



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 06:37 am BST

PDB ID : 1Y77
Title : Complete RNA Polymerase II elongation complex with substrate analogue GMPCPP
Authors : Kettenberger, H.; Armache, K.-J.; Cramer, P.
Deposited on : 2004-12-08
Resolution : 4.50 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 : 2.11
buster-report : 1.1.7 (2018)
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

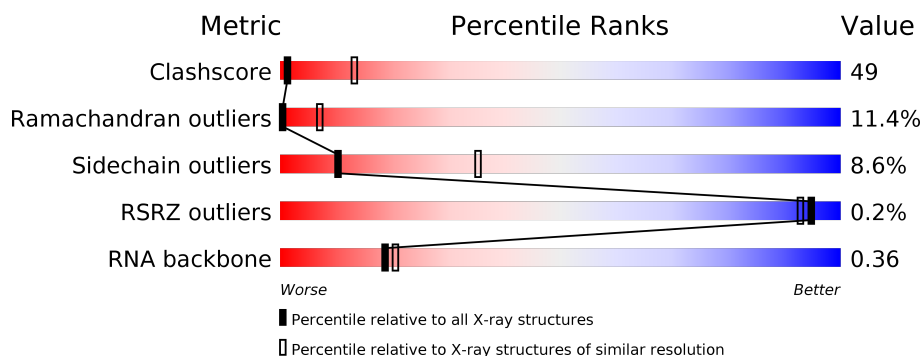
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.50 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.



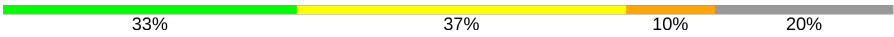
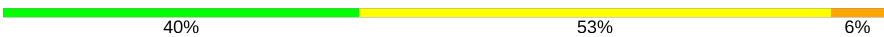
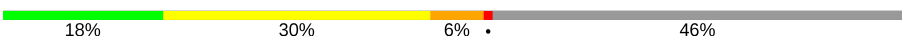


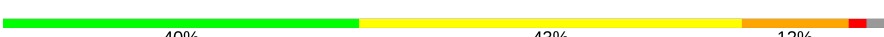
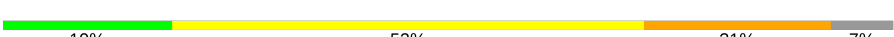
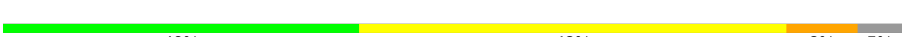

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1123 (5.20-3.80)
Ramachandran outliers	138981	1069 (5.20-3.80)
Sidechain outliers	138945	1050 (5.20-3.80)
RSRZ outliers	127900	1101 (5.30-3.70)
RNA backbone	3102	1063 (6.00-3.00)

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 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	T	19	<div> <div>16%</div> <div>21% 58% 21%</div> </div>
2	N	7	<div> <div>14%</div> <div>86%</div> </div>
3	P	10	<div> <div>80% 20%</div> </div>
4	A	1733	<div> <div>27% 43% 10% 18%</div> </div>
5	B	1224	<div> <div>29% 48% 13% 9%</div> </div>
6	C	318	<div> <div>23% 47% 12% 16%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	D	221	
8	E	215	
9	F	155	
10	G	171	
11	H	146	
12	I	122	
13	J	70	
14	K	120	
15	L	70	

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	B	1308	X	-	-	-

2 Entry composition

There are 18 unique types of molecules in this entry. The entry contains 31835 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 DNA chain called 5'-D(P*AP*GP*TP*AP*CP*TP*TP*AP*CP*T*CP*GP*CP*CP*TP*GP*GP*TP*CP*TP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	19	Total	C	N	O	P	21	0	0
			388	185	67	117	19			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	7	Total	C	N	O	P	20	0	0
			141	69	27	39	6			

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1416	Total	C	N	O	S	0	0	0
			11140	7021	1946	2111	62			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1112	Total	C	N	O	S	0	0	0
			8836	5594	1548	1639	55			

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

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II 32 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	D	177	Total	C	N	O	S	0	0	0
			1356	840	241	273	2			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

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

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

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

- Molecule 10 is a protein called DNA-directed RNA polymerase II 19 kDa polypeptide.

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

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

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

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit 9.

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

- Molecule 13 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

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

- Molecule 14 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

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

- Molecule 15 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

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

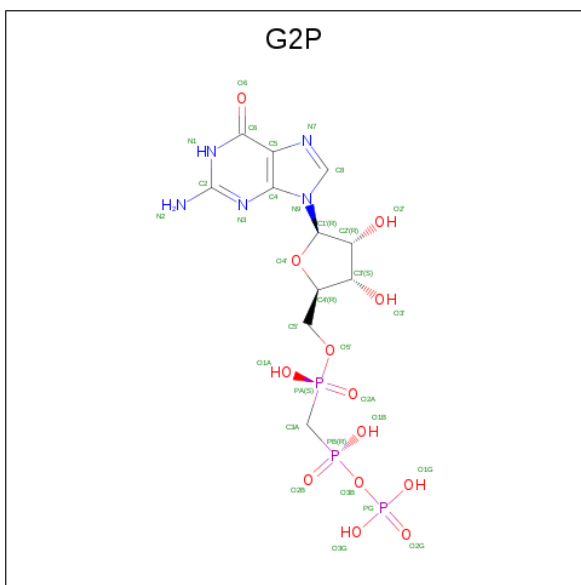
- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	J	1	Total	Zn	0	0
			1	1		
16	B	1	Total	Zn	0	0
			1	1		
16	I	2	Total	Zn	0	0
			2	2		
16	C	1	Total	Zn	0	0
			1	1		
16	A	2	Total	Zn	0	0
			2	2		
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: C₁₁H₁₈N₅O₁₃P₃).

