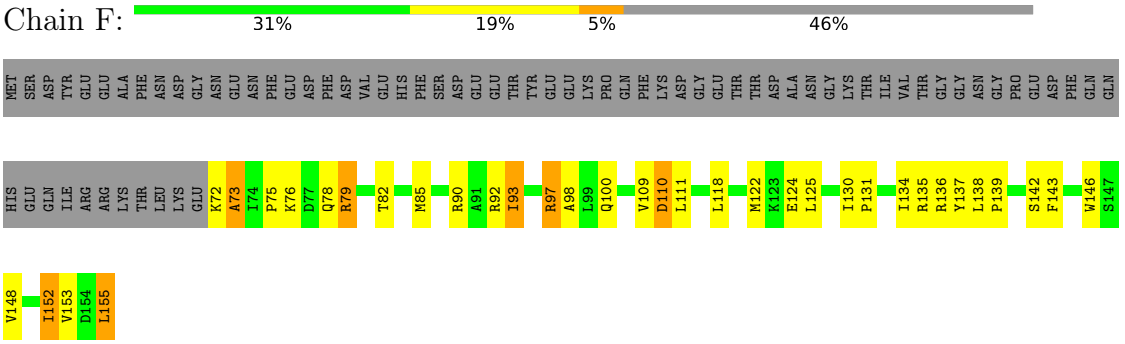
• Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1

Chain E:



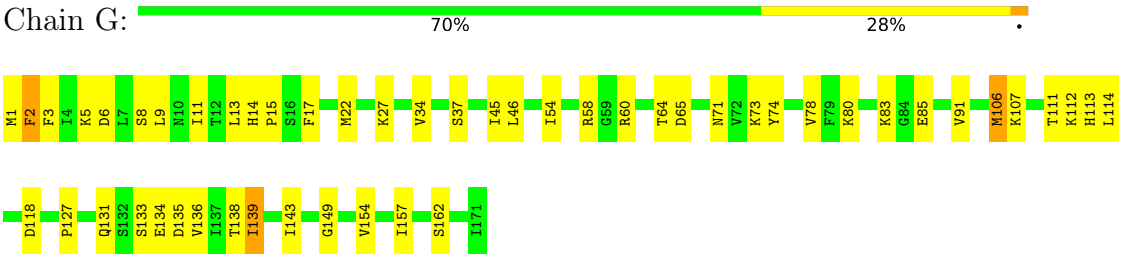
• Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

Chain F:



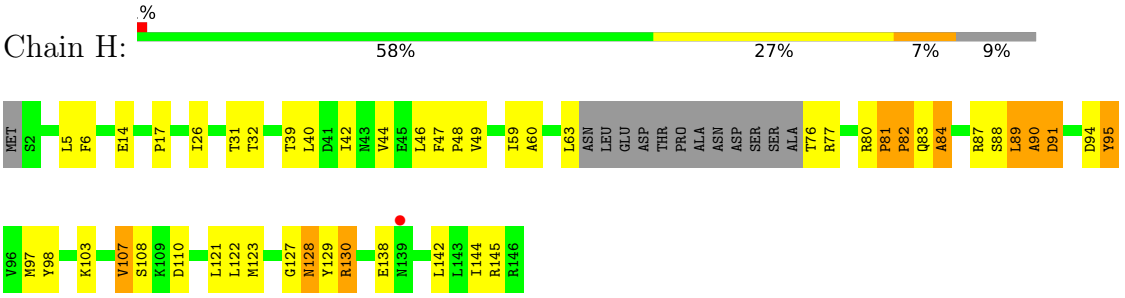
• Molecule 7: RPB7, DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7

Chain G:



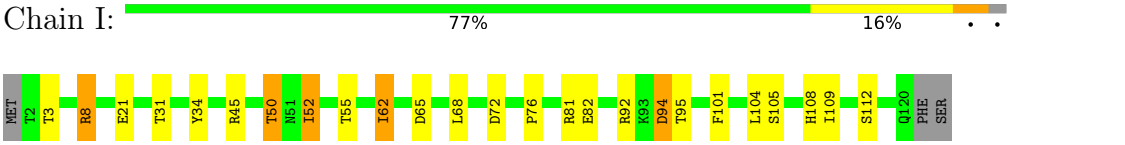
• Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

