wwPDB X-ray Structure Validation Summary Report (i)

Aug 27, 2023 – 06:55 AM EDT

PDB ID	:	3H3V
Title	:	Yeast RNAP II containing poly(A)-signal sequence in the active site
Authors	:	Dengl, S.; Cramer, P.
Deposited on	:	2009-04-17
Resolution	:	4.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

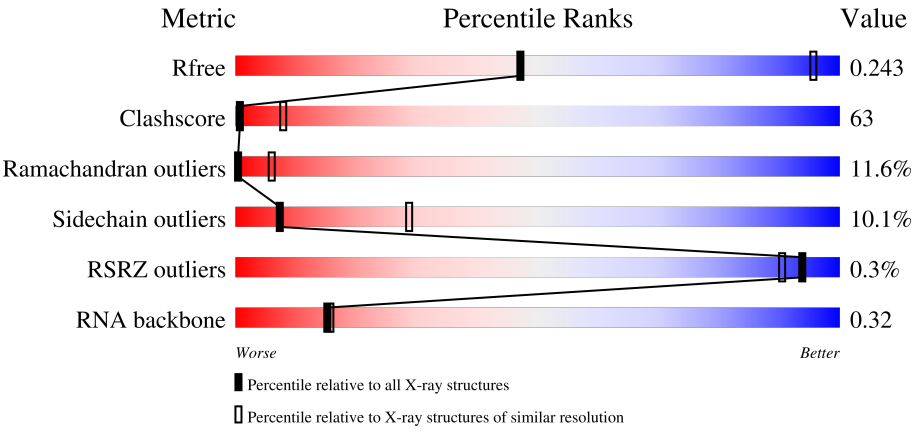
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 4.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



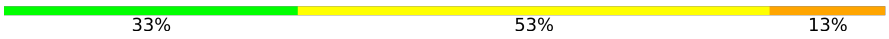
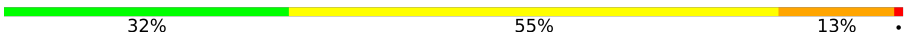
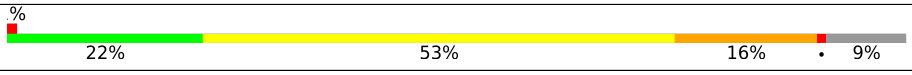
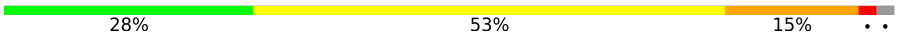
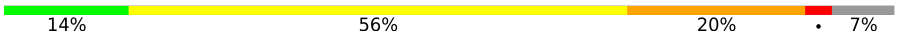
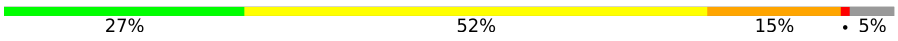
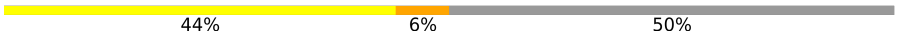
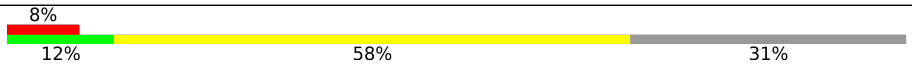
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1087 (4.30-3.70)
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)
RNA backbone	3102	1048 (5.00-3.00)

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 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	B	1733	22% 47% 11% • 18%
2	C	1224	21% 55% 14% • 9%
3	D	318	21% 49% 12% • 16%
4	E	221	19% 45% 13% • 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	F	215	
6	G	155	
7	H	171	
8	I	146	
9	J	122	
10	K	70	
11	L	120	
12	M	70	
13	N	14	
14	P	16	
15	T	26	

2 Entry composition (i)

There are 17 unique types of molecules in this entry. The entry contains 31777 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-directed RNA polymerase II subunit RPB1.

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	1108	8810	5580	1541	1634	55	0	0	0

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

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

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	177	1427	882	256	287	2	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

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

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

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

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

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

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	M	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

- Molecule 13 is a DNA chain called 5'-D(*AP*GP*CP*TP*CP*AP*AP*GP*TP*AP*GP*CP*TP*GP*CP*TP*TP*TP*AP*TP*TP*GP*CP*AP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	N	7	Total	C	N	O	P	11	0	0
			138	67	26	39	6			

- Molecule 14 is a RNA chain called 5'-D(*CP*AP*GP*CP*TP*AP*CP*TP*TP*GP*AP*GP*CP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	P	8	Total	C	N	O	P	0	0	0
			168	77	33	51	7			

- Molecule 15 is a DNA chain called 5'-R(*UP*GP*CP*AP*UP*UP*UP*CP*GP*CP*AP*AP*UP*AP*AP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	T	18	Total	C	N	O	P	8	0	0
			365	177	60	111	17			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	B	2	Total	Zn	0	0
			2	2		
16	C	1	Total	Zn	0	0
			1	1		
16	D	1	Total	Zn	0	0
			1	1		
16	J	2	Total	Zn	0	0
			2	2		
16	K	1	Total	Zn	0	0
			1	1		
16	M	1	Total	Zn	0	0
			1	1		

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

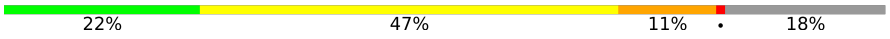
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	B	1	Total	Mg	0	0
			1	1		