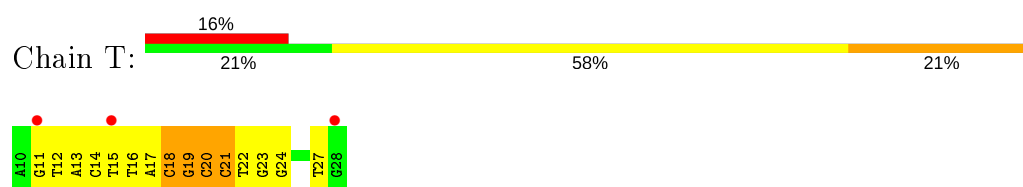


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
18	B	1	Total	C	N	O	P	32	0
			32	11	5	13	3		

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

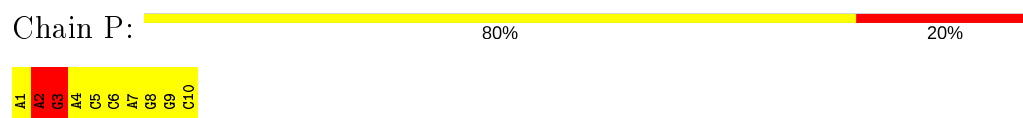
- Molecule 1: 5'-D(P*AP*GP*TP*AP*CP*TP*TP*AP*CP*T*CP*GP*CP*CP*TP*GP*GP*T P*CP*TP*G)-3'



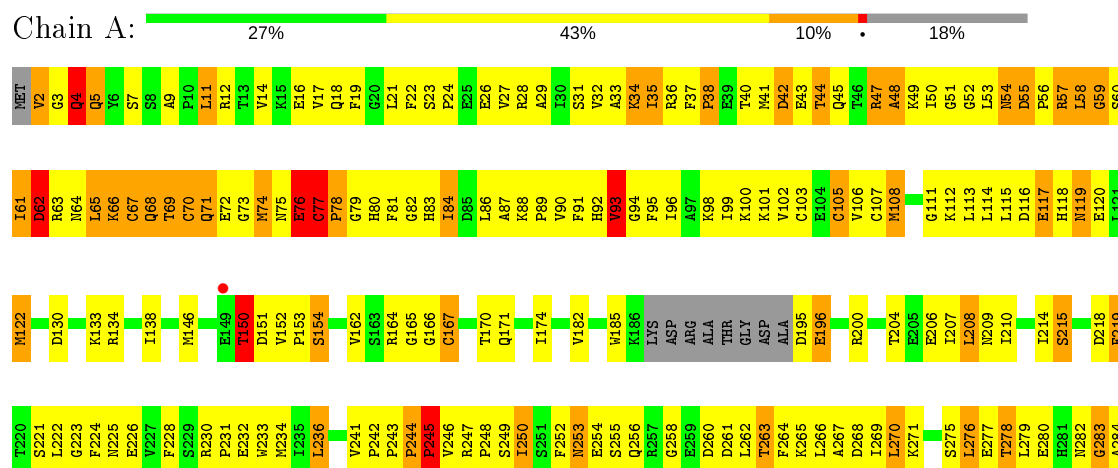
- Molecule 2: 5'-D(*AP*AP*GP*TP*AP*CP*T)-3'



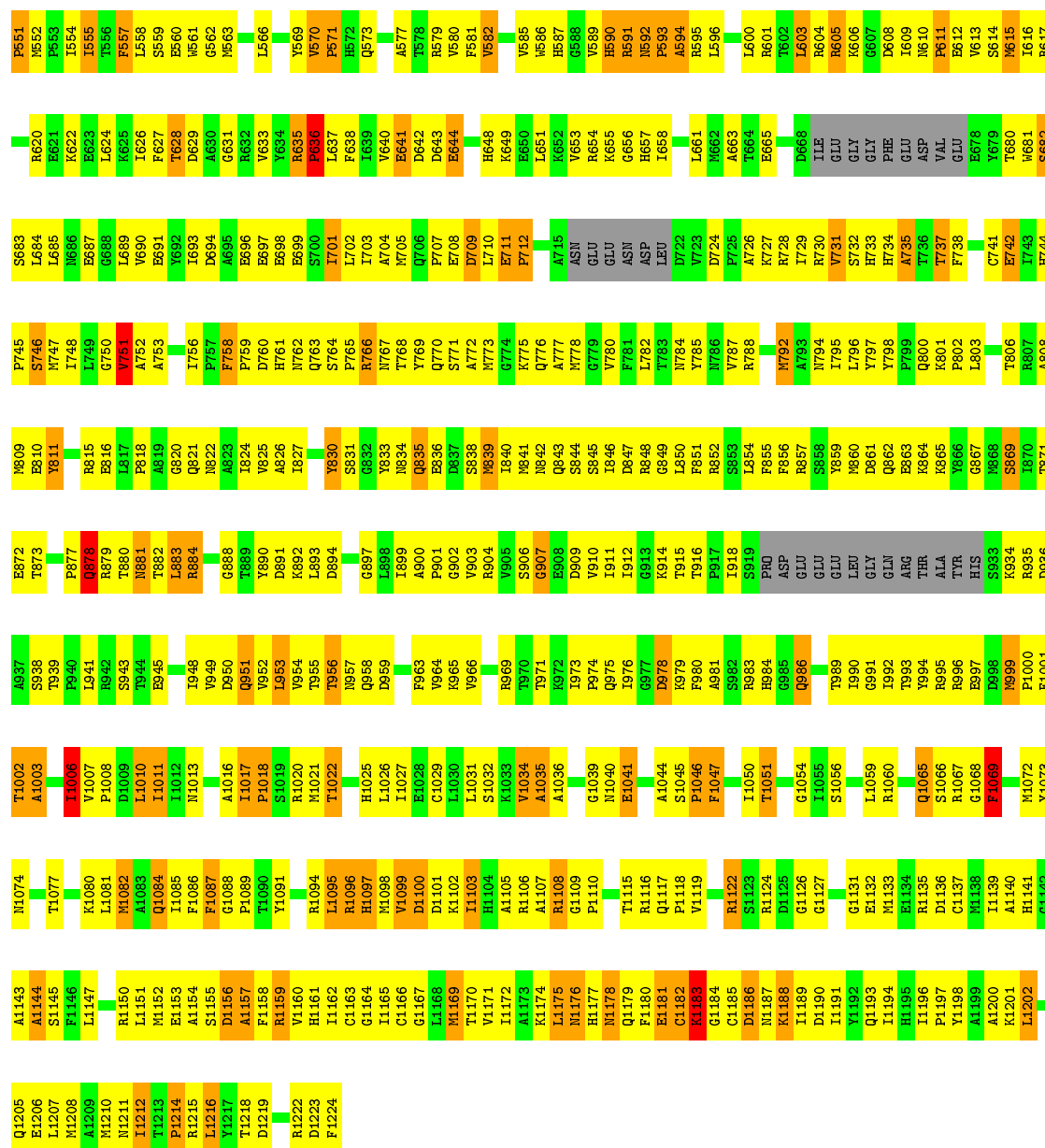
- Molecule 3: 5'-R(*AP*AP*GP*AP*CP*CP*AP*GP*GP*C)-3'