Chain H:



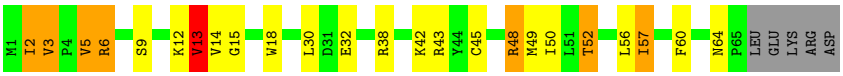
• Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9

Chain I:



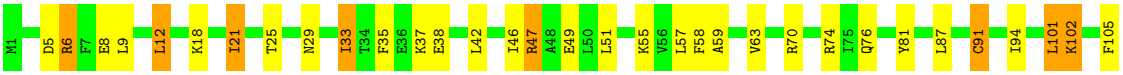
• Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5

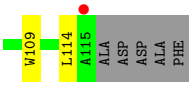
Chain J:



• Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11

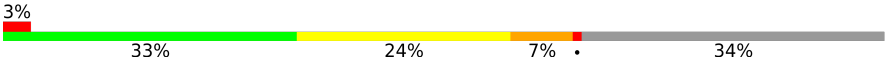
Chain K:

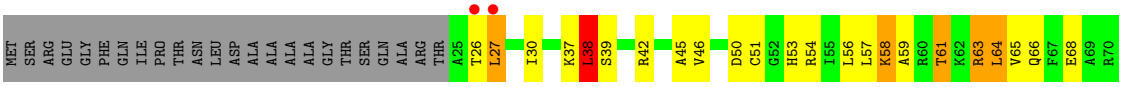




• Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4

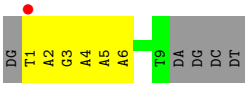
Chain L:





• Molecule 13: 5'-D(*GP*TP*AP*GP*AP*AP*AP*GP*CP*TP*AP*GP*CP*TP)-3'

Chain N:

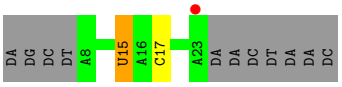


• Molecule 14: 5'-R(*CP*AP)-3'

Chain P:

• Molecule 15: 5'-D(*AP*GP*CP*TP*AP*GP*CP*TP*TP*TP*CP*BRUP*AP*CP*CP *TP* GP*AP*AP*CP*AP*AP*CP*TP*AP*AP*CP)-3'

Chain T:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	223.29Å 392.79Å 282.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.70 74.11 – 3.70	Depositor EDS
% Data completeness (in resolution range)	97.3 (50.00-3.70) 97.2 (74.11-3.70)	Depositor EDS
R_{merge}	0.65	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.92 (at 3.67Å)	Xtriage
Refinement program	BUSTER 2.11.2	Depositor
R, R_{free}	0.157 , 0.194 0.180 , 0.209	Depositor DCC
R_{free} test set	2526 reflections (1.97%)	wwPDB-VP
Wilson B-factor (Å^2)	117.0	Xtriage
Anisotropy	0.292	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 111.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.42$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.040 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.057 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	31802	wwPDB-VP
Average B, all atoms (Å^2)	136.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.97% 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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, BRU, ZN, 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	0.55	0/11397	0.86	10/15415 (0.1%)
2	B	0.52	0/9029	0.82	4/12171 (0.0%)
3	C	0.50	0/2133	0.80	1/2891 (0.0%)
4	D	0.54	0/1444	0.85	0/1935
5	E	0.47	0/1788	0.73	0/2406
6	F	0.58	0/691	0.81	0/933
7	G	0.51	0/1368	0.81	0/1844
8	H	0.53	0/1086	0.84	0/1470
9	I	0.46	0/989	0.79	0/1331
10	J	0.52	0/541	0.85	1/727 (0.1%)
11	K	0.50	0/938	0.73	0/1267
12	L	0.54	0/365	0.93	0/485
13	N	1.33	2/207 (1.0%)	1.11	0/318
14	P	1.39	1/46 (2.2%)	0.81	0/69
15	T	1.25	0/340	0.97	0/519
All	All	0.55	3/32362 (0.0%)	0.83	16/43781 (0.0%)

