wwPDB X-ray Structure Validation Summary Report (i)

Jun 1, 2026 – 12:12 PM EDT

PDB ID	:	9PVV / pdb_00009pvv
Title	:	RNA polymerase II elongation complex with dC at +1 site, 8-oxo-GTP bound in A-site.
Authors	:	Hou, P.; Oh, J.; Wang, D.
Deposited on	:	2025-08-03
Resolution	:	3.48 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:

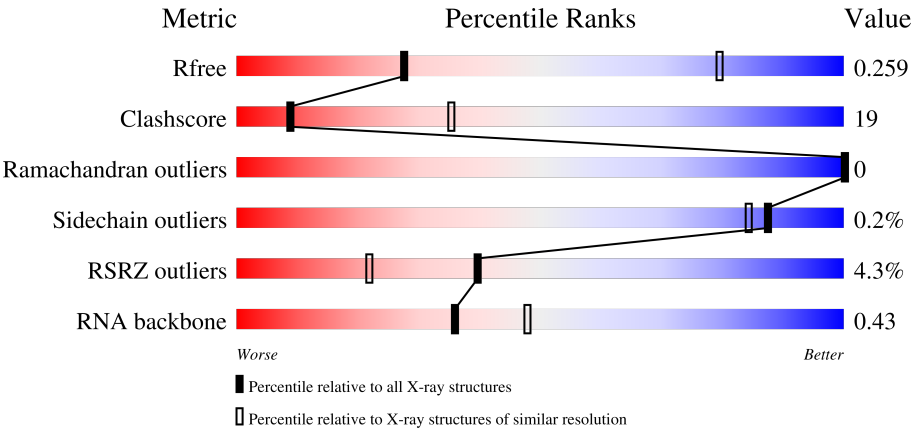
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 3.48 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}	180053	1083 (3.52-3.44)
Clashscore	190562	1139 (3.52-3.44)
Ramachandran outliers	187476	1111 (3.52-3.44)
Sidechain outliers	187428	1112 (3.52-3.44)
RSRZ outliers	180081	1082 (3.52-3.44)
RNA backbone	3983	1001 (3.94-3.02)

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	R	9	33% 56% 11%
2	T	29	48% 38% 14%
3	N	18	17% 61% 22%
4	A	1733	5% 47% 34% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	3% 56% 36% 8%
6	C	318	4% 57% 27% 16%
7	E	215	3% 61% 37%
8	F	155	% 39% 16% 45%
9	H	146	3% 47% 45% 9%
10	I	122	% 55% 41%
11	J	70	3% 51% 41% 7%
12	K	120	3% 64% 31% 5%
13	L	70	4% 46% 16% 39%

2 Entry composition (i)

There are 16 unique types of molecules in this entry. The entry contains 29102 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	9	Total	C	N	O	P	0	0	0
			193	88	40	57	8			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	25	Total	C	N	O	P	0	0	0
			497	239	76	157	25			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	14	Total	C	N	O	P	0	0	0
			296	138	66	78	14			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1393	Total	C	N	O	S	0	0	0
			10894	6875	1908	2051	60			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1123	Total	C	N	O	S	0	0	0
			8859	5607	1552	1647	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	212	Total	C	N	O	S	0	0	0
			1731	1100	305	315	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
			952	585	173	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.

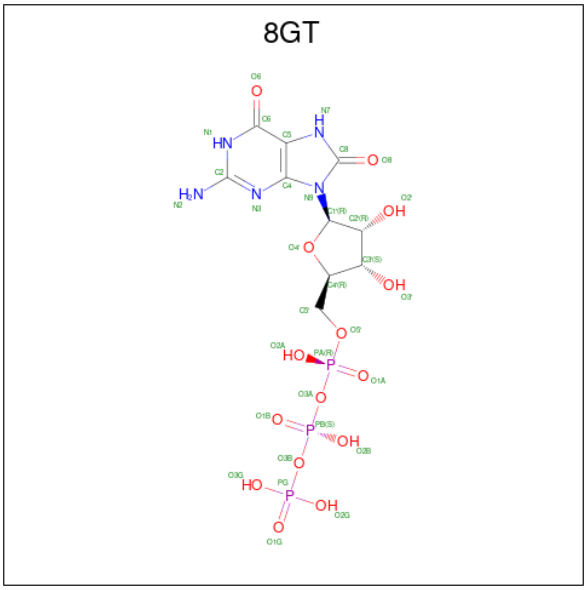
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- 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 8-OXO-GUANOSINE-5'-TRIPHOSPHATE (CCD ID: 8GT) (formula:

$C_{10}H_{16}N_5O_{15}P_3$) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 15 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

- Molecule 16 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

3 Residue-property plots (i)

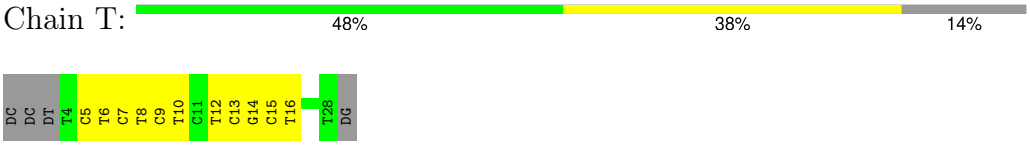
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

Chain R:

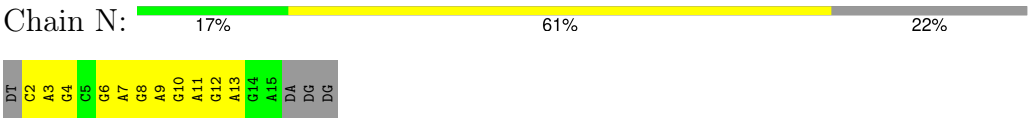
• Molecule 2: Template strand DNA

Chain T:



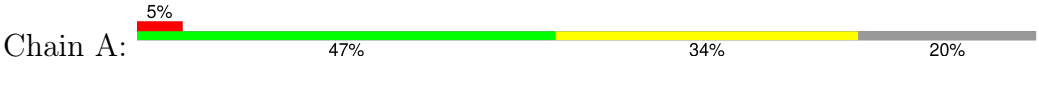
• Molecule 3: Non-template strand DNA

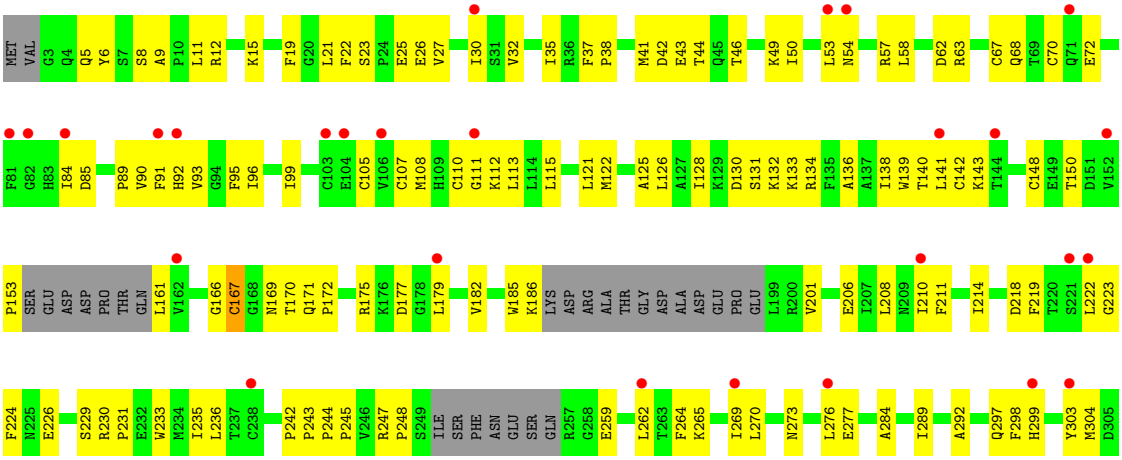
Chain N:

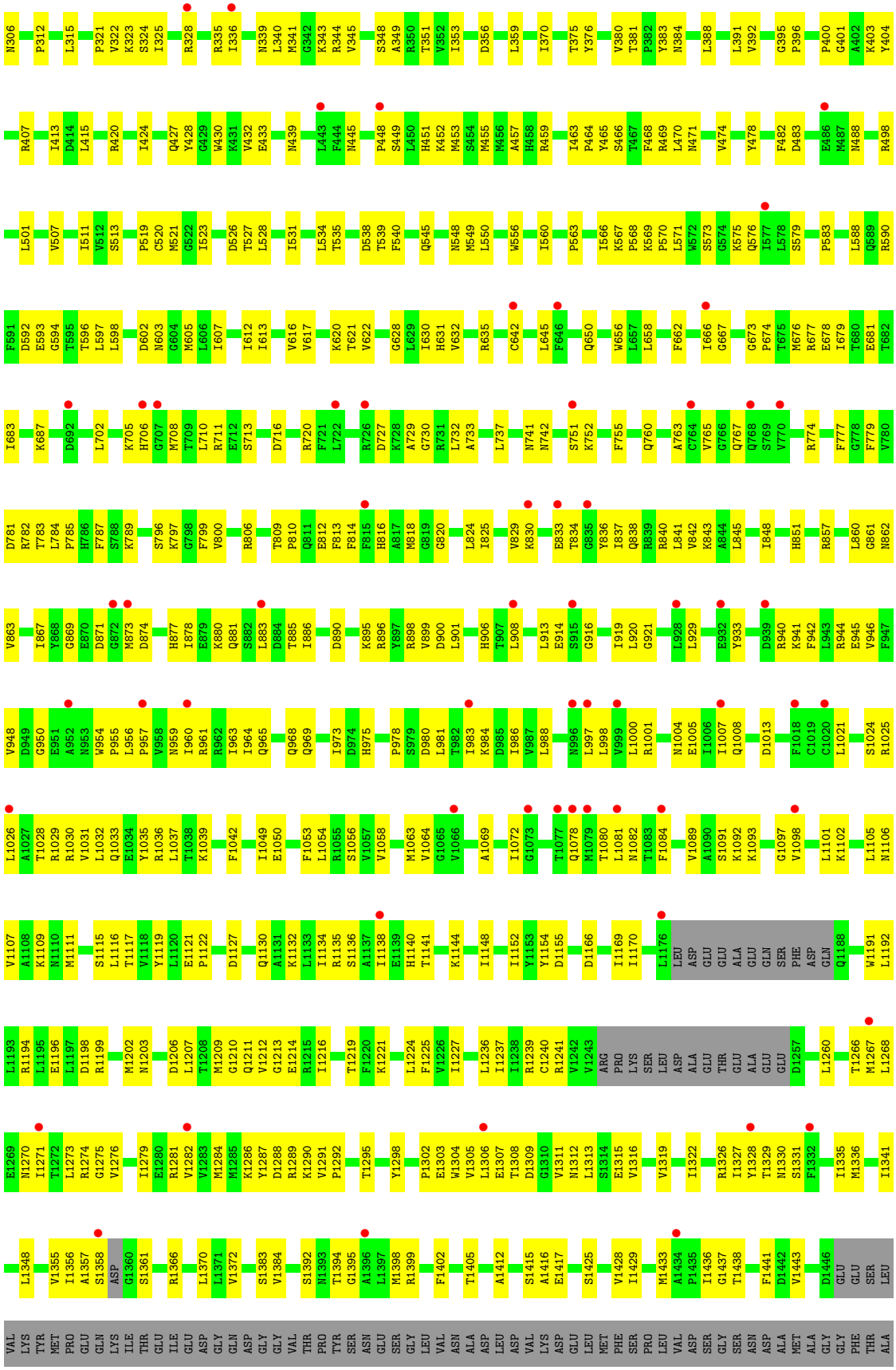


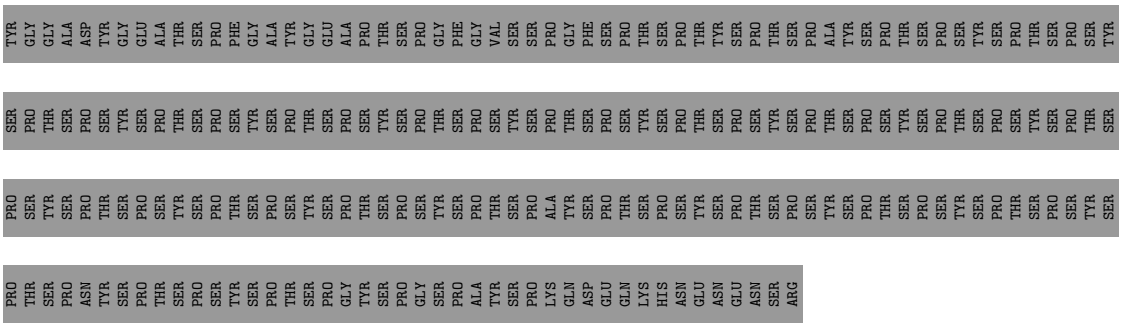
• Molecule 4: DNA-directed RNA polymerase II subunit RPB1

Chain A:



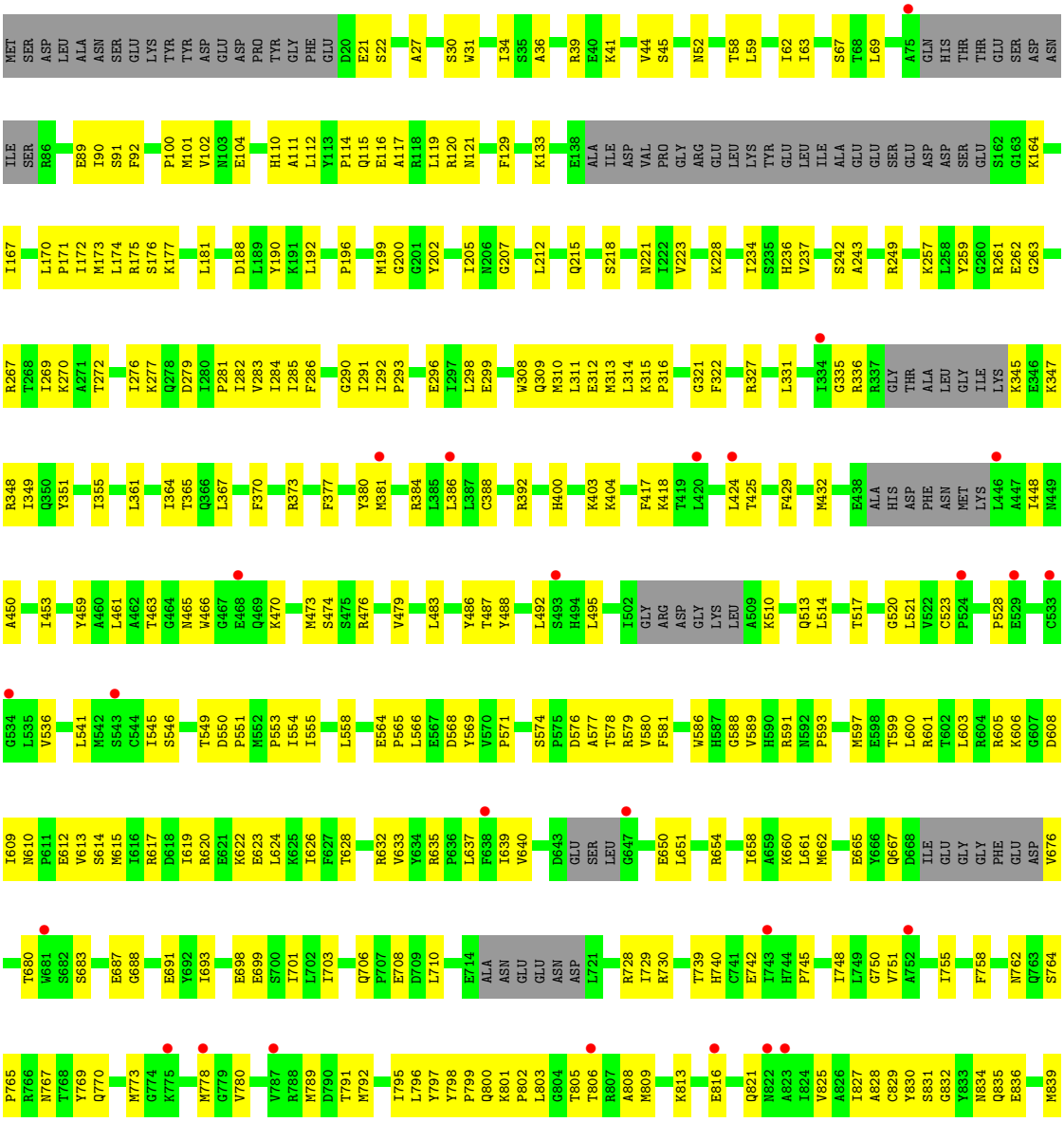


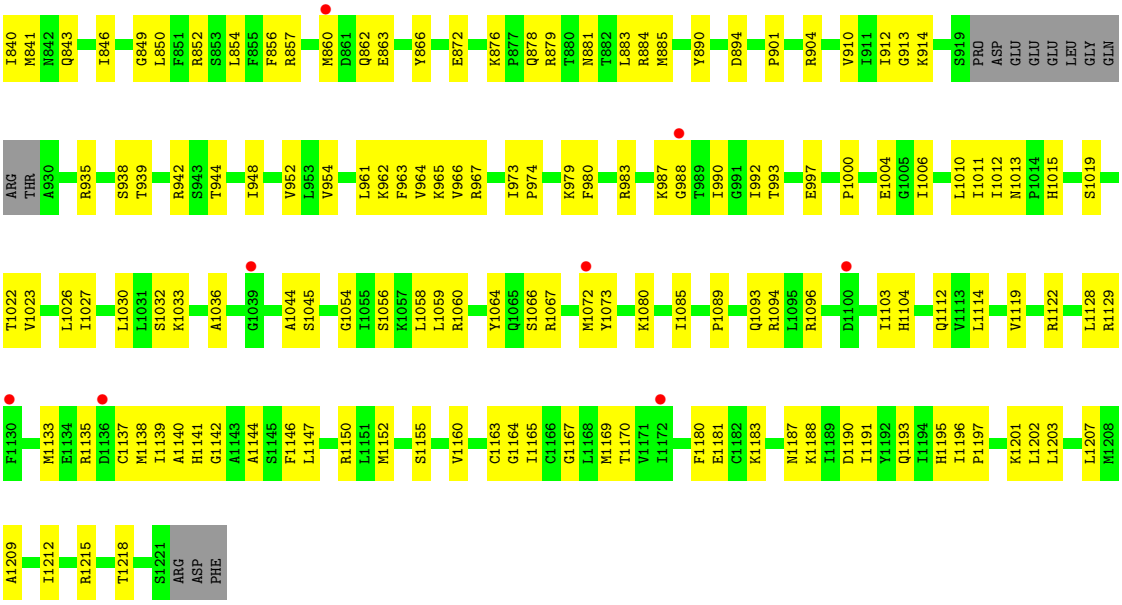

N306 P312 L315 P321 V322 K323 S324 I325 R328 R335 I336 N339 L340 M341 G342 K343 R344 V345 S348 A349 R350 T351 V352 I353 D356 L359 I370 T375 Y376 V380 T381 P382 Y383 N384 L388 L391 V392 G395 P396 P400 G401 A402 K403 Y404
R407 I413 D414 L415 R420 I424 Q427 Y428 G429 W430 K431 V432 E433 N439 L443 F444 N445 P448 S449 L450 H451 K452 M453 S454 M455 M456 A457 H458 R459 I463 P464 Y465 S466 T467 F468 R469 L470 N471 V474 Y478 F482 D483 E486 M487 N488 R498
L501 V507 I511 V512 S513 P519 C520 M521 G522 I523 D526 T527 L528 I531 L534 T535 D538 T539 F540 Q545 N548 M549 L550 W556 I560 P563 I566 K567 P568 K569 P570 L571 W572 S573 G574 K575 Q576 I577 L578 S579 P583 L588 Q589 R590
F591 D592 E593 G594 T595 T596 L597 L598 D602 N603 G604 M605 L606 I607 I612 I613 V616 V617 K620 T621 V622 G628 L629 I630 H631 V632 R635 C642 L645 F646 Q650 W656 L657 L658 F662 I666 G667 G673 P674 T675 M676 R677 E678 I679 T680 E681 T682
I683 K687 D692 L702 K705 H706 G707 M708 T709 L710 R711 E712 S713 D716 R720 F721 L722 R726 D727 K728 A729 G730 R731 L732 A733 L737 N741 N742 S751 K752 F755 Q760 A763 C764 V765 G766 Q767 Q768 S769 V770 R774 F777 G778 F779 V780
D781 R782 T783 L784 P785 H786 F787 S788 K789 S796 K797 G798 F799 V800 R806 T809 P810 Q811 E812 F813 F814 F815 H816 A817 M818 G819 G820 L824 I825 V829 K830 E833 T834 G835 Y836 I837 Q838 R839 R840 L841 V842 K843 A844 L845 I848 H851 R857 L860 G861 N862
V863 I867 Y868 G869 E870 D871 G872 M873 D874 H877 I878 E879 K880 Q881 S882 L883 D884 T885 I886 D890 K895 R896 Y897 R898 V899 D900 L901 H906 T907 L908 L913 E914 S915 G916 I919 L920 G921 L928 L929 E932 Y933 D939 R940 K941 F942 L943 R944 E945 V946 F947
V948 D949 G950 E951 A952 N953 W954 P955 L956 P957 V958 N959 I960 R961 R962 I963 I964 Q965 Q968 Q969 I973 D974 H975 P978 S979 D980 L981 T982 I983 K984 D985 I986 V987 L988 N996 L997 L998 V999 L1000 R1001 N1004 E1005 I1006 I1007 Q1008 D1013 F1018 C1019 C1020 L1021 S1024 R1025
L1026 A1027 T1028 R1029 R1030 V1031 L1032 Q1033 E1034 Y1035 R1036 L1037 T1038 K1039 F1042 I1049 E1050 F1053 L1054 R1055 S1056 V1057 V1058 M1063 V1064 G1065 V1066 A1069 I1072 G1073 T1077 Q1078 M1079 T1080 L1081 N1082 T1083 F1084 V1089 A1090 S1091 K1092 K1093 G1097 V1098 L1101 K1102 L1105 N1106
V1107 A1108 K1109 N1110 M1111 S1115 L1116 T1117 V1118 Y1119 L1120 E1121 P1122 D1127 Q1130 A1131 K1132 L1133 I1134 R1135 S1136 A1137 I1138 E1139 H1140 T1141 K1144 I1148 I1152 Y1153 Y1154 D1155 D1166 I1169 I1170 L1176 LEU ASP GLU GLU ALA GLU GLN SER PHE ASP GLN Q1188 W1191 L1192
L1193 R1194 L1195 E1196 L1197 D1198 R1199 M1202 N1203 D1206 L1207 T1208 M1209 G1210 Q1211 V1212 G1213 E1214 R1215 I1216 T1219 F1220 K1221 L1224 F1225 V1226 I1227 L1236 I1237 I1238 R1239 C1240 R1241 V1242 V1243 ARG PRO LYS SER LEU ASP ALA GLU THR GLU ALA GLU GLU D1257 L1260 T1266 M1267 L1268
E1269 N1270 I1271 T1272 L1273 R1274 G1275 V1276 I1279 E1280 R1281 V1282 V1283 M1284 M1285 K1286 Y1287 D1288 R1289 K1290 V1291 P1292 T1295 Y1298 P1302 E1303 W1304 V1305 L1306 E1307 T1308 D1309 G1310 V1311 N1312 L1313 S1314 E1315 V1316 V1319 I1322 R1326 I1327 Y1328 T1329 N1330 S1331 F1332 I1335 M1336 I1341
L1348 V1355 I1356 A1357 S1358 ASP G1360 S1361 R1366 L1370 L1371 V1372 S1383 V1384 S1392 N1393 T1394 G1395 A1396 L1397 M1398 R1399 F1402 T1405 A1412 S1415 A1416 E1417 S1425 V1428 I1429 M1433 A1434 P1435 I1436 G1437 T1438 F1441 D1442 V1443 D1446 GLU GLU SER LEU
VAL LYS TYR MET PRO GLU GLN LYS ILE THR GLU ILE GLU ASP GLY GLN ASP GLY GLY VAL THR PRO TYR SER ASN GLU SER GLY LEU VAL ASN ALA ASP LEU ASP VAL LYS ASP GLU LEU MET PHE SER PRO LEU VAL ASP SER GLY SER ASN ASP ALA MET ALA GLY GLY PHE THR ALA