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

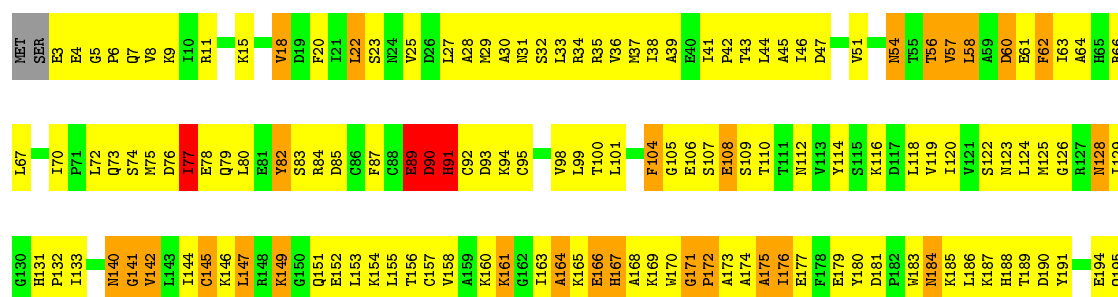


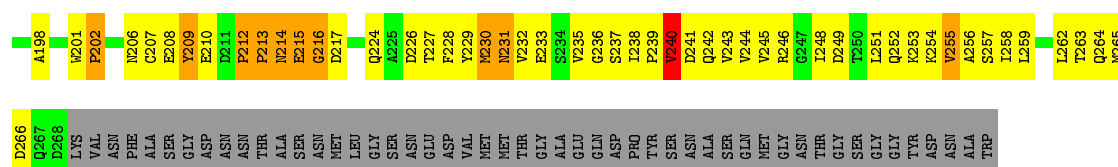
L1268	L1197	A1131	P1060	L998	E914	E846	R782	L701	G623	P554	V491	Y428	R350	P285
I1271	D1198	K1132	M1063	V999	I919	D847	T783	T709	S624	D555	P492	G429	T351	H286
R1274	M1202	I1133	V1064	L1000	I920	I848	L784	L710	S625	W556	Q493	K430	V352	E287
I1279	K1205	R1135	L1067	G1002	G921	Y852	H786	R711	G627	G558	E495	V432	I288	A288
E1280	D1206	S1136	N1004	K1003	D922	D853	K789	E712	G628	V559	E496	E433	I289	S354
R1281	L1207	A1137	N1005	G1004	Q926	H854		S713	L629	I560	T497	R434	E290	E291
V1282	T1208	T1141	E1006	E1005	I929	T855	F794	F714	L630	P661	R498	H435	A292	A292
V1283	M1209	T1147	Q1008	I1007	L929	R857	E795	V718	H631	T562	A499	H435	E293	E293
M1284	G1210	T1147	P1075	E1006	L929	R857	E795	V719	V632	P663	A499	H435	S294	S294
R1289	G1213	I1148	A1076	N1009	E932	S859	K797	R720	Q503	K567	S502	D438	L295	L295
K1290	E1214	S1150	M1079	A1010	E933	L860	K798	R721	Q503	P668	S502	D440	L296	L296
P1292	R1215	I1152	T1080	R1013	Q935	H862	R800	L722	F646	K569	Q505	P441	G365	G365
S1293	I1216	Y1153	ASN	A1014	V937	I864	E802	E724	G647	P570	L504	V442	V366	Q297
P1294	K1217	Y1154	THR	V1015	V937	I864	N802	A725	L649	W572	A506	L444	P367	F298
T1295	D1218	D1155	PHE	T1016	D939	F866	Y804	K728	K651	G574	P508	R446	K372	T302
G1296	T1219	P1156	HIS	L1017	R940	I867	L805	A729	V652	K575	Q510	Q447	T373	Y303
E1297	K1221	P1157	PHE	F1018	R941	F868	R806	R730	V653	Q576	Q510	L450	T381	D305
Y1298	N1222	P1158	ALA	C1019	F942	G869	G807	R731	I654	I577	V512	K452	P382	Y383
V1299	D1223	S1160	GLY	C1020	L943	E870	L808	L732	F655	L578	V512	K452	P382	N306
K1300	L1224	T1161	VAL	L1021	L943	E870	L808	L732	F655	L578	V512	K452	P382	N306
E1303	V1225	V1182	ALA	L1022	F947	G872	P810	N736	L658	V580	Q515	S454	K372	T302
W1304	V1226	I1163	SER	R1023	V946	H873	Q811	L737	H659	V580	Q515	S454	T373	Y303
V1305	W1228	P1164	K1092	A1027	D949	D874	E812	K738	N660	I586	N517	M455	T381	D305
L1306	W1228	E1165	K1093	T1028	W954	A875	F813	D739	G661	H587	K518	M456	P382	Y383