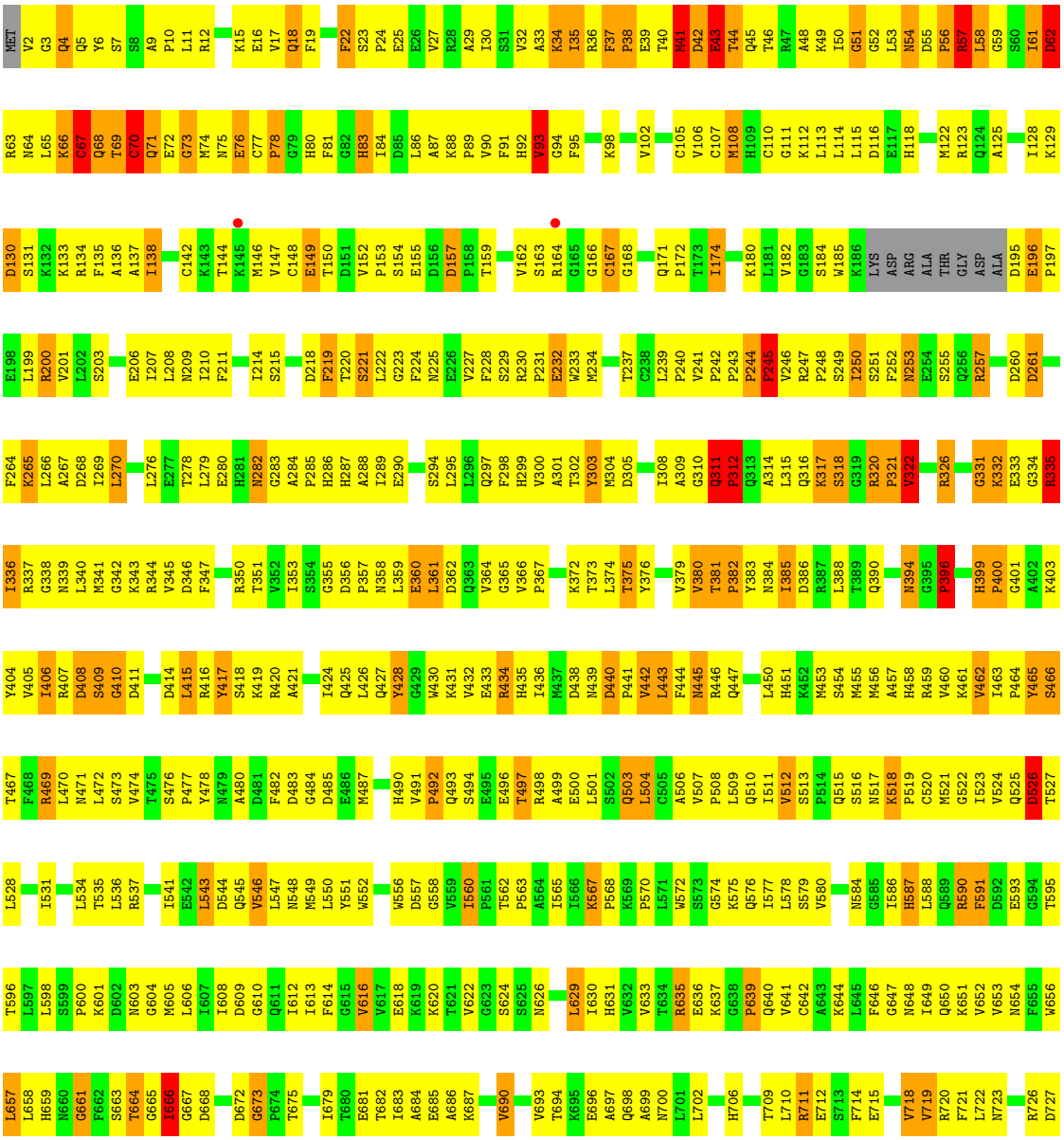
3 Residue-property plots (i)

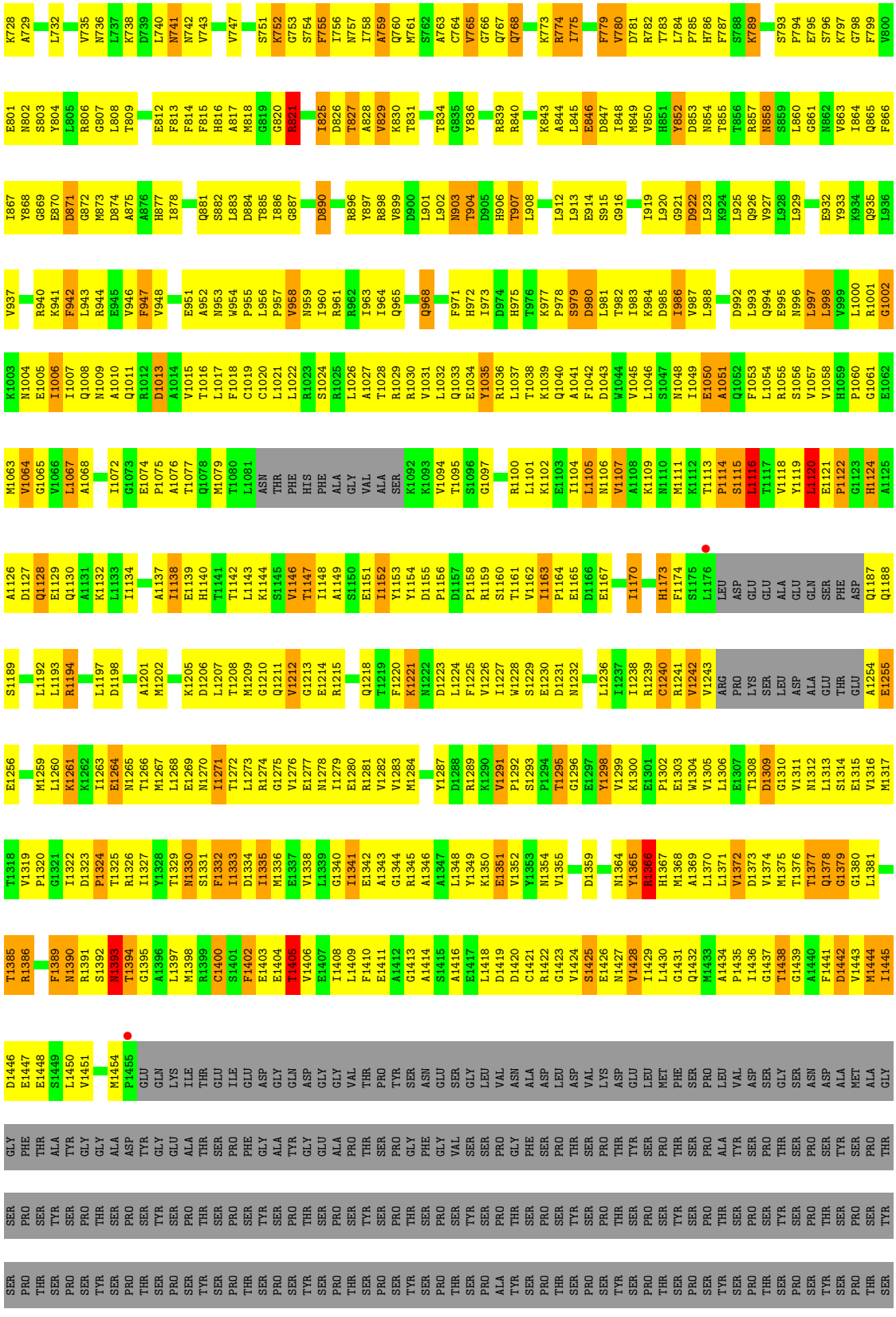
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

