wwPDB X-ray Structure Validation Summary Report (i)

Aug 15, 2023 – 04:18 AM EDT

PDB ID	:	1TWA
Title	:	RNA polymerase II complexed with ATP
Authors	:	Westover, K.D.; Bushnell, D.A.; Kornberg, R.D.
Deposited on	:	2004-06-30
Resolution	:	3.20 Å(reported)

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

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

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

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

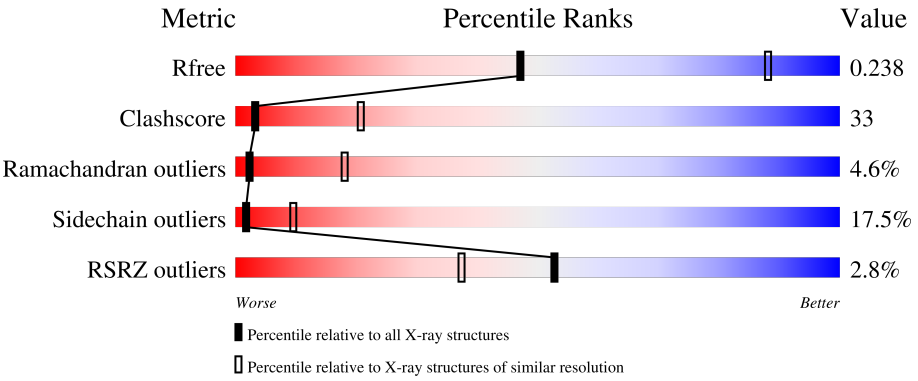
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 3.20 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.



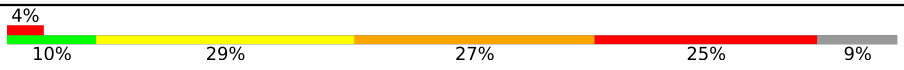
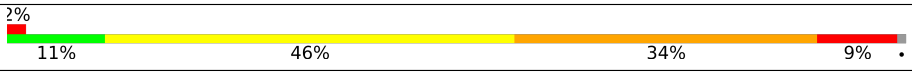
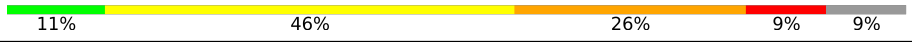
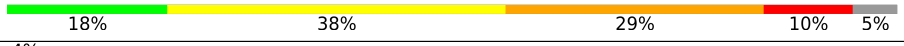
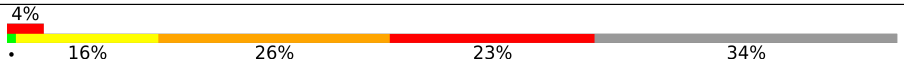
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	1733	2% 16% 35% 20% 8% 22%
2	B	1224	4% 18% 40% 23% 8% 11%
3	C	318	% 13% 38% 25% 8% 16%
4	E	215	2% 18% 41% 30% 11%
5	F	155	% 11% 21% 16% 5% 46%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	H	146	
7	I	122	
8	J	70	
9	K	120	
10	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
12	ZN	C	3002	-	-	X	-
12	ZN	J	3001	-	-	X	-

2 Entry composition (i)

There are 13 unique types of molecules in this entry. The entry contains 27757 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 largest subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1352	10635	6711	1842	2024	58	0	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1091	8690	5511	1516	1610	53	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

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

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	215	1760	1116	310	322	12	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	F	83	670	428	114	125	3	0	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II 14.2 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	I	121	Total	C	N	O	S	0	0	0
			990	610	181	188	11			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	J	64	Total	C	N	O	S	0	0	0
			525	334	92	93	6			

- Molecule 9 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

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

- Molecule 11 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	2	Total	Mn	0	0
			2	2		

- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

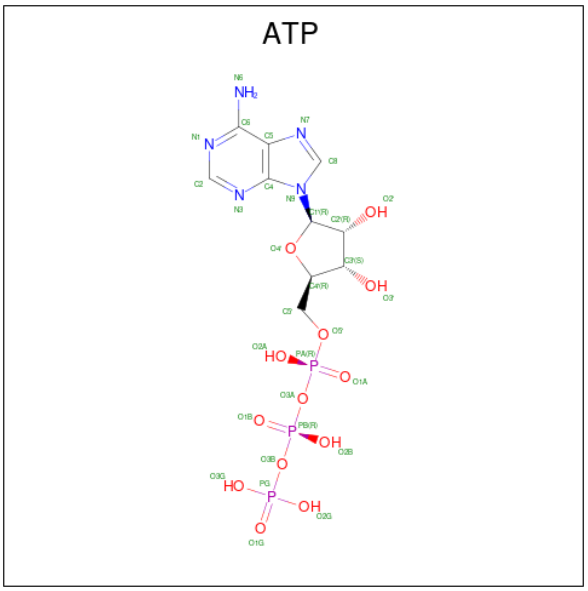
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	2	Total	Zn	0	0
			2	2		
12	B	1	Total	Zn	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	C	1	Total Zn 1 1	0	0
12	I	2	Total Zn 2 2	0	0
12	J	1	Total Zn 1 1	0	0
12	L	1	Total Zn 1 1	0	0

- Molecule 13 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



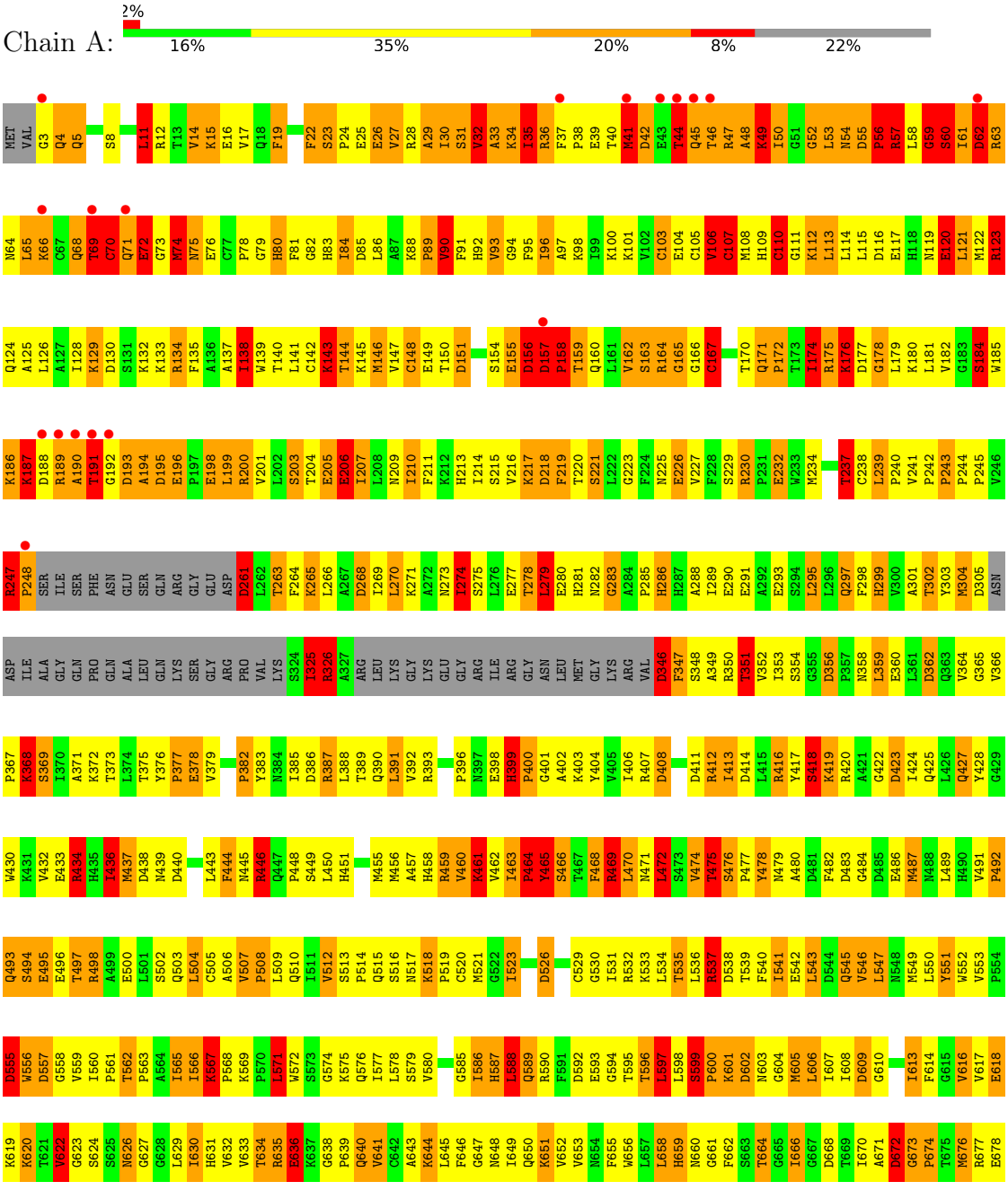
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	A	1	Total C N O P 31 10 5 13 3	0	0