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
1	A	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	N	1	DT	C1'-N1	5.97	1.57	1.49
14	P	9	C	C1'-N1	5.92	1.57	1.48
13	N	1	DT	C3'-O3'	5.27	1.50	1.44

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	56	PRO	C-N-CA	6.46	137.85	121.70
2	B	340	ALA	C-N-CA	6.45	137.81	121.70
1	A	399	HIS	N-CA-CB	6.42	122.16	110.60
2	B	881	ASN	C-N-CA	6.39	137.67	121.70
1	A	34	LYS	C-N-CA	6.35	137.58	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	34	LYS	Mainchain,Peptide

5.2 Too-close contacts (i)

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	11197	0	11257	219	0
2	B	8859	0	8901	158	0
3	C	2095	0	2051	58	0
4	D	1434	0	1460	33	0
5	E	1752	0	1776	21	0
6	F	679	0	701	25	0
7	G	1340	0	1357	33	0
8	H	1068	0	1040	28	0
9	I	971	0	927	11	0
10	J	532	0	542	20	0
11	K	920	0	929	26	0
12	L	363	0	386	9	0
13	N	184	0	103	4	0
14	P	42	0	22	0	0
15	T	325	0	179	2	0
16	A	2	0	0	0	0
16	B	1	0	0	0	0
16	C	1	0	0	0	0
16	I	2	0	0	0	0
16	J	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	L	1	0	0	0	0
17	A	1	0	0	0	0
18	T	32	0	14	0	0
All	All	31802	0	31645	570	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 9.

The worst 5 of 570 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:867:ILE:CG1	1:A:867:ILE:CD1	1.76	1.62
1:A:53:LEU:HD23	1:A:54:ASN:H	1.12	1.15
1:A:855:THR:HG21	1:A:857:ARG:HE	1.23	1.01
1:A:868:TYR:CE1	1:A:1064:VAL:HG11	2.04	0.93
1:A:55:ASP:HA	1:A:58:LEU:HB2	1.51	0.92

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	1417/1732 (82%)	1205 (85%)	143 (10%)	69 (5%)	2 22
2	B	1095/1224 (90%)	940 (86%)	109 (10%)	46 (4%)	3 25
3	C	264/318 (83%)	238 (90%)	20 (8%)	6 (2%)	6 36
4	D	174/221 (79%)	151 (87%)	9 (5%)	14 (8%)	1 12
5	E	212/215 (99%)	189 (89%)	19 (9%)	4 (2%)	8 40
6	F	82/155 (53%)	75 (92%)	6 (7%)	1 (1%)	13 48
7	G	169/171 (99%)	150 (89%)	16 (10%)	3 (2%)	8 41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
8	H	129/146 (88%)	99 (77%)	19 (15%)	11 (8%)	1 10
9	I	117/122 (96%)	98 (84%)	16 (14%)	3 (3%)	5 34
10	J	63/70 (90%)	56 (89%)	3 (5%)	4 (6%)	1 17
11	K	113/120 (94%)	105 (93%)	8 (7%)	0	100 100
12	L	44/70 (63%)	27 (61%)	8 (18%)	9 (20%)	0 1
All	All	3879/4564 (85%)	3333 (86%)	376 (10%)	170 (4%)	2 24

5 of 170 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	43	GLU
1	A	48	ALA
1	A	58	LEU
1	A	189	ARG
1	A	193	ASP