- Molecule 1: DNA-directed RNA polymerase II subunit RPB1

Chain B:

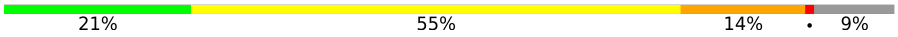


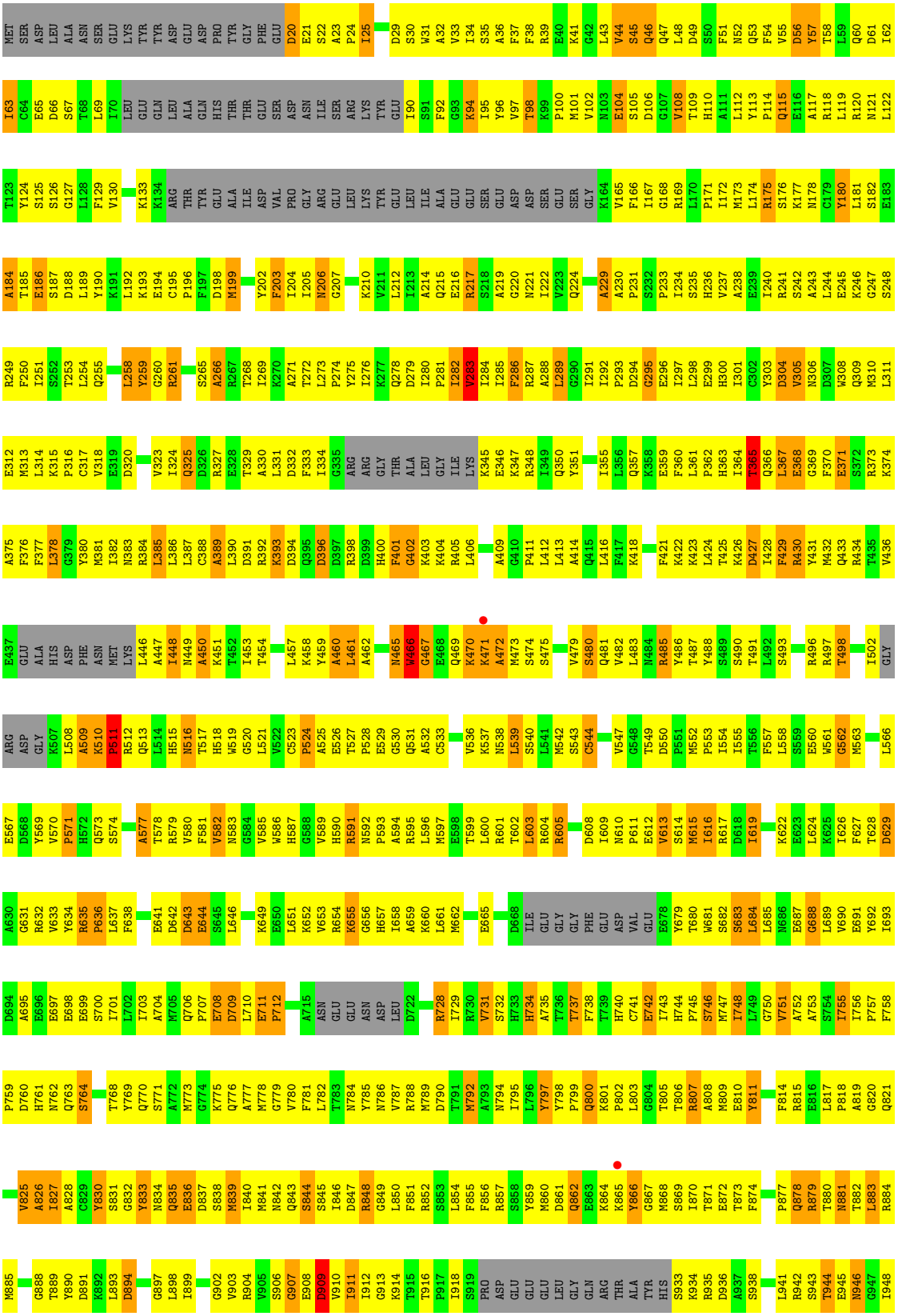


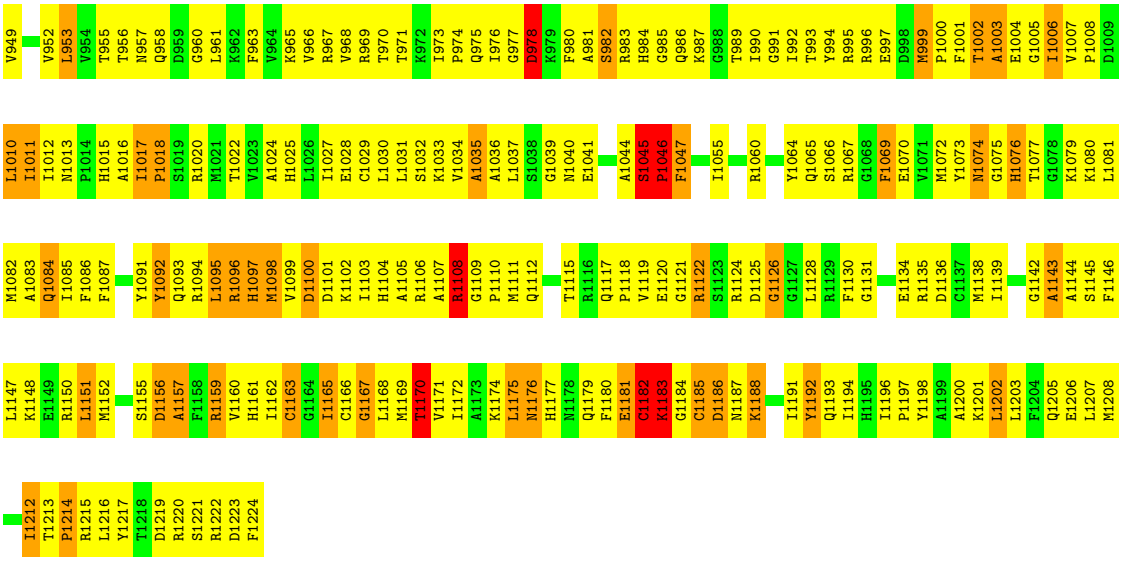

K728 A729 L732 V735 N736 L737 K738 D739 L740 N741 N742 V743 V747 S751 K752 G753 S754 F755 I756 N757 I758 A759 Q760 M761 S762 A763 C764 V765 G766 Q767 Q768 K773 R774 I775 F779 V780 D781 R782 T783 L784 P785 H786 F787 S788 K789 S793 P794 E795 S796 K797 G798 F799 V800
E801 N802 S803 Y804 L805 R806 G807 L808 T809 E812 F813 F814 F815 H816 A817 M818 G819 G820 R821 I825 D826 T827 A828 A829 K830 T831 T834 G835 Y836 R839 R840 K843 A844 L845 E846 D847 I848 M849 V850 H851 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 L912 L913 E914 S915 G916 I919 L920 G921 D922 L923 K924 L925 Q926 V927 L928 L929 E932 Y933 K934 Q935 L936
V937 R940 K941 F942 L943 R944 E945 V946 F947 V948 E951 A952 N953 W954 P955 L956 P957 V958 N959 I960 R961 R962 I963 I964 Q965 Q968 F971 H972 I973 D974 H975 T976 K977 P978 S979 D980 L981 T982 I983 K984 D985 I986 V987 L988 D992 L993 Q994 E995 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 R1025 L1026 A1027 T1028 R1029 R1030 V1031 L1032 Q1033 E1034 Y1035 R1036 L1037 T1038 K1039 Q1040 A1041 F1042 D1043 W1044 V1045 L1046 S1047 N1048 I1049 E1050 A1051 Q1052 F1053 L1054 R1055 S1056 V1057 V1058 H1059 P1060 G1061 E1062
M1063 V1064 G1065 V1066 L1067 A1068 I1072 G1073 E1074 P1075 A1076 T1077 Q1078 M1079 T1080 L1081 ASN THR PHE HIS PHE ALA GLY VAL ALA SER K1092 K1093 V1094 T1095 S1096 G1097 R1100 L1101 K1102 E1103 I1104 L1105 N1106 V1107 A1108 K1109 N1110 M1111 K1112 T1113 P1114 S1115 L1116 T1117 V1118 Y1119 L1120 E1121 P1122 G1123 H1124 A1125
A1126 D1127 Q1128 E1129 Q1130 A1131 K1132 L1133 I1134 A1137 I1138 E1139 H1140 T1141 T1142 L1143 K1144 S1145 V1146 T1147 I1148 A1149 S1150 E1151 I1152 Y1153 Y1154 D1155 P1156 D1157 P1158 R1159 S1160 T1161 V1162 I1163 P1164 E1165 D1166 E1167 I1170 H1173 F1174 S1175 L1176 LEU ASP GLU GLU ALA GLU GLN SER PHE ASP Q1187 Q1188