E1307	D1233	D1166	T1095	R1028	W954	A876	F814	L740	F662	L588	P519	A457	P382	N306
T1308	L1236	E1167	S1096	R1029	P957	H877	F815	N741	S663	Q589	M521	R459	K317	K317
D1309	I1237	I1168	G1097	R1030	V958	A817	R822	N742	G665	F591	I593	V460	Y392	Y392
N1312	I1238	I1169	V1098	V1031	R959	Q881	R822	N742	G665	F591	I593	V460	Y392	Y392
L1313	R1239	Q1171	P1099	L1032	R960	S882	G819	Q745	G667	Q525	Q525	P464	G401	R320
S1314	G1240	L1172	R1100	Q1033	R961	L883	G820	N746	D668	D526	D526	Y465	A402	R321
E1315	R1241	H1173	L1105	E1035	R962	D884	R821	V747	T596	T596	T527	S466	K403	V322
V1316	V1242	F1174	L1105	R1036	R964	I886	R822	N747	L670	I597	L528	T467	Y404	K323
V1319	ARG	LEU	ASP	L1037	T970	G837	G823	W751	A671	I597	L528	F468	Y404	K323
P1320	LYS	GLU	ASP	L1037	T970	G837	G823	W751	A671	I597	L528	F468	Y404	K323
I1322	SER	GLU	GLU	L1037	T970	G837	G823	W751	A671	I597	L528	F468	Y404	K323
P1324	ASP	GLU	GLU	L1037	T970	G837	G823	W751	A671	I597	L528	F468	Y404	K323
R1326	GLU	GLU	GLU	L1037	T970	G837	G823	W751	A671	I597	L528	F468	Y404	K323
I1327	THR	PHE	PHE	L1037	T970	G837	G823	W751	A671	I597	L528	F468	Y404	K323
Y1328	GLU	ASP	ASP	L1037	T970	G837	G823	W751	A671	I597	L528	F468	Y404	K323
T1329	A1254	Q1187	Q1187	N1048	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
N1330	E1255	Q1188	Q1188	E1050	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
S1331	E1256	S1189	S1189	A1051	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
F1332	P1190	P1190	P1190	Q1062	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
I1333	G1123	W1191	G1123	L1054	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
D1334	K1261	L1192	H1124	R1055	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
I1336	E1264	L1193	L1193	S1056	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
M1336	N1265	R1194	R1194	V1057	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
E1337	L1266	L1195	L1195	W1058	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323
V1338	M1267	E1196	Q1130	H1059	R977	R896	R823	W751	A671	I597	L528	F468	Y404	K323



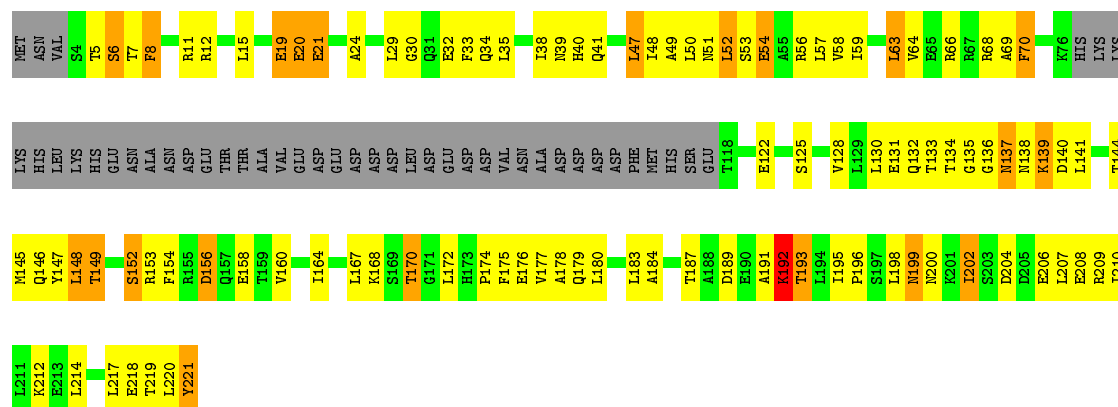
• Molecule 6: DNA-directed RNA polymerase II 45 kDa polypeptide