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	1243/1519 (82%)	1043 (84%)	200 (16%)	2 15
2	B	966/1061 (91%)	819 (85%)	147 (15%)	3 17
3	C	234/274 (85%)	198 (85%)	36 (15%)	2 17
4	D	160/200 (80%)	131 (82%)	29 (18%)	1 11
5	E	196/197 (100%)	174 (89%)	22 (11%)	6 28
6	F	74/137 (54%)	63 (85%)	11 (15%)	3 18
7	G	152/152 (100%)	141 (93%)	11 (7%)	14 45
8	H	117/128 (91%)	99 (85%)	18 (15%)	2 17
9	I	113/116 (97%)	105 (93%)	8 (7%)	14 45
10	J	60/65 (92%)	53 (88%)	7 (12%)	5 26
11	K	99/102 (97%)	85 (86%)	14 (14%)	3 20

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
12	L	40/57 (70%)	28 (70%)	12 (30%)	0 2
All	All	3454/4008 (86%)	2939 (85%)	515 (15%)	3 18

5 of 515 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	H	88	SER
9	I	31	THR
8	H	77	ARG
1	A	1442	ASP
1	A	1420	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	363	HIS
2	B	1025	HIS
2	B	366	GLN
2	B	842	ASN
4	D	132	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	1/2 (50%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

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	$\#\|Z\| > 2$
15	BRU	T	15	13,15	18,21,22	0.87	0	26,30,33	2.03	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
15	BRU	T	15	13,15	-	2/7/21/22	0/2/2/2

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
15	T	15	BRU	C4-N3-C2	-4.80	121.14	127.35
15	T	15	BRU	O4-C4-C5	-4.45	120.26	125.84
15	T	15	BRU	N3-C2-N1	3.90	120.07	114.89
15	T	15	BRU	C5-C4-N3	3.59	117.47	113.34
15	T	15	BRU	O4'-C1'-N1	3.37	113.89	107.86

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
15	T	15	BRU	C3'-C4'-C5'-O5'
15	T	15	BRU	O4'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	T	15	BRU	1	0

5.5 Carbohydrates (i)

There are no monosaccharides 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	$\#\|Z\| > 2$
18	G2P	T	1024	-	27,34,34	2.55	7 (25%)	33,54,54	2.85	7 (21%)

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
18	G2P	T	1024	-	-	2/15/38/38	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	T	1024	G2P	C6-N1	7.77	1.46	1.33
18	T	1024	G2P	PA-O5'	5.08	1.64	1.57
18	T	1024	G2P	C4-N3	4.06	1.42	1.35
18	T	1024	G2P	PB-O3B	3.97	1.62	1.58
18	T	1024	G2P	C2-N1	3.94	1.42	1.35

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
18	T	1024	G2P	PB-O3B-PG	11.91	174.55	132.62
18	T	1024	G2P	C5-C6-N1	-7.52	113.15	123.43
18	T	1024	G2P	C2-N1-C6	4.09	122.42	115.93
18	T	1024	G2P	O2A-PA-C3A	3.07	117.19	109.07
18	T	1024	G2P	O1G-PG-O2G	2.49	120.42	110.68

There are no chirality outliers.

