wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 11:42 PM GMT

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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

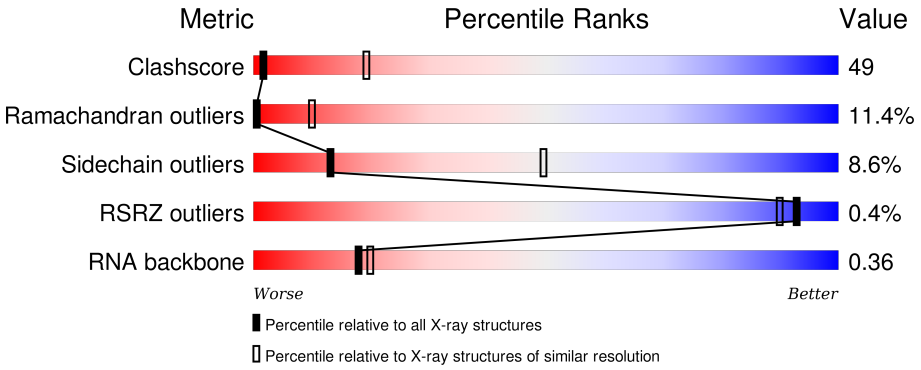
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 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	102246	1003 (5.30-3.62)
Ramachandran outliers	100387	1117 (5.40-3.60)
Sidechain outliers	100360	1099 (5.40-3.60)
RSRZ outliers	91569	1075 (5.40-3.60)
RNA backbone	2183	1090 (6.00-2.80)

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. 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	T	19	26% 21% 58% 21%
2	N	7	14% 14% 86%
3	P	10	10% 80% 20%
4	A	1733	27% 43% 10% • 18%
5	B	1224	29% 48% 13% • 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	C	318	23% 47% 12% 16%
7	D	221	33% 37% 10% 20%
8	E	215	40% 53% 6%
9	F	155	18% 30% 6% 46%
10	G	171	37% 53% 9%
11	H	146	33% 49% 10% 9%
12	I	122	40% 43% 12%
13	J	70	19% 53% 21% 7%
14	K	120	40% 48% 8% 5%
15	L	70	14% 31% 20% 34%

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 (i)

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
			Total	C	N	O	P			
1	T	19	388	185	67	117	19	21	0	0

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

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

- 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
			Total	C	N	O	P			
3	P	10	214	97	44	64	9	0	0	0

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

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

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

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

- 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			

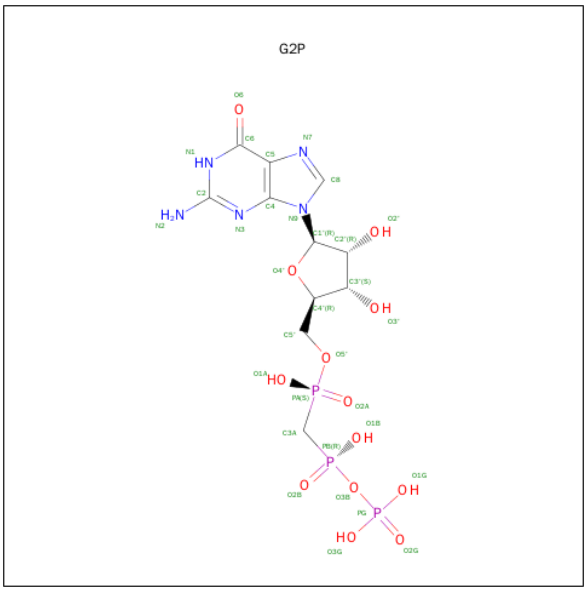
- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

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: $C_{11}H_{18}N_5O_{13}P_3$).



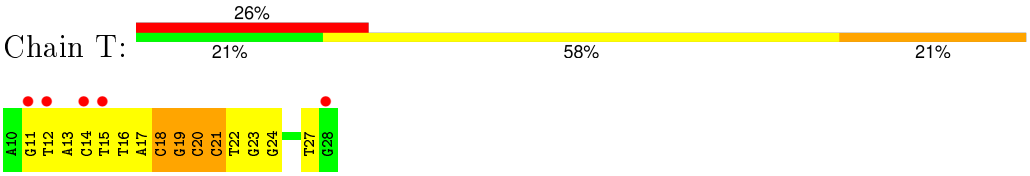
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
18	B	1	Total	C	N	O	P	32	0
			32	11	5	13	3		

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 errors 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*TP*CP*TP*G)-3'

Chain T:

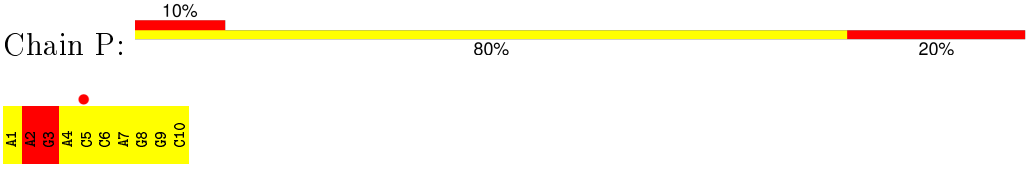


• Molecule 2: 5'-D(*AP*AP*GP*TP*AP*CP*T)-3'

Chain N:

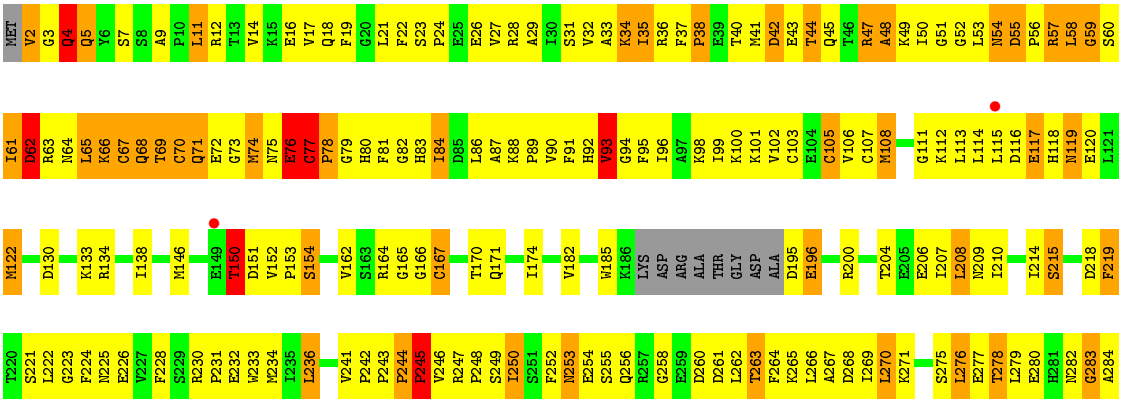
• Molecule 3: 5'-R(*AP*AP*GP*AP*CP*CP*AP*GP*GP*C)-3'

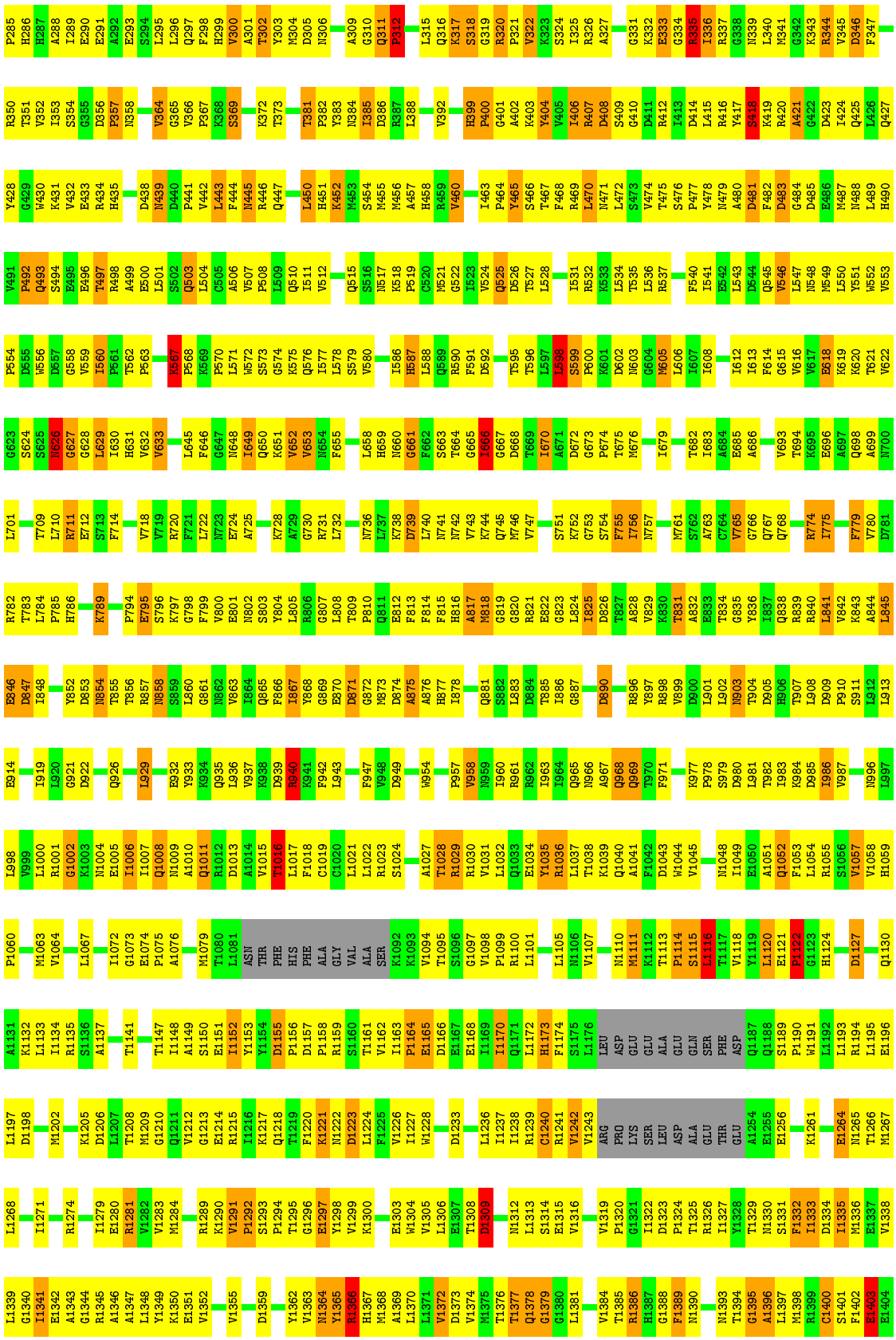
Chain P:

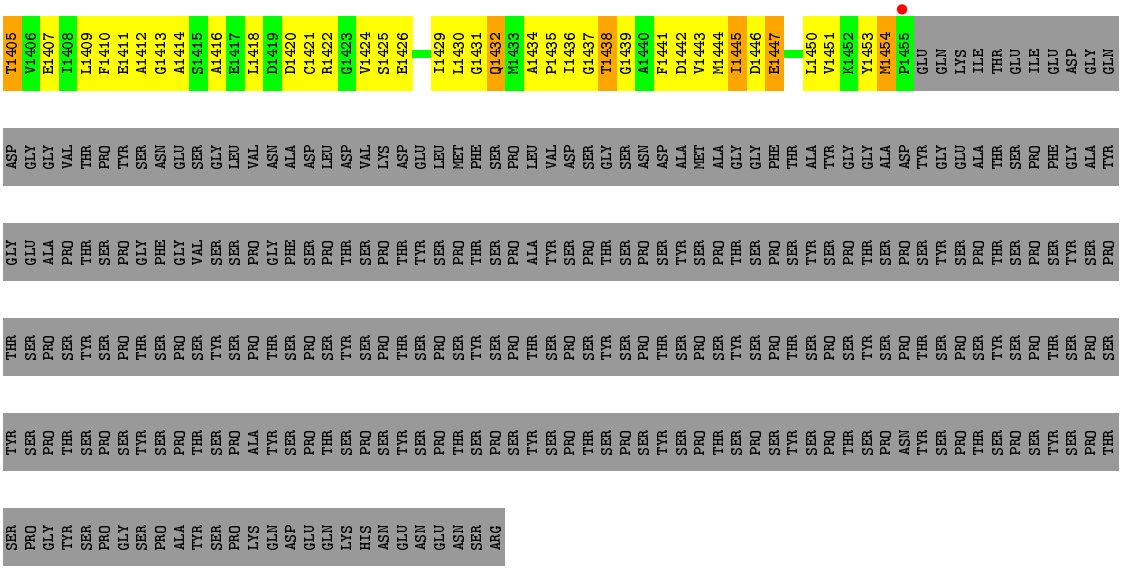


• Molecule 4: DNA-directed RNA polymerase II largest subunit

Chain A:

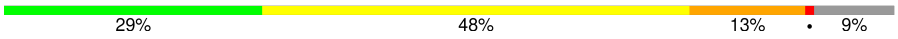


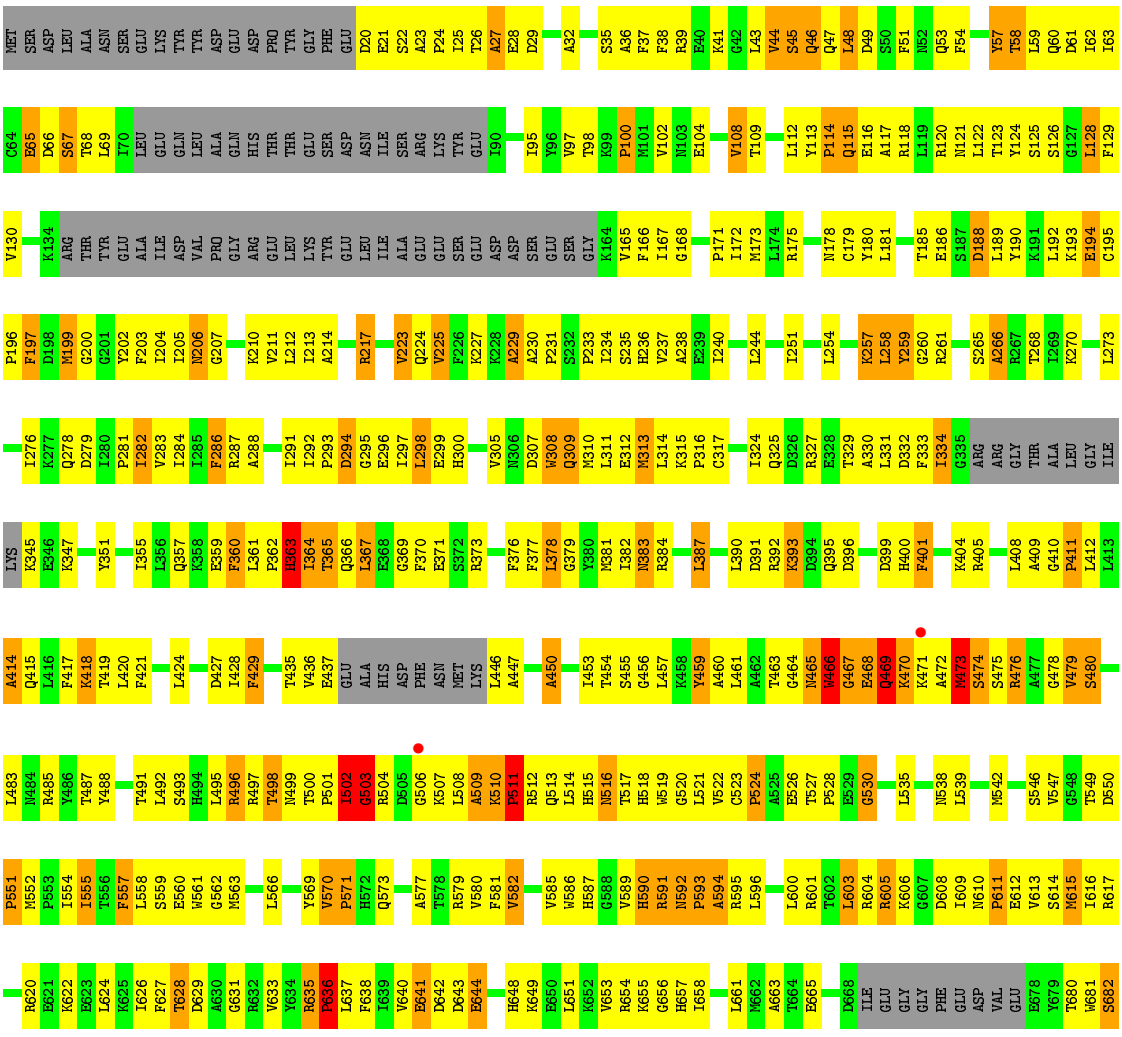

P285 H286 H287 A288 I289 E290 E291 A292 E293 S294 L295 L296 Q297 F298 H299 V300 A301 T302 Y303 M304 D305 N306 A309 G310 Q311 P312 L315 Q316 K317 S318 G319 R320 P321 V322 K323 S324 I325 R326 A327 G331 K332 E333 G334 R335 I336 R337 G338 N339 L340 M341 G342 K343 R344 V345 D346 F347
R350 T351 V352 I353 S354 G355 D356 P357 N358 V364 G365 V366 P367 K368 S369 K372 T373 T381 P382 Y383 N384 I385 D386 R387 L388 V392 H399 P400 G401 A402 K403 Y404 V405 I406 R407 D408 S409 G410 D411 R412 I413 D414 L415 R416 Y417 S418 K419 R420 A421 G422 D423 I424 Q425 L426 Q427
Y428 G429 W430 K431 V432 E433 R434 H435 D438 N439 D440 P441 V442 L443 F444 N445 R446 Q447 L450 H451 K452 M453 S454 M455 M456 A457 H458 R459 V460 I463 P464 Y465 S466 T467 F468 R469 L470 N471 L472 S473 V474 T475 S476 P477 Y478 N479 A480 D481 F482 D483 G484 D485 E486 M487 N488 L489 H490
V491 P492 Q493 S494 E495 E496 T497 R498 A499 E500 L501 S502 Q503 L504 C505 A506 V507 P508 L509 Q510 I511 V512 Q515 S516 M517 K518 P519 Q520 M521 G522 I523 V524 Q525 D526 T527 L528 I531 R532 K533 L534 T535 L536 R537 F540 I541 E542 L543 D544 Q545 V546 L547 N548 M549 L550 Y551 W552 V553
P554 D555 W556 D557 G558 V559 I560 P561 T562 P563 K567 P568 K569 P570 L571 W572 S573 G574 K575 Q576 I577 L578 S579 V580 I586 H587 L588 Q589 R590 F591 D592 T595 T596 L597 L598 S599 P600 K601 D602 M603 G604 M605 L606 I607 I608 I612 I613 F614 G615 V616 V617 E618 K619 K620 T621 V622
G623 S624 S625 N626 G627 G628 L629 I630 H631 V632 V633 L645 F646 K647 N648 I649 Q650 K651 V652 V653 N654 F655 L658 H659 N660 G661 F662 S663 T664 G665 I666 G667 D668 T669 I670 A671 D672 G673 S674 T675 H676 I679 T682 I683 A684 E685 A686 V693 T694 K695 E696 A697 Q698 A699 N700
L701 T709 L710 R711 E712 S713 F714 V718 V719 R720 F721 L722 N723 E724 A725 K728 A729 G730 R731 L732 N736 L737 K738 D739 L740 N741 N742 V743 K744 Q745 M746 V747 S751 K752 G753 S754 F755 I756 N757 M761 S762 A763 C764 V765 G766 Q767 Q768 R774 I775 F779 V780 D781
R782 T783 L784 P785 H786 K789 P794 E795 S796 K797 G798 F799 V800 E801 N802 S803 Y804 L805 R806 G807 L808 T809 P810 Q811 E812 F813 F814 F815 H816 A817 M818 G819 G820 R821 E822 G823 L824 I825 D826 T827 A828 V829 K830 T831 A832 E833 T834 G835 Y836 I837 Q838 R839 R840 L841 V842 K843 A844 L845
E846 D847 I848 Y852 D853 N854 T855 T856 R857 N858 S859 L860 G861 N862 V863 I864 Q865 F866 I867 Y868 G869 E870 D871 G872 M873 D874 A875 A876 H877 I878 Q881 S882 L883 D884 T885 I886 G887 D890 R896 Y897 R898 V899 D900 L901 L902 N903 T904 D905 H906 T907 L908 D909 P910 S911 I912 L913
E914 I919 L920 G921 D922 Q926 L929 E932 Y933 K934 Q935 L936 V937 K938 D939 R940 K941 F942 L943 F947 V948 D949 W954 P957 V958 N959 I960 R961 R962 I963 I964 Q965 N966 A967 Q968 Q969 T970 F971 K977 P978 S979 D980 L981 T982 I983 K984 D985 I986 V987 N996 L997
L998 V999 L1000 R1001 G1002 K1003 N1004 E1005 I1006 I1007 Q1008 N1009 A1010 Q1011 R1012 D1013 A1014 V1015 T1016 L1017 F1018 C1019 C1020 L1021 L1022 R1023 S1024 A1027 T1028 R1029 R1030 V1031 L1032 Q1033 E1034 Y1035 R1036 L1037 T1038 K1039 Q1040 A1041 F1042 D1043 W1044 V1045 N1048 I1049 E1050 A1051 Q1052 F1053 L1054 R1055 S1056 V1057 V1058 H1059
P1060 M1063 V1064 L1067 I1072 G1073 E1074 P1075 A1076 M1079 T1080 L1081 ASN THR PHE HIS PHE ALA GLY VAL ALA SER K1092 K1093 V1094 V1095 S1096 G1097 V1098 P1099 R1100 L1101 L1105 N1106 V1107 N1110 M1111 K1112 T1113 P1114 S1115 L1116 T1117 V1118 Y1119 L1120 E1121 P1122 G1123 H1124 D1127 Q1130
A1131 K1132 L1133 I1134 R1135 S1136 A1137 T1141 T1147 I1148 A1149 S1150 E1151 I1152 Y1153 Y1154 D1155 P1156 D1157 P1158 R1159 S1160 T1161 V1162 I1163 P1164 E1165 D1166 E1167 E1168 I1169 I1170 Q1171 L1172 H1173 F1174 S1175 L1176 LEU ASP GLU GLU ALA GLU GLN SER PHE ASP Q1187 Q1188 S1189 P1190 W1191 L1192 L1193 R1194 L1195 E1196
L1197 D1198 M1202 K1205 D1206 L1207 T1208 M1209 G1210 Q1211 V1212 G1213 E1214 R1215 I1216 K1217 Q1218 T1219 F1220 K1221 N1222 D1223 L1224 F1225 V1226 I1227 W1228 D1233 L1236 I1237 I1238 R1239 C1240 R1241 V1242 V1243 ARG PRO LYS SER LEU ASP ALA GLU THR GLU A1254 E1255 E1256 K1261 E1264 N1265 T1266 M1267
L1268 I1271 R1274 I1279 E1280 R1281 V1282 V1283 M1284 R1289 K1290 V1291 P1292 S1293 P1294 T1295 G1296 E1297 Y1298 V1299 K1300 E1303 W1304 V1305 L1306 E1307 T1308 D1309 N1312 L1313 S1314 E1315 V1316 V1319 P1320 G1321 I1322 D1323 P1324 T1325 R1326 I1327 Y1328 T1329 N1330 S1331 F1332 I1333 D1334 I1335 M1336 E1337 V1338
L1339 G1340 I1341 E1342 A1343 G1344 R1345 A1346 A1347 L1348 Y1349 K1350 E1351 V1352 V1355 D1359 Y1362 V1363 N1364 Y1365 R1366 H1367 M1368 A1369 L1370 L1371 V1372 D1373 V1374 M1375 T1376 T1377 Q1378 G1379 G1380 L1381 V1384 T1385 R1386 H1387 G1388 F1389 N1390 N1393 T1394 G1395 A1396 L1397 M1398 R1399 C1400 S1401 F1402 E1403 E1404