Chain C: 23% 47% 12% 16%

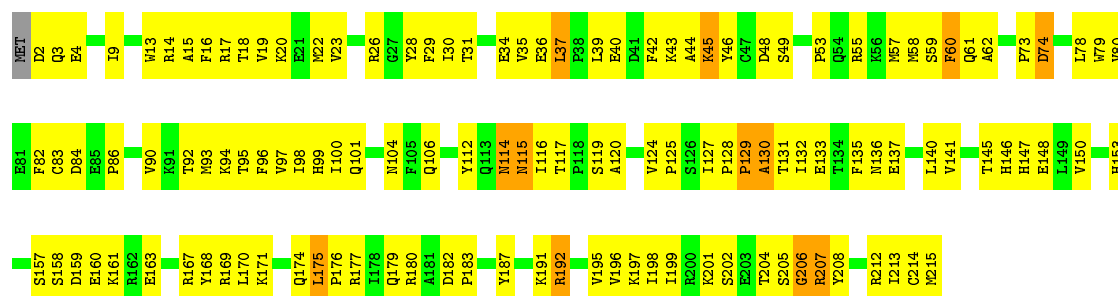




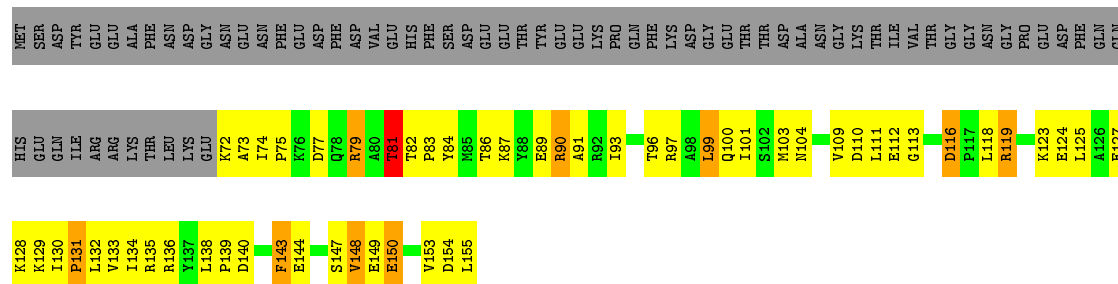
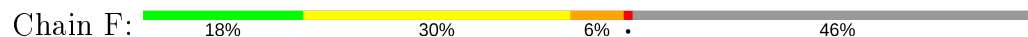
• Molecule 7: DNA-directed RNA polymerase II 32 kDa polypeptide



• Molecule 8: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

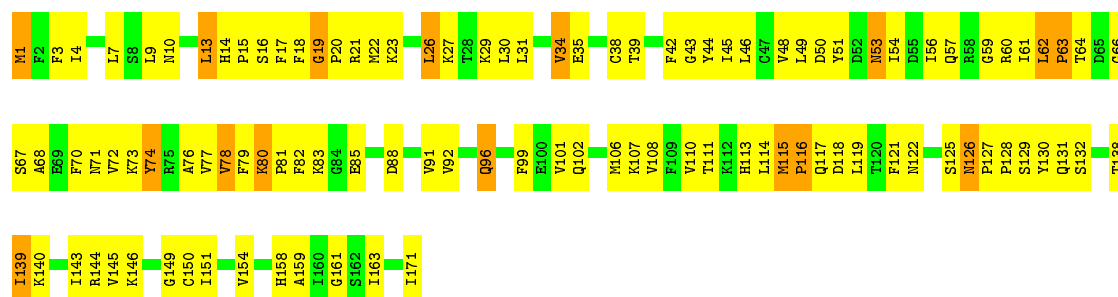


• Molecule 9: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide



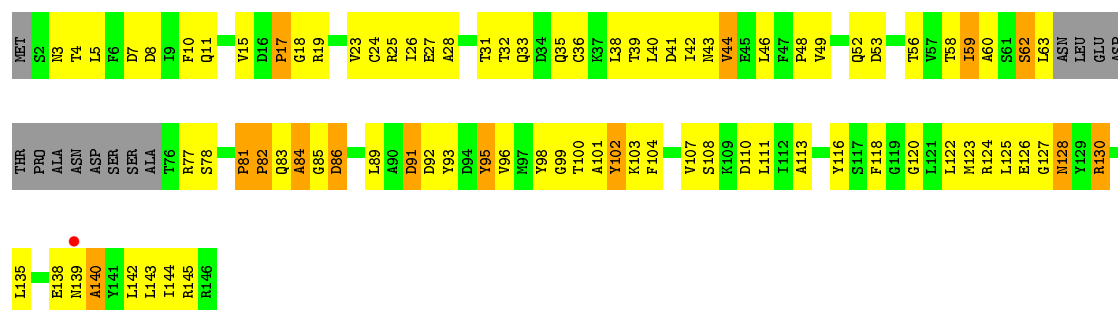
• Molecule 10: DNA-directed RNA polymerase II 19 kDa polypeptide

Chain G: 



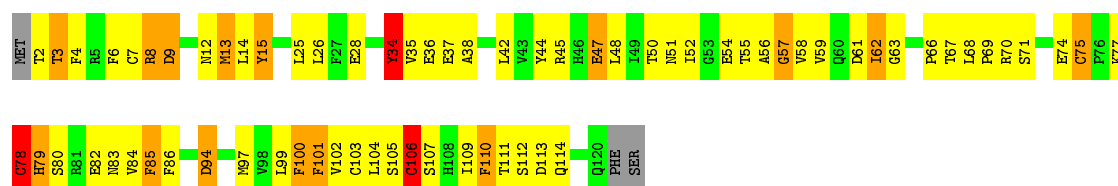
- Molecule 11: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

Chain H: 




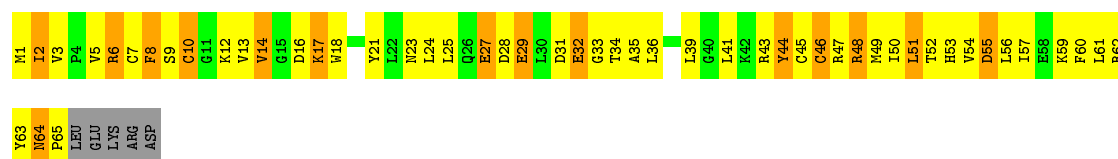
- Molecule 12: DNA-directed RNA polymerase II subunit 9

Chain I: 



- Molecule 13: DNA-directed RNA polymerases I/II/III subunit 10

Chain J: 



- Molecule 14: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K: 