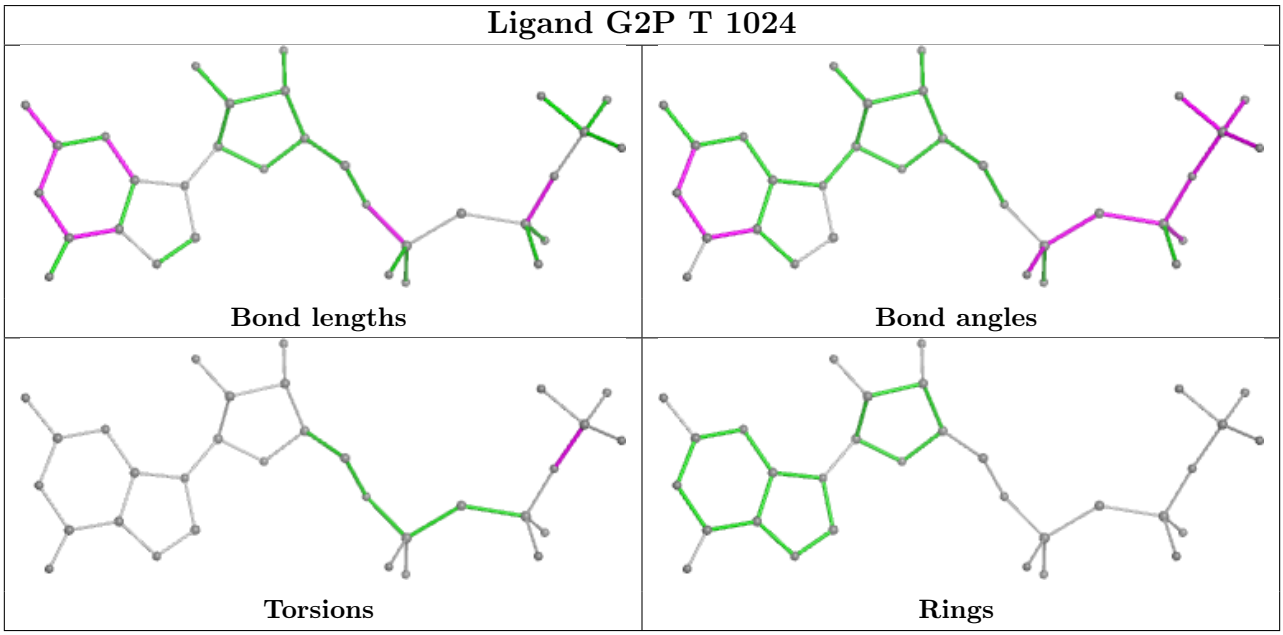
All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
18	T	1024	G2P	PB-O3B-PG-O1G
18	T	1024	G2P	PB-O3B-PG-O3G

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.



5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	B	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	934:LYS	C	935:ARG	N	5.74
1	B	351:TYR	C	352:ALA	N	3.20

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, 95^{th} 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	1425/1732 (82%)	-0.15	4 (0%) 94 90	68, 122, 180, 239	0
2	B	1115/1224 (91%)	-0.02	8 (0%) 87 81	71, 135, 194, 224	0
3	C	266/318 (83%)	-0.28	0 100 100	89, 120, 158, 195	0
4	D	178/221 (80%)	-0.14	1 (0%) 89 83	95, 134, 183, 201	0
5	E	214/215 (99%)	-0.12	2 (0%) 84 76	97, 158, 205, 214	0
6	F	84/155 (54%)	-0.34	0 100 100	75, 103, 132, 149	0
7	G	171/171 (100%)	-0.13	0 100 100	94, 119, 151, 181	0
8	H	133/146 (91%)	-0.10	1 (0%) 86 78	135, 163, 198, 220	0
9	I	119/122 (97%)	-0.27	0 100 100	123, 168, 204, 216	0
10	J	65/70 (92%)	-0.23	0 100 100	101, 117, 161, 173	0
11	K	115/120 (95%)	-0.30	1 (0%) 84 76	87, 119, 157, 171	0
12	L	46/70 (65%)	0.02	2 (4%) 35 26	110, 185, 203, 208	0
13	N	9/14 (64%)	0.84	1 (11%) 5 4	236, 246, 283, 284	0
14	P	2/2 (100%)	1.47	0 100 100	216, 216, 216, 220	0
15	T	15/27 (55%)	0.26	1 (6%) 17 12	191, 217, 273, 276	0
All	All	3957/4607 (85%)	-0.12	21 (0%) 91 85	68, 129, 194, 284	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
13	N	1	DT	4.6
1	A	1176	LEU	3.7
2	B	339	THR	3.2
2	B	708	GLU	3.1
12	L	27	LEU	3.0

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^{th} 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
15	BRU	T	15	20/21	0.81	0.22	214,225,227,233	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