• Molecule 5: DNA-directed RNA polymerase II subunit RPB2

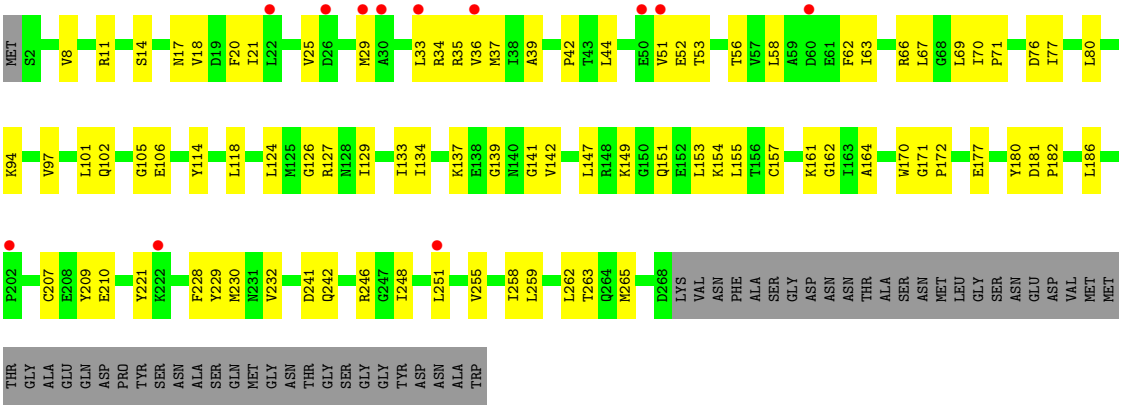
Chain B:





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

Chain C:

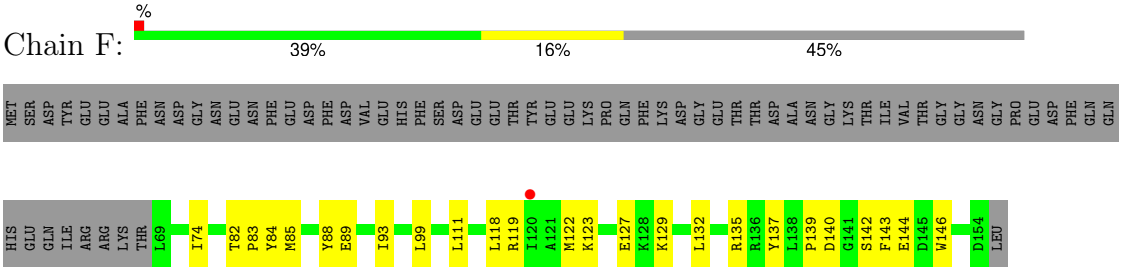


• Molecule 7: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:

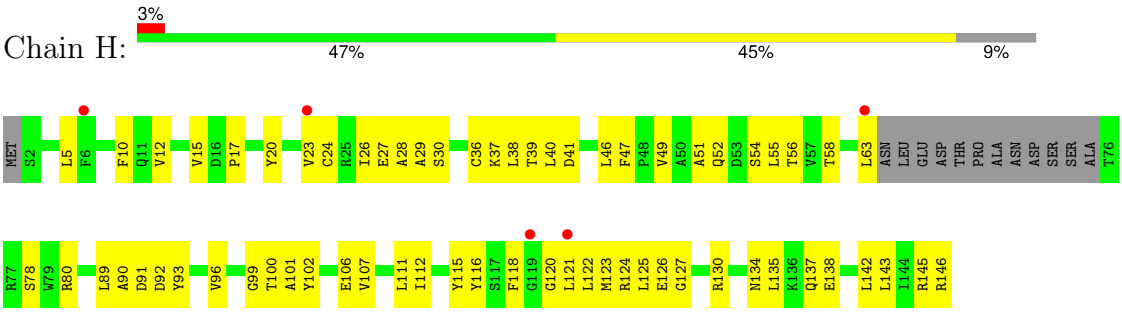
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F:



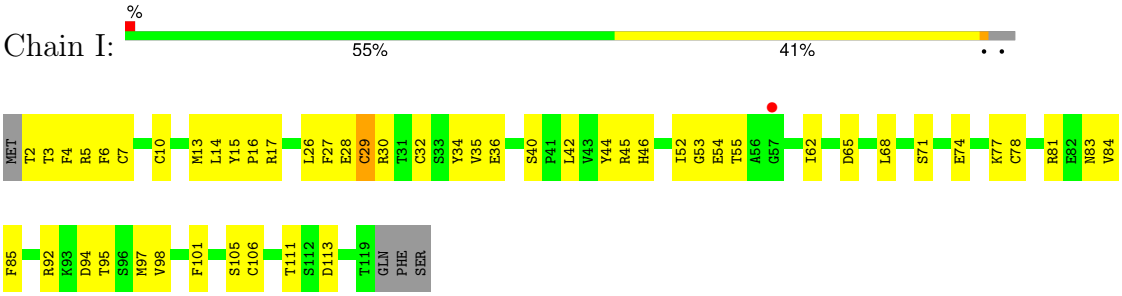
• Molecule 9: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:



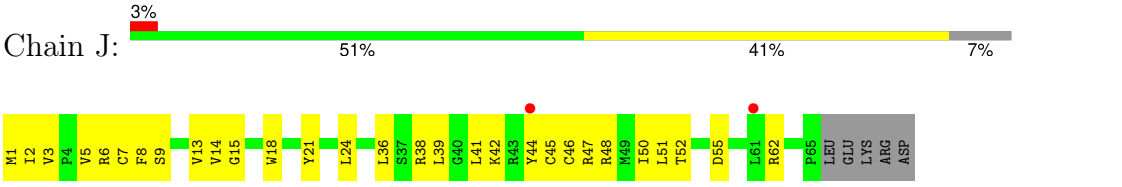
• Molecule 10: DNA-directed RNA polymerase II subunit RPB9

Chain I:



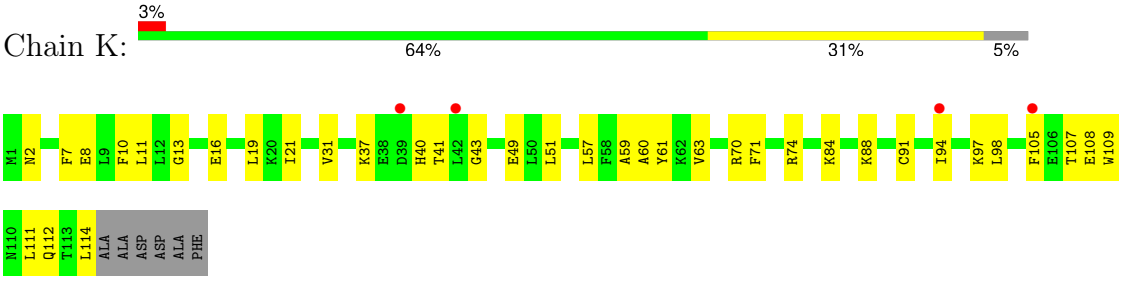
• Molecule 11: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:



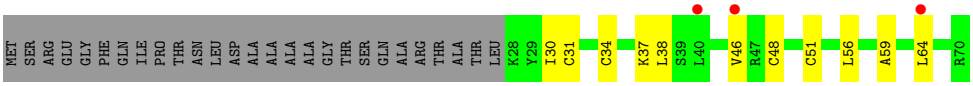
• Molecule 12: DNA-directed RNA polymerase II subunit RPB11

Chain K:



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

Chain L:


MET
SER
ARG
GLU
GLY
PHE
GLN
ILE
PRO
THR
ASN
LEU
ASP
ALA
ALA
ALA
ALA
GLY
THR
SER
GLN
ALA
ARG
THR
ALA
THR
LEU
K28
Y29
I30
C31
C34
K37
L38
S39
L40
V46
R47
C48
C51
L56
A59
L64
R70

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	159.52Å 222.70Å 191.19Å 90.00° 97.82° 90.00°	Depositor
Resolution (Å)	38.52 – 3.48 38.52 – 3.48	Depositor EDS
% Data completeness (in resolution range)	98.7 (38.52-3.48) 98.7 (38.52-3.48)	Depositor EDS
R_{merge}	0.30	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.78 (at 3.48Å)	Xtriage
Refinement program	PHENIX (1.21.2_5419: ???)	Depositor
R, R_{free}	0.218 , 0.258 0.218 , 0.259	Depositor DCC
R_{free} test set	2000 reflections (2.37%)	wwPDB-VP
Wilson B-factor (Å^2)	89.4	Xtriage
Anisotropy	0.842	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 116.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	29102	wwPDB-VP
Average B, all atoms (Å^2)	118.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.86% 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: ZN, MG, 8GT

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	0.16	0/217	0.32	0/337
2	T	0.22	0/551	0.44	0/845
3	N	0.21	0/335	0.36	0/516
4	A	0.18	0/11089	0.40	0/15001
5	B	0.17	0/9030	0.37	0/12186
6	C	0.16	0/2139	0.34	0/2899
7	E	0.17	0/1767	0.36	0/2378
8	F	0.16	0/696	0.33	0/943
9	H	0.18	0/1082	0.43	0/1466
10	I	0.16	0/970	0.41	0/1308
11	J	0.16	0/541	0.36	0/727
12	K	0.17	0/937	0.37	0/1265
13	L	0.21	0/339	0.43	0/450
All	All	0.17	0/29693	0.38	0/40321

There are no bond length outliers.