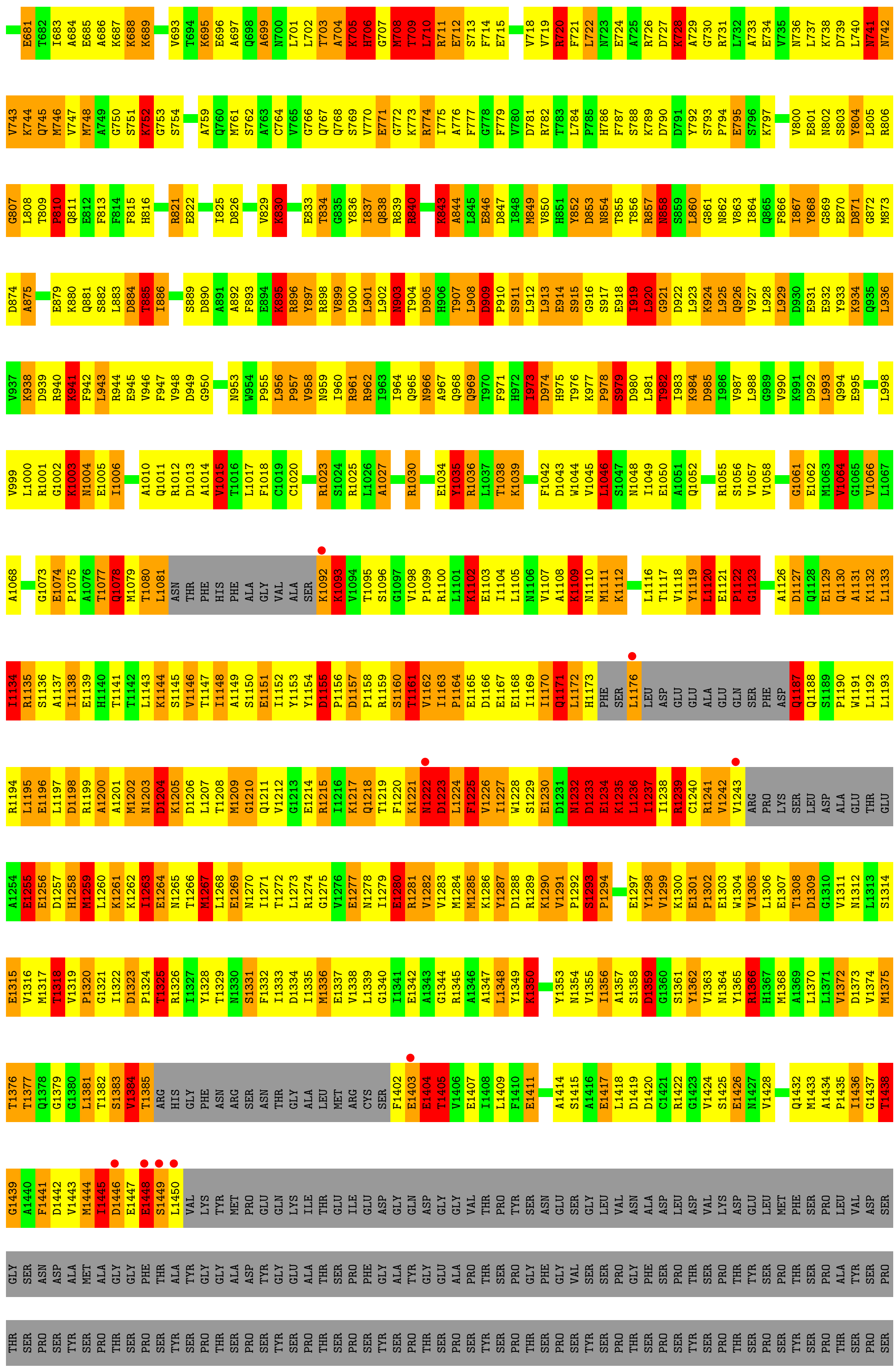
3 Residue-property plots (i)