- Molecule 15: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	222.46Å 393.07Å 283.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 4.50 48.95 – 4.49	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-4.50) 78.3 (48.95-4.49)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.98 (at 4.45Å)	Xtriage
Refinement program	unknown	Depositor
R, R_{free}	0.297 , (Not available) 0.270 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	137.1	Xtriage
Anisotropy	0.467	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 54.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.019 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.030 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	31835	wwPDB-VP
Average B, all atoms (Å ²)	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.97% 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: 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	T	1.18	0/433	1.16	2/666 (0.3%)
2	N	1.35	0/158	0.85	0/242
3	P	1.14	0/240	2.32	6/373 (1.6%)
4	A	0.59	6/11339 (0.1%)	0.78	16/15334 (0.1%)
5	B	0.57	5/9008 (0.1%)	0.86	16/12146 (0.1%)
6	C	0.63	2/2133 (0.1%)	1.10	5/2891 (0.2%)
7	D	0.49	0/1365	0.73	2/1837 (0.1%)
8	E	0.44	0/1788	0.64	0/2406
9	F	0.54	0/691	0.78	0/933
10	G	0.54	0/1368	0.75	0/1844
11	H	0.40	0/1086	0.66	0/1470
12	I	0.44	0/989	0.75	0/1331
13	J	0.54	0/541	0.82	1/727 (0.1%)
14	K	0.48	0/937	0.68	0/1265
15	L	0.46	0/365	0.75	0/485
All	All	0.58	13/32441 (0.0%)	0.84	48/43950 (0.1%)

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	T	0	3
4	A	0	3
5	B	1	7
6	C	0	4
All	All	1	17

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	119	ASN	C-N	27.87	1.98	1.34
4	A	150	THR	C-N	-17.84	0.93	1.34
5	B	503	GLY	C-N	14.88	1.68	1.34
6	C	89	GLU	C-N	-14.40	1.00	1.34
5	B	363	HIS	C-N	-13.51	1.02	1.34

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	503	GLY	O-C-N	-24.37	83.71	122.70
6	C	90	ASP	CB-CG-OD1	-24.17	96.54	118.30
5	B	469	GLN	O-C-N	-23.96	84.36	122.70
6	C	89	GLU	O-C-N	-23.94	84.40	122.70
3	P	3	G	O5'-P-OP2	23.57	138.99	110.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	B	364	ILE	CB

5 of 17 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	150	THR	Mainchain
4	A	77	CYS	Peptide
1	T	19	DG	Sidechain
1	T	20	DC	Sidechain
1	T	21	DC	Sidechain

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	T	388	0	216	51	0
2	N	141	0	81	8	0
3	P	214	0	111	20	0
4	A	11140	0	11218	1205	0
5	B	8836	0	8869	984	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	C	2095	0	2049	244	0
7	D	1356	0	1319	118	0
8	E	1752	0	1776	152	0
9	F	679	0	701	83	0
10	G	1340	0	1357	154	0
11	H	1068	0	1040	104	0
12	I	971	0	928	99	0
13	J	532	0	542	94	0
14	K	919	0	929	81	0
15	L	363	0	387	43	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
16	L	1	0	0	0	0
17	A	1	0	0	0	0
18	B	32	0	12	0	0
All	All	31835	0	31535	3108	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 49.

The worst 5 of 3108 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:825:ILE:CG2	5:B:508:LEU:HD11	1.45	1.45
5:B:503:GLY:C	5:B:504:ARG:N	1.68	1.44
1:T:16:DT:C4'	4:A:1403:GLU:OE2	1.77	1.30
1:T:20:DC:H4'	4:A:447:GLN:NE2	1.44	1.29
4:A:56:PRO:O	4:A:57:ARG:HG3	1.28	1.28

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	A	1406/1733 (81%)	962 (68%)	280 (20%)	164 (12%)	0	6
5	B	1096/1224 (90%)	740 (68%)	225 (20%)	131 (12%)	0	6
6	C	264/318 (83%)	158 (60%)	67 (25%)	39 (15%)	0	4
7	D	173/221 (78%)	121 (70%)	34 (20%)	18 (10%)	0	9
8	E	212/215 (99%)	149 (70%)	48 (23%)	15 (7%)	1	16
9	F	82/155 (53%)	64 (78%)	14 (17%)	4 (5%)	2	23
10	G	169/171 (99%)	131 (78%)	26 (15%)	12 (7%)	1	16
11	H	129/146 (88%)	84 (65%)	30 (23%)	15 (12%)	0	6
12	I	117/122 (96%)	81 (69%)	27 (23%)	9 (8%)	1	15
13	J	63/70 (90%)	37 (59%)	11 (18%)	15 (24%)	0	1
14	K	112/120 (93%)	89 (80%)	18 (16%)	5 (4%)	2	24
15	L	44/70 (63%)	18 (41%)	13 (30%)	13 (30%)	0	0
All	All	3867/4565 (85%)	2634 (68%)	793 (20%)	440 (11%)	0	7

5 of 440 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	4	GLN
4	A	5	GLN
4	A	48	ALA
4	A	54	ASN
4	A	55	ASP

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	A	1239/1520 (82%)	1135 (92%)	104 (8%)	11	36
5	B	964/1061 (91%)	876 (91%)	88 (9%)	9	32

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	C	234/274 (85%)	211 (90%)	23 (10%)	8	28
7	D	140/200 (70%)	124 (89%)	16 (11%)	5	24
8	E	196/197 (100%)	187 (95%)	9 (5%)	27	53
9	F	74/137 (54%)	65 (88%)	9 (12%)	5	22
10	G	152/152 (100%)	142 (93%)	10 (7%)	16	43
11	H	117/128 (91%)	111 (95%)	6 (5%)	24	50
12	I	113/116 (97%)	99 (88%)	14 (12%)	4	22
13	J	60/65 (92%)	55 (92%)	5 (8%)	11	36
14	K	99/102 (97%)	92 (93%)	7 (7%)	14	41
15	L	40/57 (70%)	37 (92%)	3 (8%)	13	40
All	All	3428/4009 (86%)	3134 (91%)	294 (9%)	10	35