There are no bond angle outliers.

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	R	193	0	98	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	T	497	0	285	10	0
3	N	296	0	155	16	0
4	A	10894	0	10937	491	0
5	B	8859	0	8816	340	0
6	C	2101	0	2056	68	0
7	E	1731	0	1758	70	0
8	F	684	0	692	22	0
9	H	1064	0	1029	58	0
10	I	952	0	897	43	0
11	J	532	0	542	24	0
12	K	919	0	929	34	0
13	L	337	0	352	9	0
14	T	33	0	11	6	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	29102	0	28557	1075	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 19.

The worst 5 of 1075 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)
9:H:89:LEU:HD13	9:H:91:ASP:C	1.47	1.39
9:H:89:LEU:HD13	9:H:91:ASP:O	1.28	1.31
7:E:94:LYS:HG2	7:E:123:LEU:HD13	1.30	1.08
9:H:89:LEU:CD1	9:H:91:ASP:C	2.36	0.98
4:A:705:LYS:HE3	4:A:708:MET:SD	2.06	0.95

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	
4	A	1379/1733 (80%)	1344 (98%)	35 (2%)	0	100	100
5	B	1103/1224 (90%)	1086 (98%)	17 (2%)	0	100	100
6	C	265/318 (83%)	264 (100%)	1 (0%)	0	100	100
7	E	210/215 (98%)	207 (99%)	3 (1%)	0	100	100
8	F	84/155 (54%)	82 (98%)	2 (2%)	0	100	100
9	H	129/146 (88%)	127 (98%)	2 (2%)	0	100	100
10	I	116/122 (95%)	115 (99%)	1 (1%)	0	100	100
11	J	63/70 (90%)	63 (100%)	0	0	100	100
12	K	112/120 (93%)	111 (99%)	1 (1%)	0	100	100
13	L	41/70 (59%)	40 (98%)	1 (2%)	0	100	100
All	All	3502/4173 (84%)	3439 (98%)	63 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	
4	A	1200/1520 (79%)	1198 (100%)	2 (0%)	87	84
5	B	955/1061 (90%)	954 (100%)	1 (0%)	88	87
6	C	235/274 (86%)	235 (100%)	0	100	100
7	E	193/197 (98%)	193 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
8	F	73/137 (53%)	73 (100%)	0	100 100
9	H	116/128 (91%)	116 (100%)	0	100 100
10	I	110/116 (95%)	107 (97%)	3 (3%)	39 62
11	J	60/65 (92%)	59 (98%)	1 (2%)	53 69
12	K	99/102 (97%)	99 (100%)	0	100 100
13	L	37/57 (65%)	37 (100%)	0	100 100
All	All	3078/3657 (84%)	3071 (100%)	7 (0%)	87 84

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	I	7	CYS
10	I	10	CYS
11	J	7	CYS
10	I	29	CYS
5	B	1163	CYS

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	1193	GLN
5	B	1211	ASN
6	C	267	GLN
4	A	1378	GLN
4	A	1110	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	7/9 (77%)	1 (14%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	5	A

There are no RNA pucker outliers to report.

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 oligosaccharides 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	8GT	T	101	16	34,35,35	4.96	21 (61%)	48,56,56	1.96	10 (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
14	8GT	T	101	16	-	8/22/38/38	0/3/3/3

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	T	101	8GT	C4-N3	10.02	1.57	1.34
14	T	101	8GT	C6-N1	9.55	1.56	1.38
14	T	101	8GT	C3'-C4'	-9.39	1.29	1.53
14	T	101	8GT	C2'-C1'	-9.08	1.24	1.53
14	T	101	8GT	O4'-C1'	8.69	1.62	1.42

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
14	T	101	8GT	O6-C6-C5	-5.63	113.68	127.26
14	T	101	8GT	N9-C4-N3	5.49	132.99	126.13
14	T	101	8GT	C5-C6-N1	4.71	125.08	112.13
14	T	101	8GT	C2-N3-C4	4.20	119.54	112.30
14	T	101	8GT	C2'-C1'-N9	-3.76	110.86	115.90

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
14	T	101	8GT	C5'-O5'-PA-O3A
14	T	101	8GT	C5'-O5'-PA-O2A
14	T	101	8GT	C4'-C5'-O5'-PA
14	T	101	8GT	O4'-C1'-N9-C8
14	T	101	8GT	O4'-C1'-N9-C4

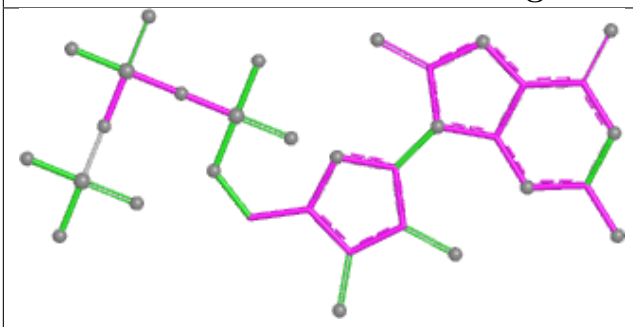
There are no ring outliers.