S1189 L1192 L1193 R1194 L1197 D1198 A1201 M1202 K1205 D1206 L1207 T1208 M1209 G1210 Q1211 V1212 G1213 E1214 R1215 Q1218 T1219 F1220 K1221 N1222 D1223 L1224 F1225 V1226 I1227 W1228 S1229 E1230 D1231 N1232 L1236 I1237 I1238 R1239 C1240 R1241 V1242 V1243 ARG PRO LYS SER LEU ASP ALA GLU THR GLU A1254 E1255
E1256 M1259 L1260 K1261 K1262 I1263 E1264 N1265 T1266 M1267 L1268 E1269 N1270 I1271 T1272 L1273 R1274 G1275 V1276 E1277 N1278 I1279 E1280 R1281 V1282 V1283 M1284 Y1287 D1288 R1289 K1290 V1291 P1292 S1293 P1294 T1295 G1296 E1297 Y1298 V1299 K1300 E1301 P1302 E1303 W1304 V1305 L1306 E1307 T1308 D1309 G1310 V1311 N1312 L1313 S1314 E1315 V1316 M1317
T1318 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 Y1353 N1354 V1355 D1359 N1364 Y1365 R1366 H1367 M1368 A1369 L1370 L1371 V1372 D1373 V1374 M1375 T1376 T1377 Q1378 G1379 G1380 L1381
T1385 R1386 F1389 N1390 R1391 S1392 N1393 T1394 G1395 A1396 L1397 M1398 R1399 C1400 S1401 F1402 E1403 E1404 T1405 V1406 E1407 I1408 L1409 F1410 E1411 A1412 G1413 A1414 S1415 A1416 E1417 L1418 D1419 D1420 C1421 R1422 G1423 V1424 S1425 E1426 N1427 V1428 I1429 L1430 G1431 Q1432 M1433 A1434 P1435 I1436 G1437 T1438 G1439 A1440 F1441 D1442 V1443 M1444 I1445
D1446 E1447 E1448 S1449 L1450 V1451 M1454 P1455 GLU GLN LYS ILE THR GLU ILE GLU ASP GLY GLN ASP GLY 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 SER GLY SER PRO ASN ASP SER ALA MET GLY
GLY PHE THR ALA TYR GLY GLY ALA ASP TYR GLY GLU ALA PRO THR SER PRO GLY PHE GLY VAL SER SER GLY GLU ALA THR SER PRO PHE GLY GLY ALA TYR GLY GLU ALA PRO THR SER PRO GLY PHE GLY GLY ALA TYR SER PRO GLY GLY SER PRO ALA TYR SER PRO LYS GLN ASP GLU GLN LYS HIS ASN GLU ASN GLU ASN SER ARG
SER PRO SER TYR SER PRO PRO THR SER TYR SER PRO SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR
SER PRO THR SER PRO SER PRO THR SER TYR SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR SER PRO SER PRO SER TYR
PRO SER TYR SER PRO THR SER TYR SER PRO PRO ASN TYR SER PRO SER PRO ASN GLU ASN GLU ASN SER ARG