5 of 294 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	469	GLN
5	B	953	LEU
12	I	78	CYS
5	B	498	THR
5	B	682	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 80 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	236	HIS
5	B	734	HIS
10	G	126	ASN
5	B	363	HIS
5	B	484	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	P	9/10 (90%)	1 (11%)	1 (11%)

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	P	3	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	P	2	A

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

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 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	B	1308	-	26,34,34	3.23	12 (46%)	30,54,54	3.86	15 (50%)

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	B	1308	-	1/1/7/7	6/18/38/38	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)
18	B	1308	G2P	C5-C6	-6.14	1.42	1.52
18	B	1308	G2P	C4-N9	-5.96	1.39	1.47
18	B	1308	G2P	C3'-C4'	-5.48	1.39	1.53
18	B	1308	G2P	C5'-C4'	5.24	1.67	1.51
18	B	1308	G2P	O2'-C2'	5.21	1.55	1.43

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
18	B	1308	G2P	O4'-C4'-C5'	13.10	152.48	109.37
18	B	1308	G2P	O5'-C5'-C4'	6.77	132.31	108.99
18	B	1308	G2P	C5'-C4'-C3'	-6.59	90.48	115.18
18	B	1308	G2P	C4-C5-N7	6.42	110.97	102.46
18	B	1308	G2P	C5-C6-N1	-5.43	111.49	118.19

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
18	B	1308	G2P	C4'

5 of 6 torsion outliers are listed below:

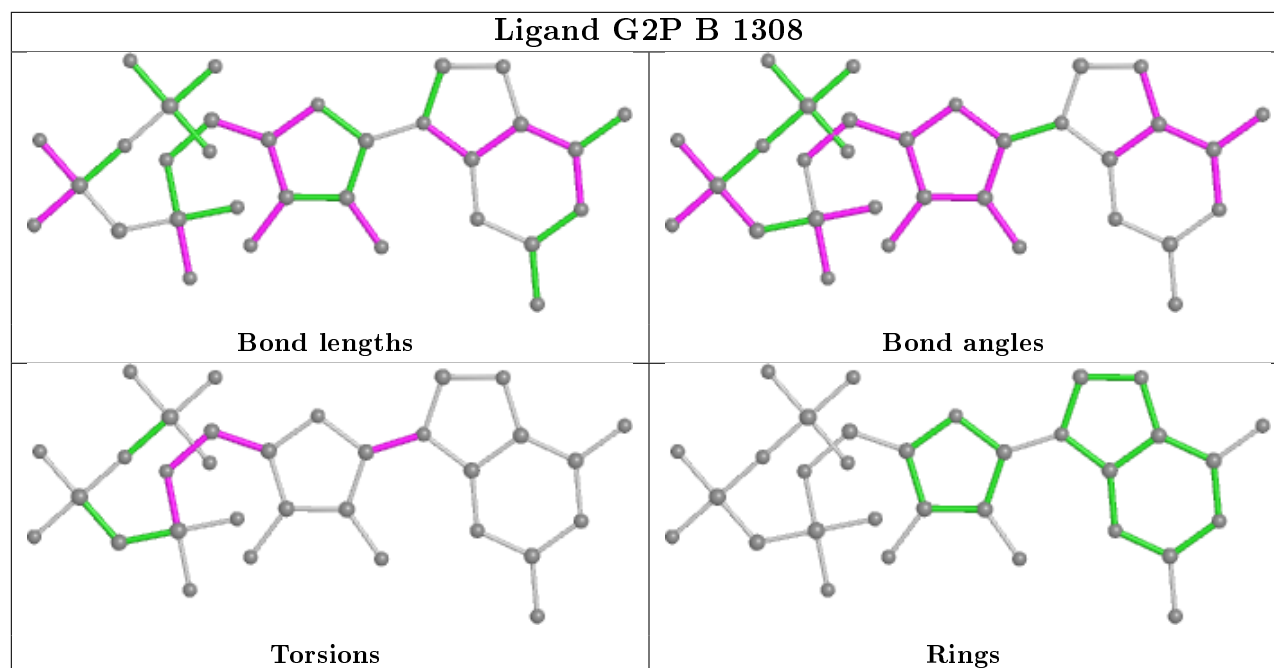
Mol	Chain	Res	Type	Atoms
18	B	1308	G2P	C4'-C5'-O5'-PA
18	B	1308	G2P	O4'-C1'-N9-C4
18	B	1308	G2P	C2'-C1'-N9-C8
18	B	1308	G2P	C2'-C1'-N9-C4
18	B	1308	G2P	O4'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient

equivalents in the CSD to analyse the geometry.