1 monomer is involved in 6 short contacts:

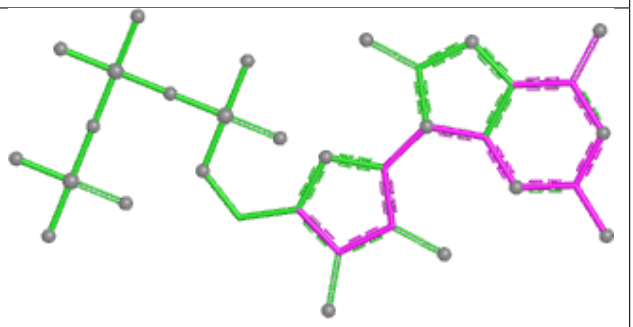
Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	T	101	8GT	6	0

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

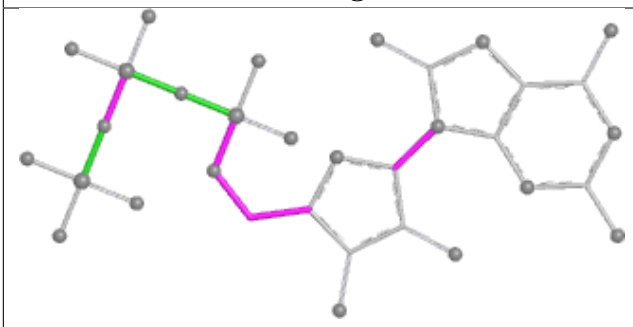
Ligand 8GT T 101



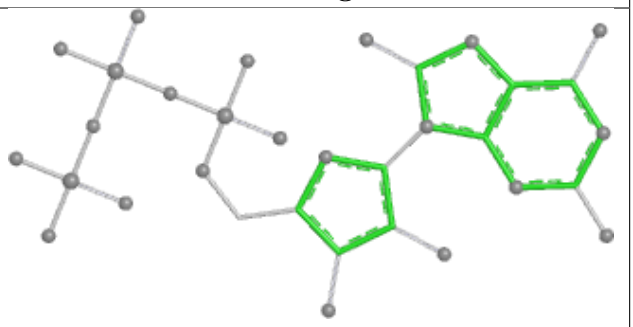
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 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	R	9/9 (100%)	-0.27	0 100 100	78, 84, 138, 140	0
2	T	25/29 (86%)	0.31	0 100 100	78, 121, 257, 265	0
3	N	14/18 (77%)	0.48	0 100 100	144, 231, 286, 286	0
4	A	1393/1733 (80%)	0.45	87 (6%) 26 17	52, 117, 202, 273	0
5	B	1123/1224 (91%)	0.37	34 (3%) 52 31	50, 96, 167, 297	0
6	C	267/318 (83%)	0.42	12 (4%) 38 21	45, 87, 136, 218	0
7	E	212/215 (98%)	0.36	7 (3%) 49 28	98, 149, 214, 243	0
8	F	86/155 (55%)	0.28	1 (1%) 76 53	84, 117, 166, 205	0
9	H	133/146 (91%)	0.37	5 (3%) 44 25	87, 129, 178, 267	0
10	I	118/122 (96%)	0.40	1 (0%) 82 61	80, 142, 199, 225	0
11	J	65/70 (92%)	0.52	2 (3%) 51 30	56, 78, 130, 153	0
12	K	114/120 (95%)	0.28	4 (3%) 47 27	58, 92, 134, 181	0
13	L	43/70 (61%)	0.67	3 (6%) 22 15	81, 146, 219, 305	0
All	All	3602/4229 (85%)	0.41	156 (4%) 40 22	45, 109, 193, 305	0

The worst 5 of 156 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	B	1136	ASP	5.6
7	E	97	VAL	5.1
4	A	1282	VAL	5.1
7	E	93	MET	4.5
4	A	1098	VAL	4.3

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates (i)

There are no oligosaccharides in this entry.

6.4 Ligands (i)

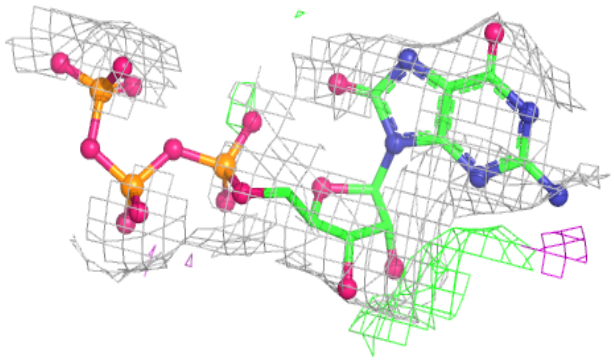
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

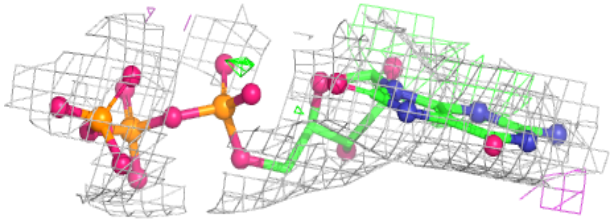
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
16	MG	A	1803	1/1	0.72	0.15	123,123,123,123	0
15	ZN	L	101	1/1	0.88	0.09	297,297,297,297	0
15	ZN	A	1801	1/1	0.88	0.08	272,272,272,272	0
14	8GT	T	101	33/33	0.93	0.09	46,68,125,132	0
15	ZN	I	202	1/1	0.93	0.09	245,245,245,245	0
16	MG	A	1804	1/1	0.97	0.05	103,103,103,103	0
15	ZN	A	1802	1/1	0.98	0.03	169,169,169,169	0
15	ZN	I	201	1/1	0.98	0.15	204,204,204,204	0
15	ZN	B	1301	1/1	0.99	0.03	191,191,191,191	0
15	ZN	C	401	1/1	0.99	0.14	154,154,154,154	0
15	ZN	J	101	1/1	0.99	0.10	104,104,104,104	0

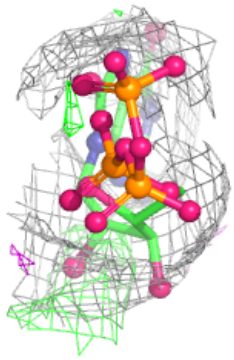
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around 8GT 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.