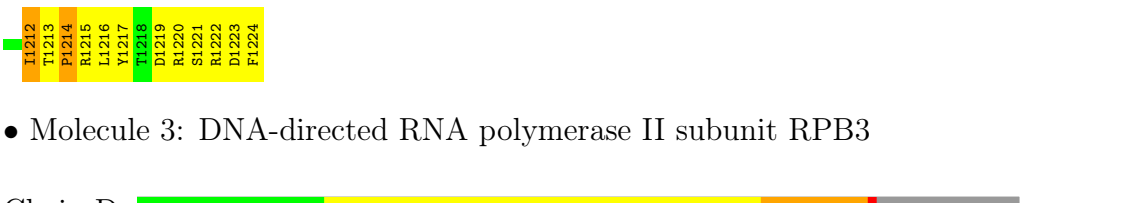
• Molecule 2: DNA-directed RNA polymerase II subunit RPB2

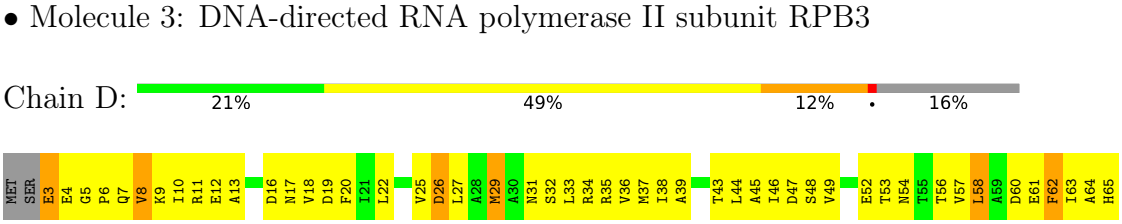
Chain C:

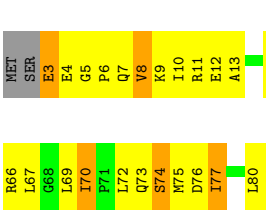






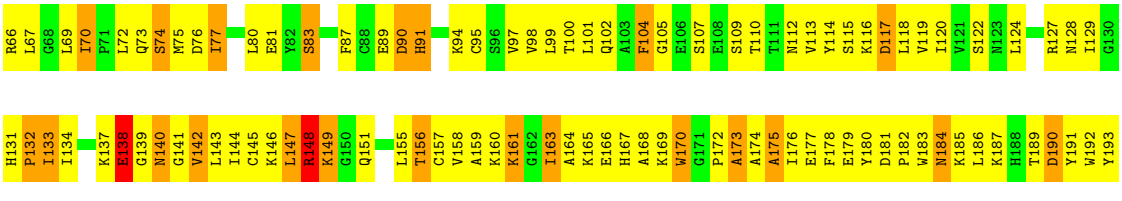


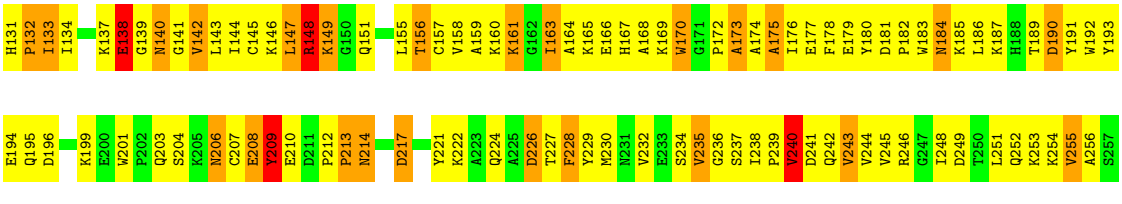


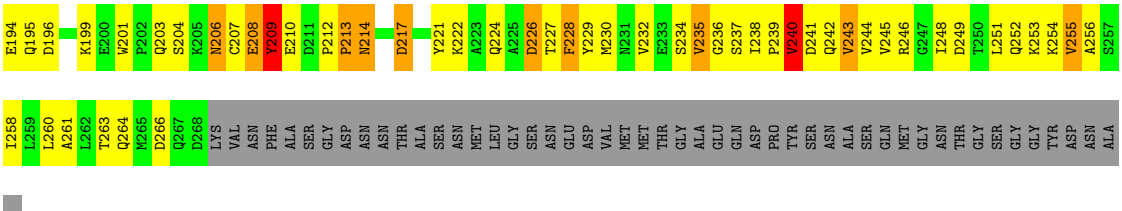


• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain D:

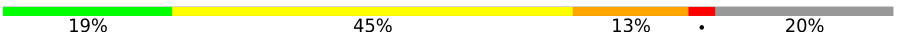


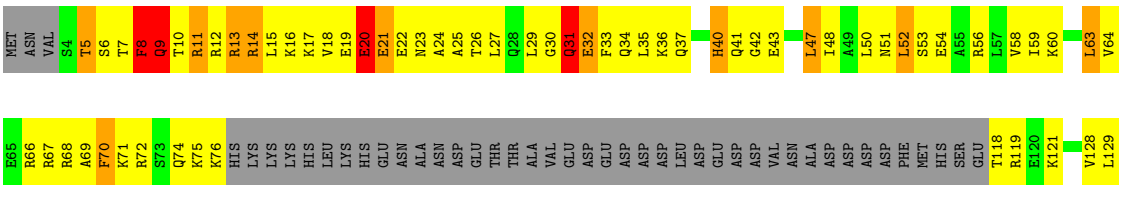


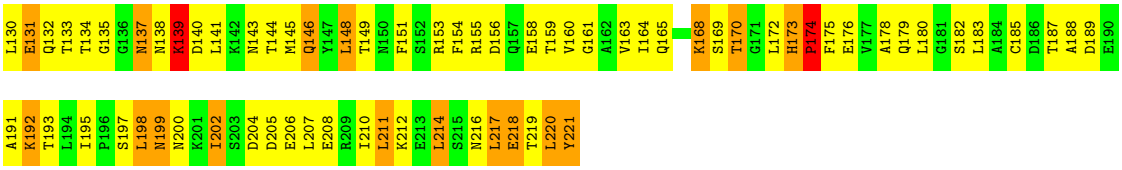


• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain E:

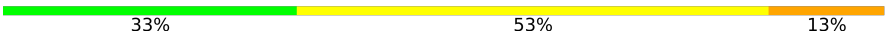


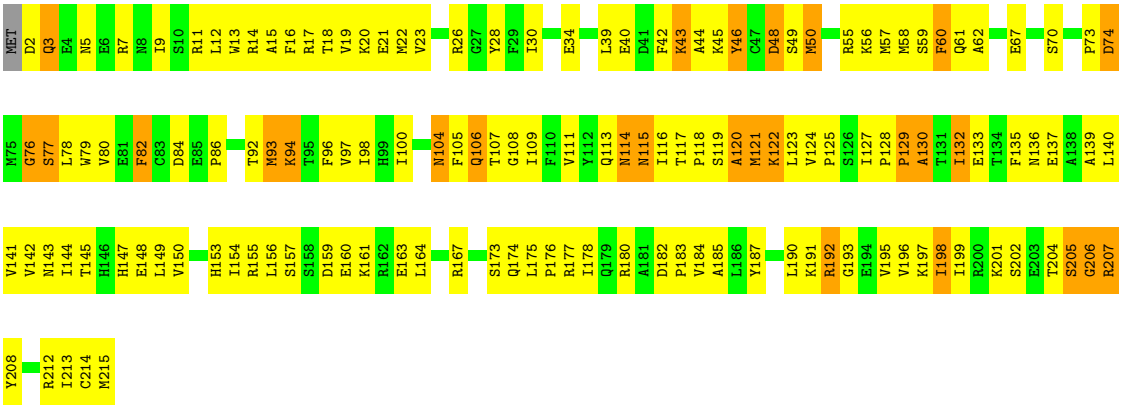




• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

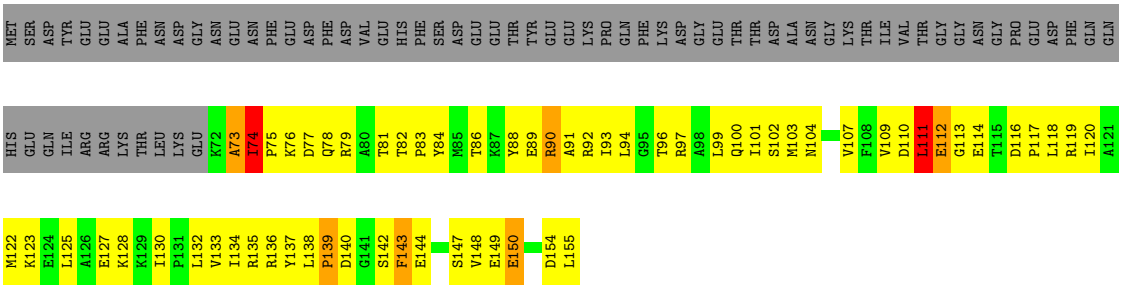
Chain F:





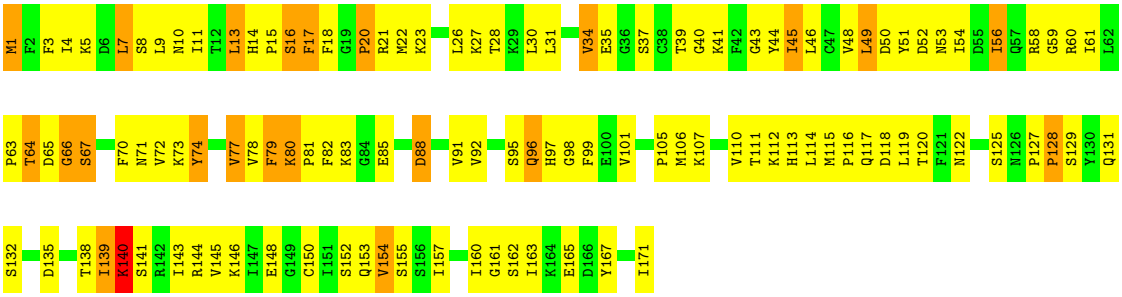
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain G:



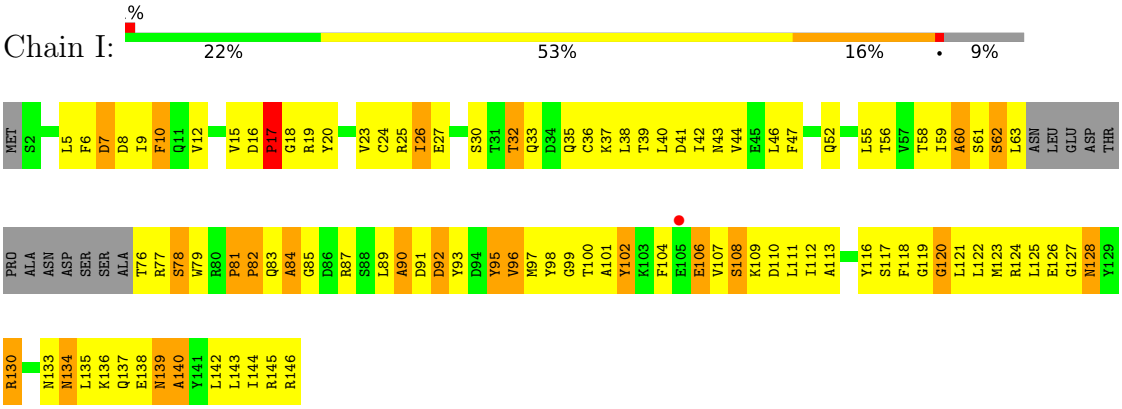
• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain H:



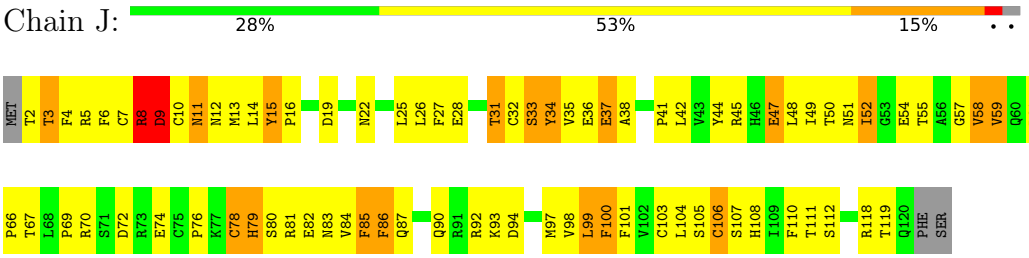
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain I:



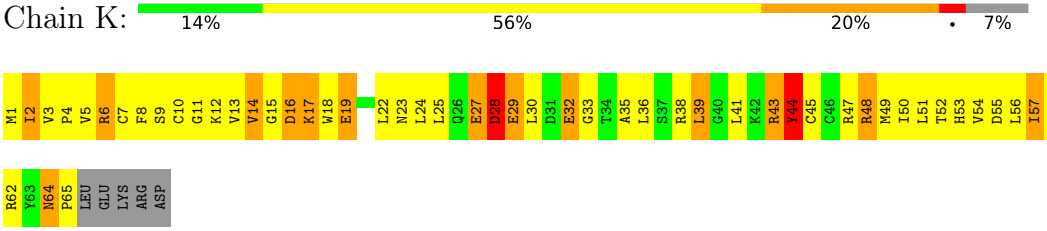
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain J:



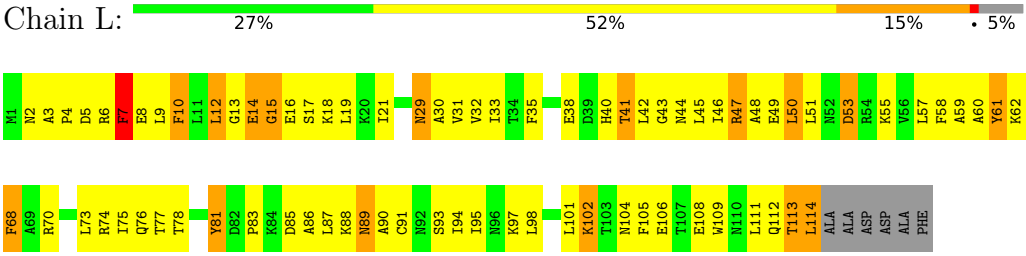
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain K:



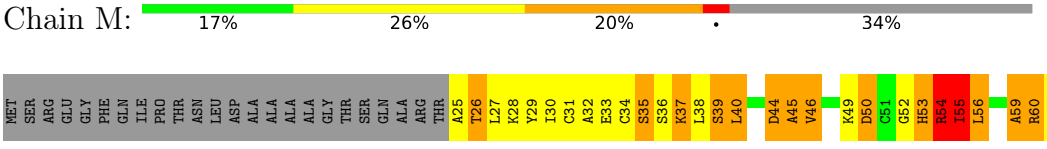
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

Chain L:



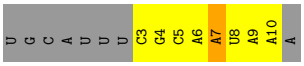
• Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4

Chain M:

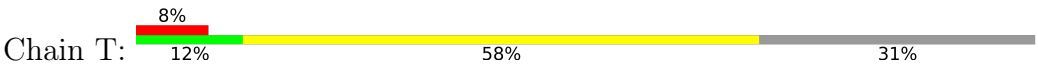


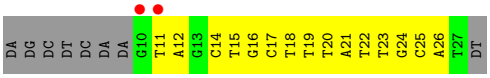
• Molecule 13: 5'-D(*AP*GP*CP*TP*CP*AP*AP*GP*TP*AP*GP*CP*TP*GP*CP*TP*TP*TP*AP*TP*TP*GP*CP*AP*TP*T)-3'

• Molecule 14: 5'-D(*CP*AP*GP*CP*TP*AP*CP*TP*TP*GP*AP*GP*CP*T)-3'



• Molecule 15: 5'-R(*UP*GP*CP*AP*UP*UP*UP*CP*GP*CP*AP*AP*UP*AP*AP*A)-3'





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	222.47Å 391.62Å 284.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 4.00 48.99 – 4.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-4.00) 100.0 (48.99-4.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	2.57 (at 4.00Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.205 , 0.241 0.215 , 0.243	Depositor DCC
R_{free} test set	16372 reflections (8.09%)	wwPDB-VP
Wilson B-factor (Å^2)	131.8	Xtriage
Anisotropy	0.506	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 82.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.024 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.024 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	31777	wwPDB-VP
Average B, all atoms (Å^2)	138.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.92% 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

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	B	0.43	0/11339	0.71	4/15334 (0.0%)
2	C	0.43	0/8981	0.69	1/12108 (0.0%)
3	D	0.43	0/2133	0.71	0/2891
4	E	0.43	0/1437	0.69	0/1925
5	F	0.42	0/1788	0.67	0/2406
6	G	0.48	0/691	0.77	0/933
7	H	0.47	0/1368	0.73	0/1844
8	I	0.41	0/1086	0.69	0/1470
9	J	0.40	0/989	0.66	0/1331
10	K	0.44	0/541	0.75	0/727
11	L	0.45	0/937	0.69	0/1265
12	M	0.48	0/366	0.72	0/485
13	N	0.70	0/154	0.88	0/235
14	P	0.55	0/188	0.94	0/291
15	T	0.42	0/407	0.95	0/627
All	All	0.44	0/32405	0.71	5/43872 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
10	K	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	311	GLN	N-CA-C	5.60	126.11	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	56	PRO	N-CA-C	-5.54	97.71	112.10
1	B	425	GLN	N-CA-C	-5.38	96.48	111.00
2	C	1163	CYS	N-CA-C	-5.21	96.94	111.00
1	B	440	ASP	N-CA-C	-5.02	97.44	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
10	K	44	TYR	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	B	11140	0	11215	1481	0
2	C	8810	0	8847	1266	0
3	D	2095	0	2051	295	0
4	E	1427	0	1451	175	0
5	F	1752	0	1776	183	0
6	G	679	0	701	98	0
7	H	1340	0	1357	185	0
8	I	1068	0	1040	166	0
9	J	971	0	929	123	0
10	K	532	0	542	122	0
11	L	919	0	929	135	0
12	M	364	0	388	57	0
13	N	138	0	80	8	0
14	P	168	0	88	15	0
15	T	365	0	208	48	0
16	B	2	0	0	0	0
16	C	1	0	0	0	0
16	D	1	0	0	0	0
16	J	2	0	0	0	0
16	K	1	0	0	0	0
16	M	1	0	0	0	0
17	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	31777	0	31602	3977	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 63.