5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
5	B	3
4	A	2
6	C	2

The worst 5 of 7 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	119:ASN	C	120:GLU	N	1.98
1	B	503:GLY	C	504:ARG	N	1.68
1	C	90:ASP	C	91:HIS	N	1.15
1	B	511:PRO	C	512:ARG	N	1.12
1	B	363:HIS	C	364:ILE	N	1.03

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	T	18/19 (94%)	0.91	3 (16%) 1 2	63, 85, 96, 104	0
2	N	6/7 (85%)	1.29	0 100 100	51, 51, 51, 51	0
3	P	10/10 (100%)	0.49	0 100 100	71, 75, 93, 93	0
4	A	1416/1733 (81%)	-0.37	2 (0%) 95 94	45, 110, 184, 223	0
5	B	1112/1224 (90%)	-0.33	2 (0%) 95 93	45, 119, 188, 216	0
6	C	266/318 (83%)	-0.40	0 100 100	60, 102, 161, 182	0
7	D	177/221 (80%)	-0.39	0 100 100	75, 131, 170, 187	0
8	E	214/215 (99%)	-0.34	0 100 100	79, 164, 209, 215	0
9	F	84/155 (54%)	-0.41	0 100 100	50, 81, 128, 147	0
10	G	171/171 (100%)	-0.34	0 100 100	79, 106, 137, 162	0
11	H	133/146 (91%)	0.04	1 (0%) 86 79	124, 161, 197, 206	0
12	I	119/122 (97%)	-0.13	0 100 100	97, 151, 181, 222	0
13	J	65/70 (92%)	-0.57	0 100 100	64, 102, 142, 149	0
14	K	114/120 (95%)	-0.34	0 100 100	65, 106, 136, 143	0
15	L	46/70 (65%)	-0.02	0 100 100	99, 158, 191, 199	0
All	All	3951/4601 (85%)	-0.33	8 (0%) 95 93	45, 116, 188, 223	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	B	471	LYS	3.7
4	A	1455	PRO	2.6
1	T	28	DG	2.5
11	H	139	ASN	2.5
1	T	15	DT	2.2

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 ⓘ

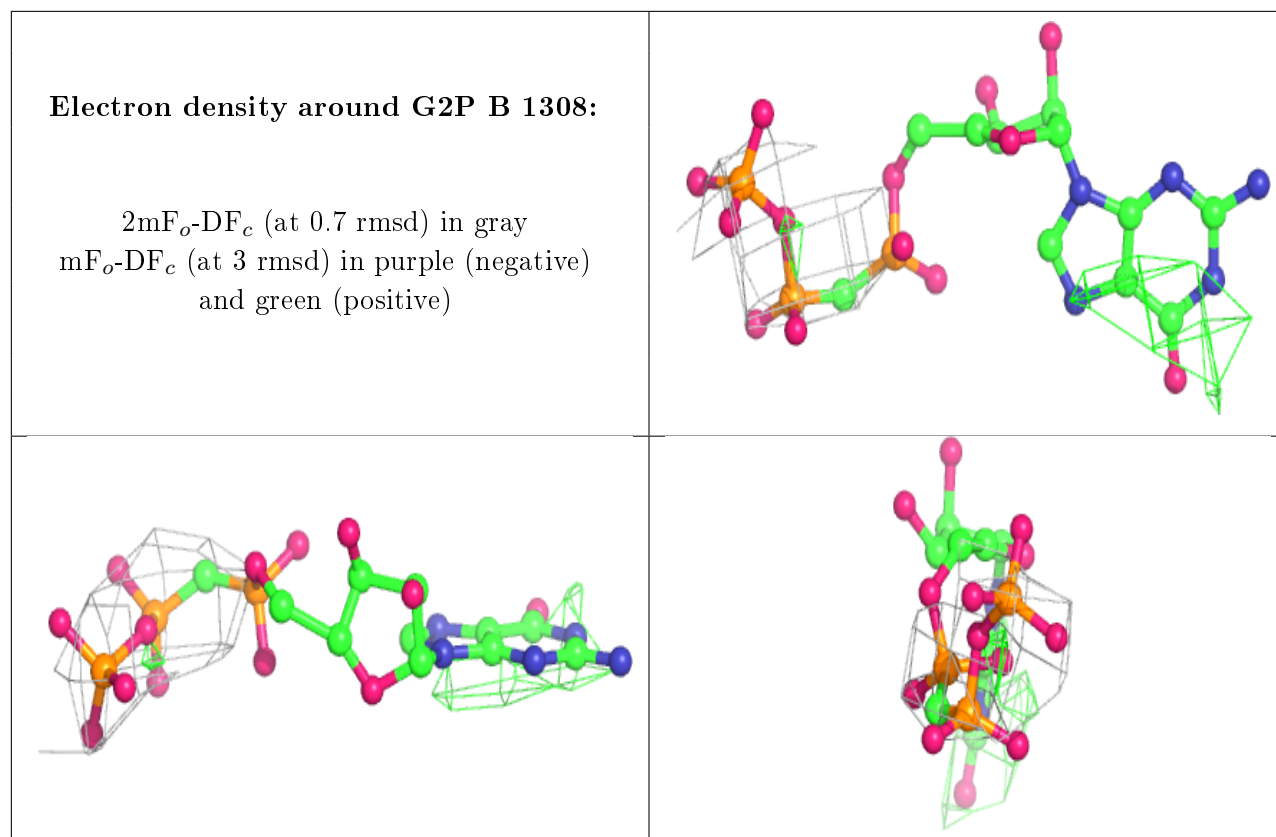
There are no carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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(\AA^2)	Q<0.9
16	ZN	A	1734	1/1	0.91	0.07	112,112,112,112	0
16	ZN	C	319	1/1	0.94	0.06	63,63,63,63	0
16	ZN	A	1735	1/1	0.95	0.06	68,68,68,68	0
16	ZN	B	1307	1/1	0.96	0.15	67,67,67,67	0
16	ZN	I	203	1/1	0.96	0.10	108,108,108,108	0
16	ZN	L	105	1/1	0.97	0.06	130,130,130,130	0
16	ZN	I	204	1/1	0.98	0.27	186,186,186,186	0
18	G2P	B	1308	32/32	-	-	0,0,0,0	32
16	ZN	J	101	1/1	0.96	0.19	90,90,90,90	0
17	MG	A	1736	1/1	0.97	0.20	77,77,77,77	0

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



6.5 Other polymers [i](#)

There are no such residues in this entry.