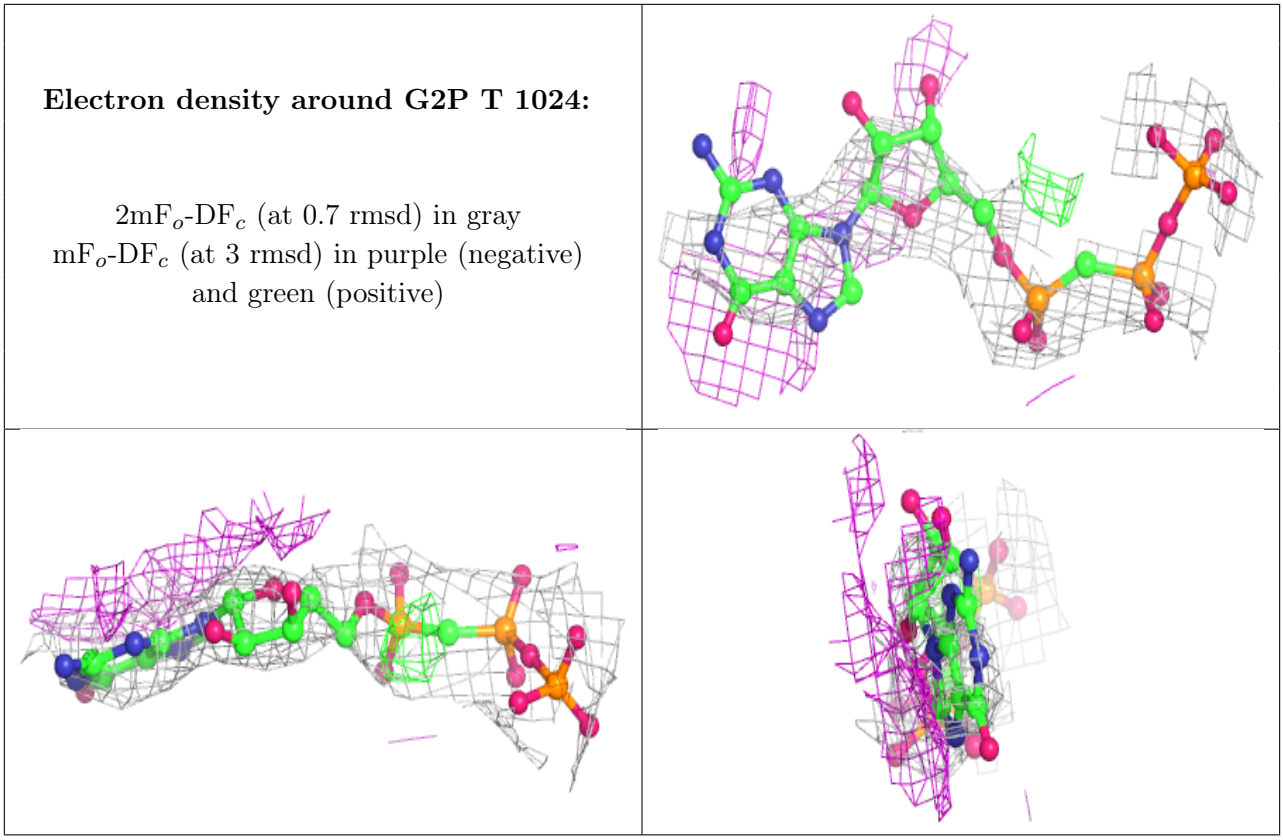
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
18	G2P	T	1024	32/32	0.67	0.47	221,226,247,248	0
17	MG	A	2458	1/1	0.97	0.22	100,100,100,100	0
16	ZN	I	1122	1/1	0.98	0.09	217,217,217,217	0
16	ZN	L	1071	1/1	0.98	0.08	194,194,194,194	0
16	ZN	J	1066	1/1	0.99	0.21	95,95,95,95	0
16	ZN	A	2456	1/1	0.99	0.12	157,157,157,157	0
16	ZN	B	2225	1/1	1.00	0.21	97,97,97,97	0
16	ZN	C	1269	1/1	1.00	0.20	96,96,96,96	0
16	ZN	I	1121	1/1	1.00	0.15	142,142,142,142	0
16	ZN	A	2457	1/1	1.00	0.18	82,82,82,82	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 1024:

$2mF_o$-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.