The worst 5 of 3977 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)
15:T:21:DA:C2'	15:T:22:DT:H5'	1.58	1.32
15:T:21:DA:H2"	15:T:22:DT:C5'	1.65	1.25
15:T:20:DT:C2'	15:T:21:DA:H5'	1.73	1.18
11:L:47:ARG:HB3	11:L:47:ARG:HH11	1.11	1.15
15:T:20:DT:H2'	15:T:21:DA:H5'	1.13	1.12

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	1406/1733 (81%)	965 (69%)	284 (20%)	157 (11%)	0 7
2	C	1090/1224 (89%)	719 (66%)	243 (22%)	128 (12%)	0 6
3	D	264/318 (83%)	163 (62%)	70 (26%)	31 (12%)	0 6
4	E	173/221 (78%)	107 (62%)	43 (25%)	23 (13%)	0 4
5	F	212/215 (99%)	154 (73%)	36 (17%)	22 (10%)	0 8
6	G	82/155 (53%)	62 (76%)	13 (16%)	7 (8%)	1 12
7	H	169/171 (99%)	129 (76%)	26 (15%)	14 (8%)	1 13
8	I	129/146 (88%)	79 (61%)	30 (23%)	20 (16%)	0 3
9	J	117/122 (96%)	80 (68%)	24 (20%)	13 (11%)	0 7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
10	K	63/70 (90%)	38 (60%)	12 (19%)	13 (21%)	0 2
11	L	112/120 (93%)	81 (72%)	24 (21%)	7 (6%)	1 18
12	M	44/70 (63%)	22 (50%)	8 (18%)	14 (32%)	0 0
All	All	3861/4565 (85%)	2599 (67%)	813 (21%)	449 (12%)	0 6

5 of 449 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	4	GLN
1	B	41	MET
1	B	48	ALA
1	B	54	ASN
1	B	57	ARG

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	1239/1520 (82%)	1113 (90%)	126 (10%)	7 28
2	C	962/1061 (91%)	867 (90%)	95 (10%)	8 29
3	D	234/274 (85%)	209 (89%)	25 (11%)	6 27
4	E	159/200 (80%)	136 (86%)	23 (14%)	3 18
5	F	196/197 (100%)	182 (93%)	14 (7%)	14 42
6	G	74/137 (54%)	69 (93%)	5 (7%)	16 44
7	H	152/152 (100%)	136 (90%)	16 (10%)	7 27
8	I	117/128 (91%)	107 (92%)	10 (8%)	10 37
9	J	113/116 (97%)	103 (91%)	10 (9%)	10 35
10	K	60/65 (92%)	54 (90%)	6 (10%)	7 29
11	L	99/102 (97%)	84 (85%)	15 (15%)	3 16
12	M	40/57 (70%)	36 (90%)	4 (10%)	7 29
All	All	3445/4009 (86%)	3096 (90%)	349 (10%)	7 29

5 of 349 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	104	PHE
6	G	111	LEU
3	D	190	ASP
4	E	137	ASN
7	H	140	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 106 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	770	GLN
2	C	1179	GLN
9	J	108	HIS
2	C	821	GLN
2	C	1074	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	7/16 (43%)	1 (14%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
14	P	7	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	B	1416/1733 (81%)	-0.37	4 (0%) 94 90	76, 129, 175, 200	0
2	C	1108/1224 (90%)	-0.32	2 (0%) 95 93	79, 140, 185, 200	0
3	D	266/318 (83%)	-0.41	0 100 100	93, 126, 168, 188	0
4	E	177/221 (80%)	-0.34	0 100 100	106, 142, 182, 190	0
5	F	214/215 (99%)	-0.33	0 100 100	99, 160, 187, 200	0
6	G	84/155 (54%)	-0.52	0 100 100	73, 107, 138, 147	0
7	H	171/171 (100%)	-0.34	0 100 100	98, 127, 164, 174	0
8	I	133/146 (91%)	0.02	1 (0%) 86 79	134, 165, 186, 196	0
9	J	119/122 (97%)	-0.31	0 100 100	122, 165, 189, 200	0
10	K	65/70 (92%)	-0.59	0 100 100	92, 121, 158, 169	0
11	L	114/120 (95%)	-0.39	0 100 100	94, 128, 156, 170	0
12	M	46/70 (65%)	-0.03	0 100 100	120, 172, 195, 198	0
13	N	7/14 (50%)	1.23	3 (42%) 0 0	199, 200, 200, 200	1 (14%)
14	P	8/16 (50%)	0.06	0 100 100	198, 199, 200, 200	0
15	T	18/26 (69%)	0.65	2 (11%) 5 5	178, 199, 200, 200	1 (5%)
All	All	3946/4621 (85%)	-0.33	12 (0%) 94 90	73, 136, 184, 200	2 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1176	LEU	3.9
13	N	7	DC	2.9
15	T	11	DT	2.8
2	C	471	LYS	2.8
1	B	1455	PRO	2.3

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 monosaccharides in this entry.

6.4 Ligands (i)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
17	MG	B	1736	1/1	0.86	0.10	153,153,153,153	0
16	ZN	J	124	1/1	0.92	0.03	200,200,200,200	0
16	ZN	M	71	1/1	0.98	0.05	174,174,174,174	0
16	ZN	D	319	1/1	0.99	0.13	98,98,98,98	0
16	ZN	J	123	1/1	0.99	0.16	136,136,136,136	0
16	ZN	B	1734	1/1	0.99	0.04	149,149,149,149	0
16	ZN	K	71	1/1	0.99	0.25	113,113,113,113	0
16	ZN	B	1735	1/1	0.99	0.12	110,110,110,110	0
16	ZN	C	1225	1/1	0.99	0.23	104,104,104,104	0

6.5 Other polymers (i)

There are no such residues in this entry.