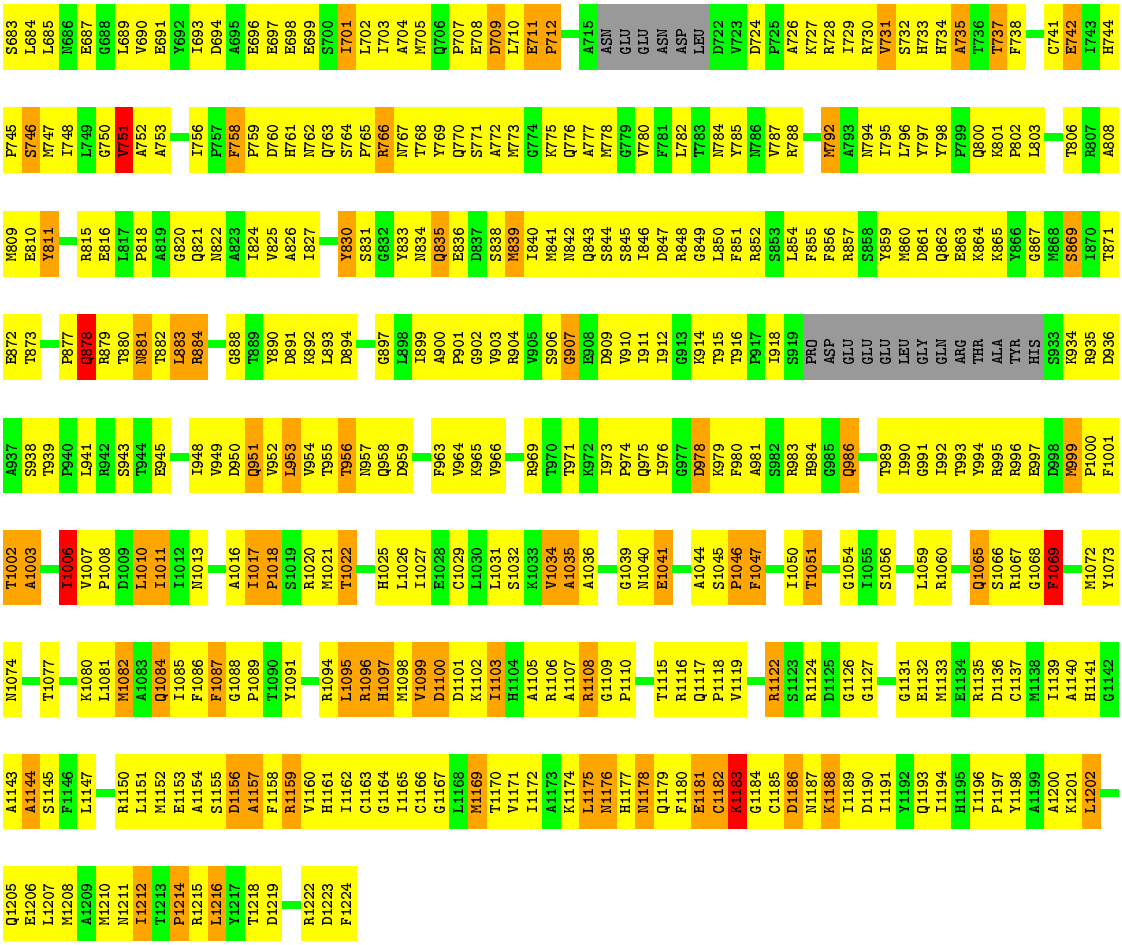


• Molecule 5: DNA-directed RNA polymerase II 140 kDa polypeptide

Chain B:

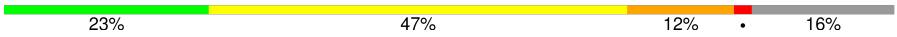


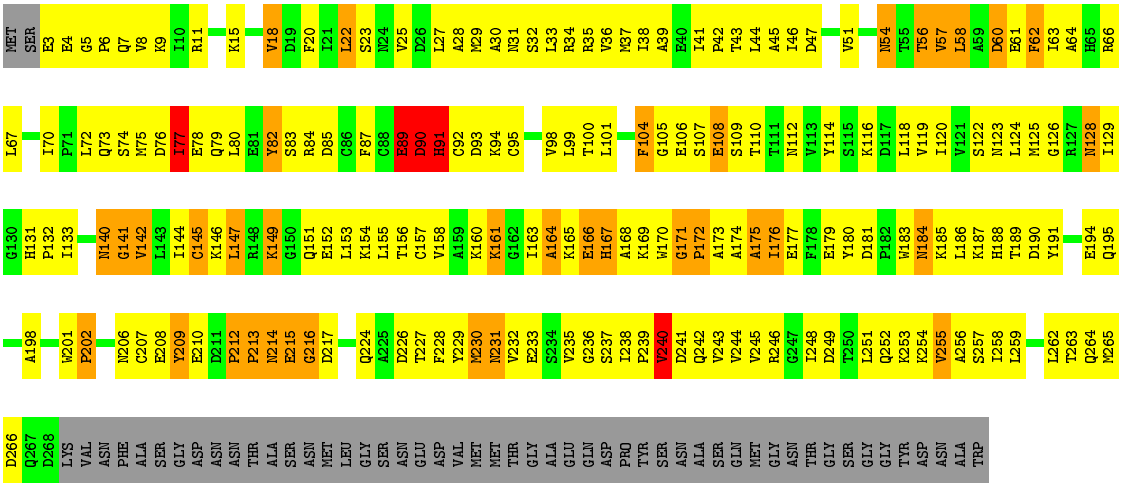




• Molecule 6: DNA-directed RNA polymerase II 45 kDa polypeptide

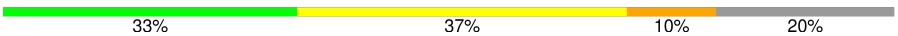
Chain C:

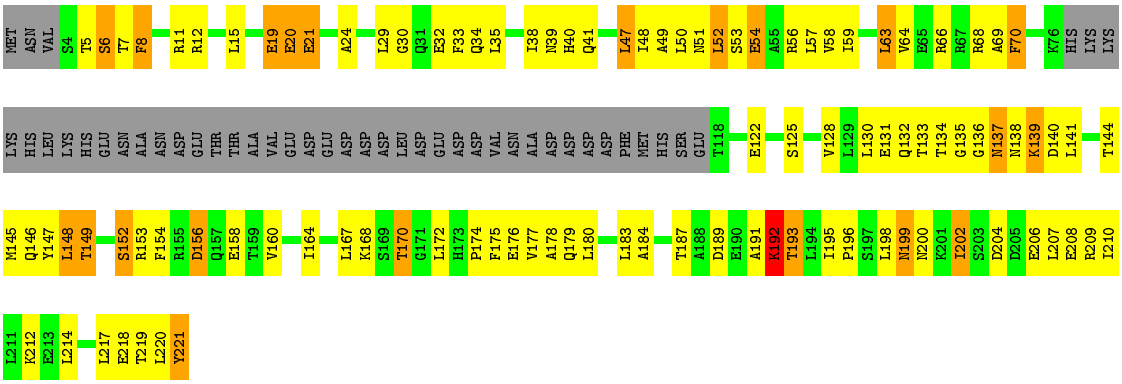




• Molecule 7: DNA-directed RNA polymerase II 32 kDa polypeptide

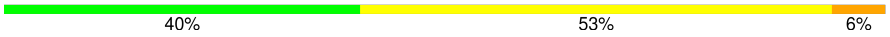
Chain D:

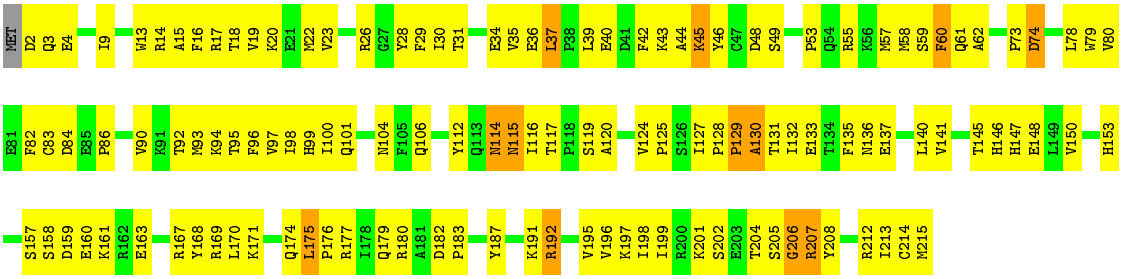




• Molecule 8: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

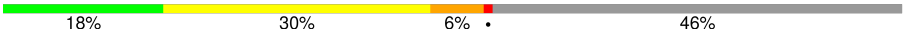
Chain E:

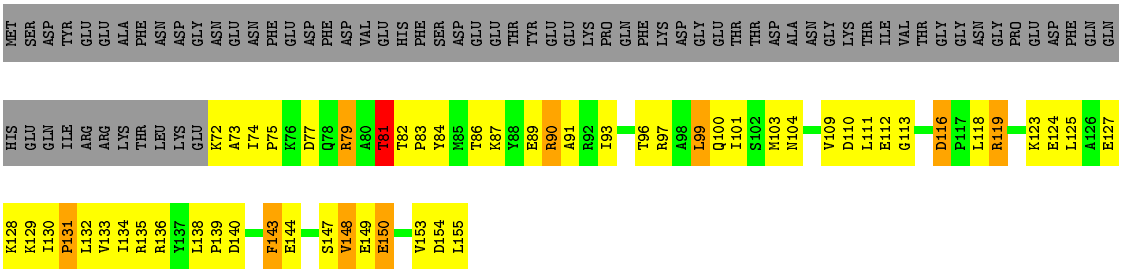




• Molecule 9: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

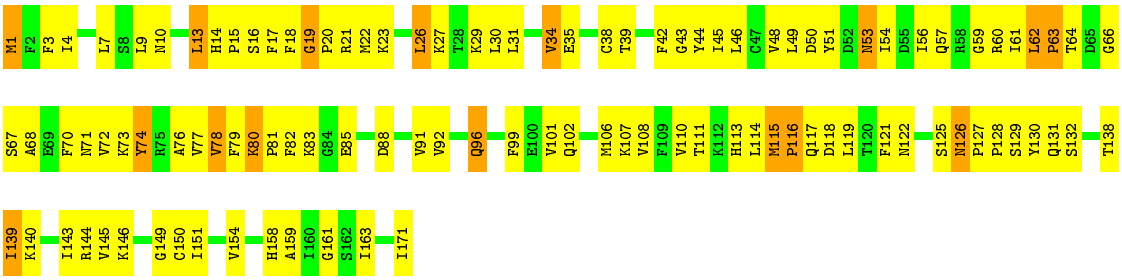
Chain F:





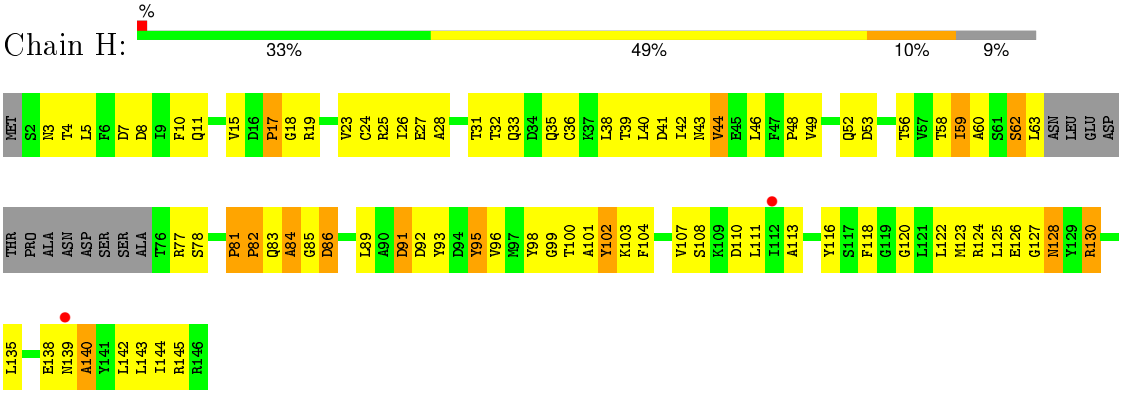
• Molecule 10: DNA-directed RNA polymerase II 19 kDa polypeptide

Chain G:



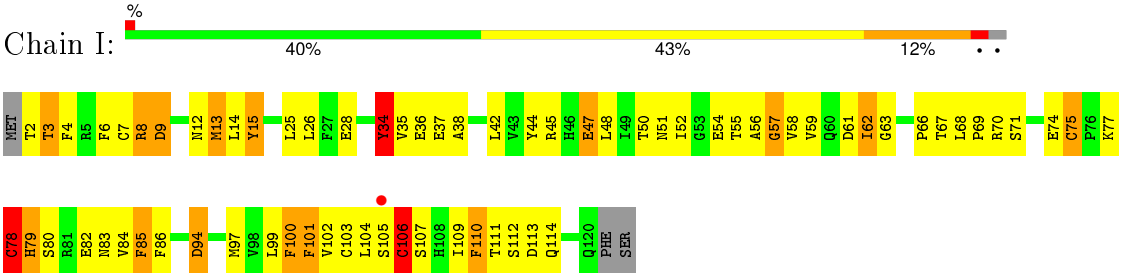
• Molecule 11: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

Chain H:



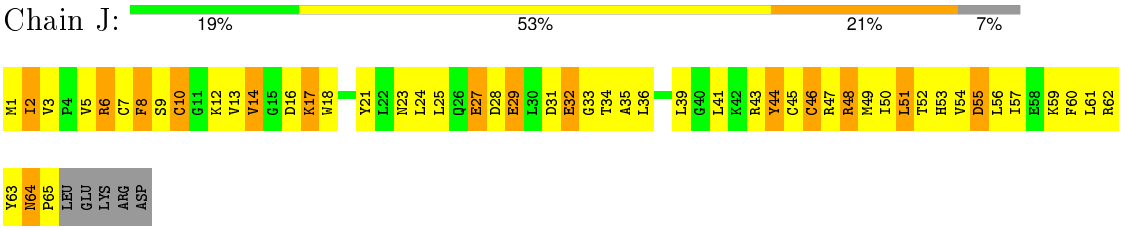
- Molecule 12: DNA-directed RNA polymerase II subunit 9

Chain I:



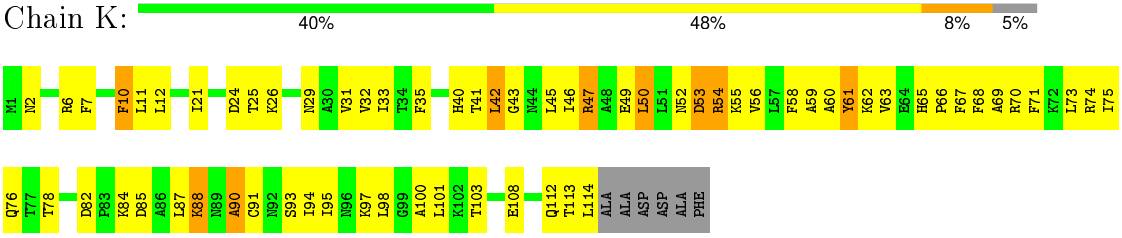
- Molecule 13: DNA-directed RNA polymerases I/II/III subunit 10

Chain J:



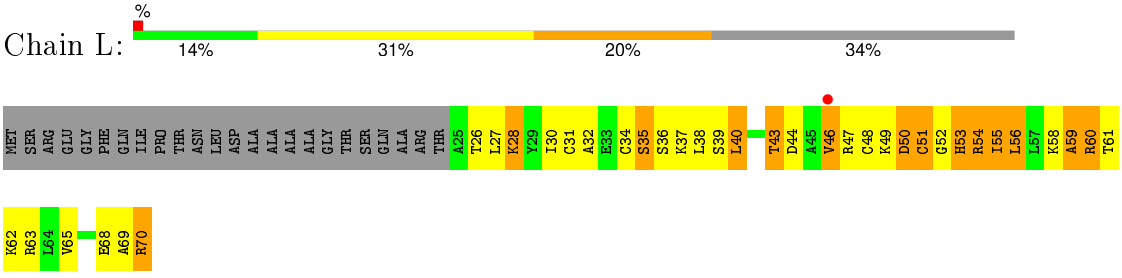
- Molecule 14: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K:



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

Chain L:



4 Data and refinement statistics (i)

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
$< I/\sigma(I) >$ [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.	DCC
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 , 54.3	EDS
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
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.28$	Xtriage
Outliers	0 of 58178 reflections	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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(^{o})	Ideal(^{o})
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 (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	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 (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	1406/1733 (81%)	962 (68%)	280 (20%)	164 (12%)	0 9
5	B	1096/1224 (90%)	740 (68%)	225 (20%)	131 (12%)	0 9
6	C	264/318 (83%)	158 (60%)	67 (25%)	39 (15%)	0 5
7	D	173/221 (78%)	121 (70%)	34 (20%)	18 (10%)	1 12
8	E	212/215 (99%)	149 (70%)	48 (23%)	15 (7%)	1 23
9	F	82/155 (53%)	64 (78%)	14 (17%)	4 (5%)	3 32
10	G	169/171 (99%)	131 (78%)	26 (15%)	12 (7%)	1 23
11	H	129/146 (88%)	84 (65%)	30 (23%)	15 (12%)	0 9
12	I	117/122 (96%)	81 (69%)	27 (23%)	9 (8%)	1 20
13	J	63/70 (90%)	37 (59%)	11 (18%)	15 (24%)	0 1
14	K	112/120 (93%)	89 (80%)	18 (16%)	5 (4%)	3 33
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 10

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 (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	1239/1520 (82%)	1135 (92%)	104 (8%)	14 51
5	B	964/1061 (91%)	876 (91%)	88 (9%)	12 46

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
6	C	234/274 (85%)	211 (90%)	23 (10%)	10 42
7	D	140/200 (70%)	124 (89%)	16 (11%)	7 35
8	E	196/197 (100%)	187 (95%)	9 (5%)	33 70
9	F	74/137 (54%)	65 (88%)	9 (12%)	6 32
10	G	152/152 (100%)	142 (93%)	10 (7%)	21 60
11	H	117/128 (91%)	111 (95%)	6 (5%)	29 67
12	I	113/116 (97%)	99 (88%)	14 (12%)	6 32
13	J	60/65 (92%)	55 (92%)	5 (8%)	14 51
14	K	99/102 (97%)	92 (93%)	7 (7%)	18 58
15	L	40/57 (70%)	37 (92%)	3 (8%)	17 56
All	All	3428/4009 (86%)	3134 (91%)	294 (9%)	13 49

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

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	2.60	9 (34%)	33,54,54	4.41	17 (51%)

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	0/15/38/38	0/3/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
18	B	1308	G2P	C3'-C4'	-5.20	1.39	1.53
18	B	1308	G2P	O3'-C3'	-4.27	1.32	1.43
18	B	1308	G2P	PB-O1B	-4.25	1.46	1.56
18	B	1308	G2P	O4'-C4'	-4.11	1.35	1.45
18	B	1308	G2P	PA-O1A	-3.04	1.48	1.56

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
18	B	1308	G2P	C2'-C1'-N9	-10.75	97.87	114.29
18	B	1308	G2P	C5-C6-N1	-8.85	111.49	123.59
18	B	1308	G2P	C5'-C4'-C3'	-6.23	90.48	115.21
18	B	1308	G2P	O5'-PA-O2A	-4.78	101.28	113.98
18	B	1308	G2P	O3'-C3'-C2'	-4.55	97.04	111.83

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
18	B	1308	G2P	C4'

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	T	18/19 (94%)	1.15	5 (27%) 1 2	63, 85, 96, 104	0
2	N	6/7 (85%)	1.62	1 (16%) 2 4	51, 51, 51, 51	0
3	P	10/10 (100%)	0.64	1 (10%) 9 8	71, 75, 93, 93	0
4	A	1416/1733 (81%)	-0.32	3 (0%) 95 94	45, 110, 184, 223	0
5	B	1112/1224 (90%)	-0.28	2 (0%) 95 94	45, 119, 188, 216	0
6	C	266/318 (83%)	-0.34	0 100 100	60, 102, 161, 182	0
7	D	177/221 (80%)	-0.34	0 100 100	75, 131, 170, 187	0
8	E	214/215 (99%)	-0.28	0 100 100	79, 164, 209, 215	0
9	F	84/155 (54%)	-0.36	0 100 100	50, 81, 128, 147	0
10	G	171/171 (100%)	-0.29	0 100 100	79, 106, 137, 162	0
11	H	133/146 (91%)	0.12	2 (1%) 76 67	124, 161, 197, 206	0
12	I	119/122 (97%)	-0.04	1 (0%) 87 82	97, 151, 181, 222	0
13	J	65/70 (92%)	-0.52	0 100 100	64, 102, 142, 149	0
14	K	114/120 (95%)	-0.28	0 100 100	65, 106, 136, 143	0
15	L	46/70 (65%)	0.04	1 (2%) 65 56	99, 158, 191, 199	0
All	All	3951/4601 (85%)	-0.27	16 (0%) 93 90	45, 116, 188, 223	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	B	471	LYS	3.8
1	T	28	DG	2.7
4	A	1455	PRO	2.7
11	H	139	ASN	2.5
1	T	11	DG	2.4

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

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

6.3 Carbohydrates (i)

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
16	ZN	B	1307	1/1	0.96	0.15	-0.80	67,67,67,67	0
16	ZN	J	101	1/1	0.96	0.19	-1.05	90,90,90,90	0
16	ZN	C	319	1/1	0.94	0.06	-1.18	63,63,63,63	0
16	ZN	I	203	1/1	0.96	0.09	-1.21	108,108,108,108	0
16	ZN	I	204	1/1	0.98	0.28	-1.22	186,186,186,186	0
16	ZN	L	105	1/1	0.97	0.06	-1.82	130,130,130,130	0
16	ZN	A	1735	1/1	0.95	0.06	-2.12	68,68,68,68	0
16	ZN	A	1734	1/1	0.91	0.07	-3.43	112,112,112,112	0
18	G2P	B	1308	32/32	-	-	-	0,0,0,0	32
17	MG	A	1736	1/1	0.97	0.20	-	77,77,77,77	0

6.5 Other polymers (i)

There are no such residues in this entry.