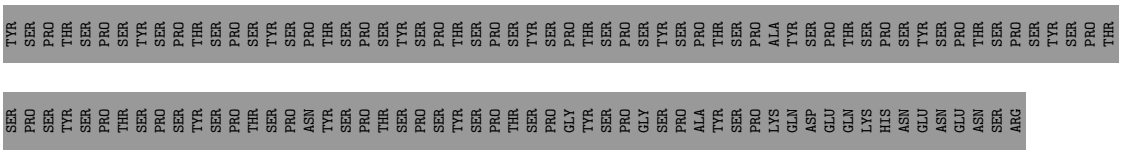
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

• Molecule 1: DNA-directed RNA polymerase II largest subunit

Chain A:

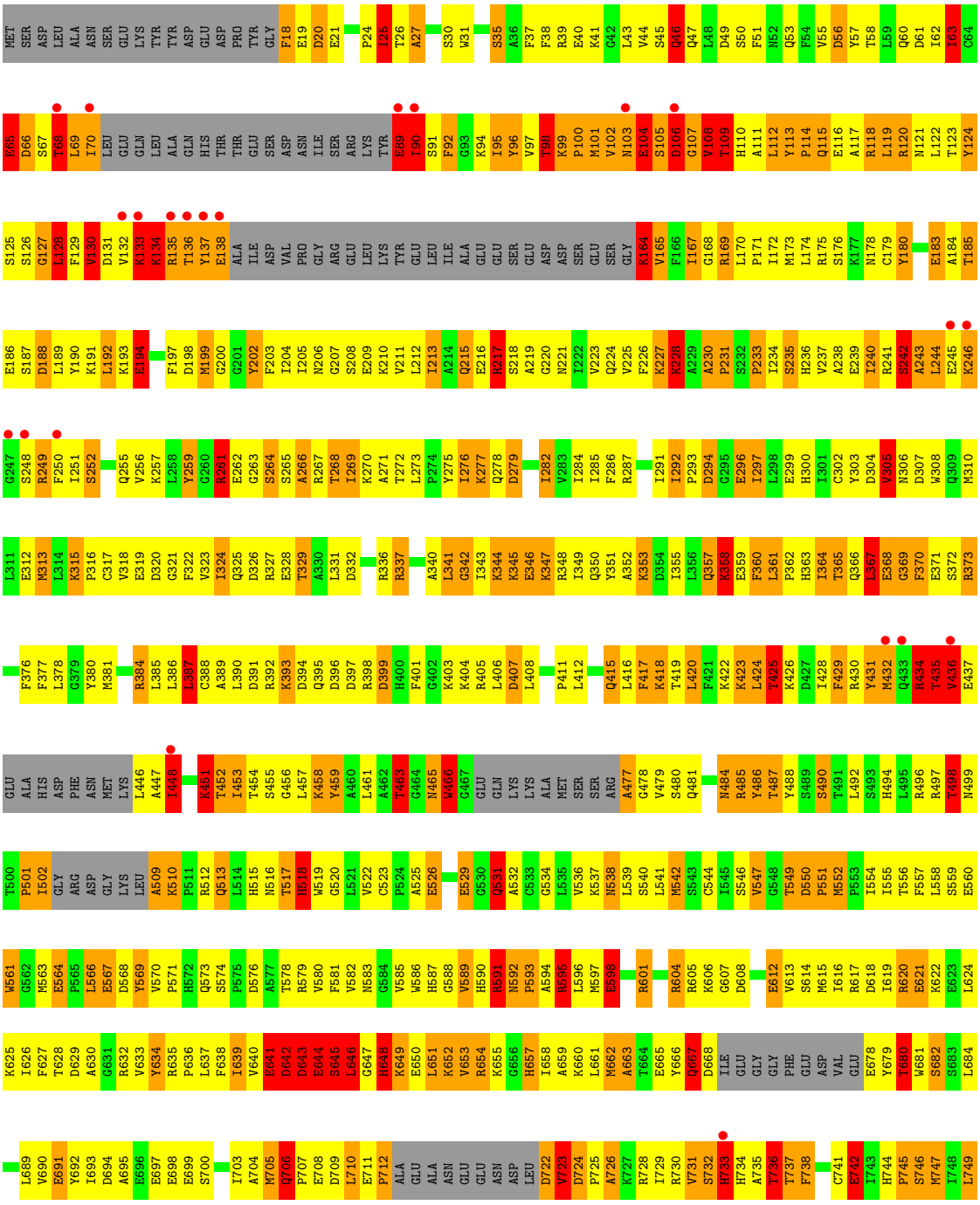


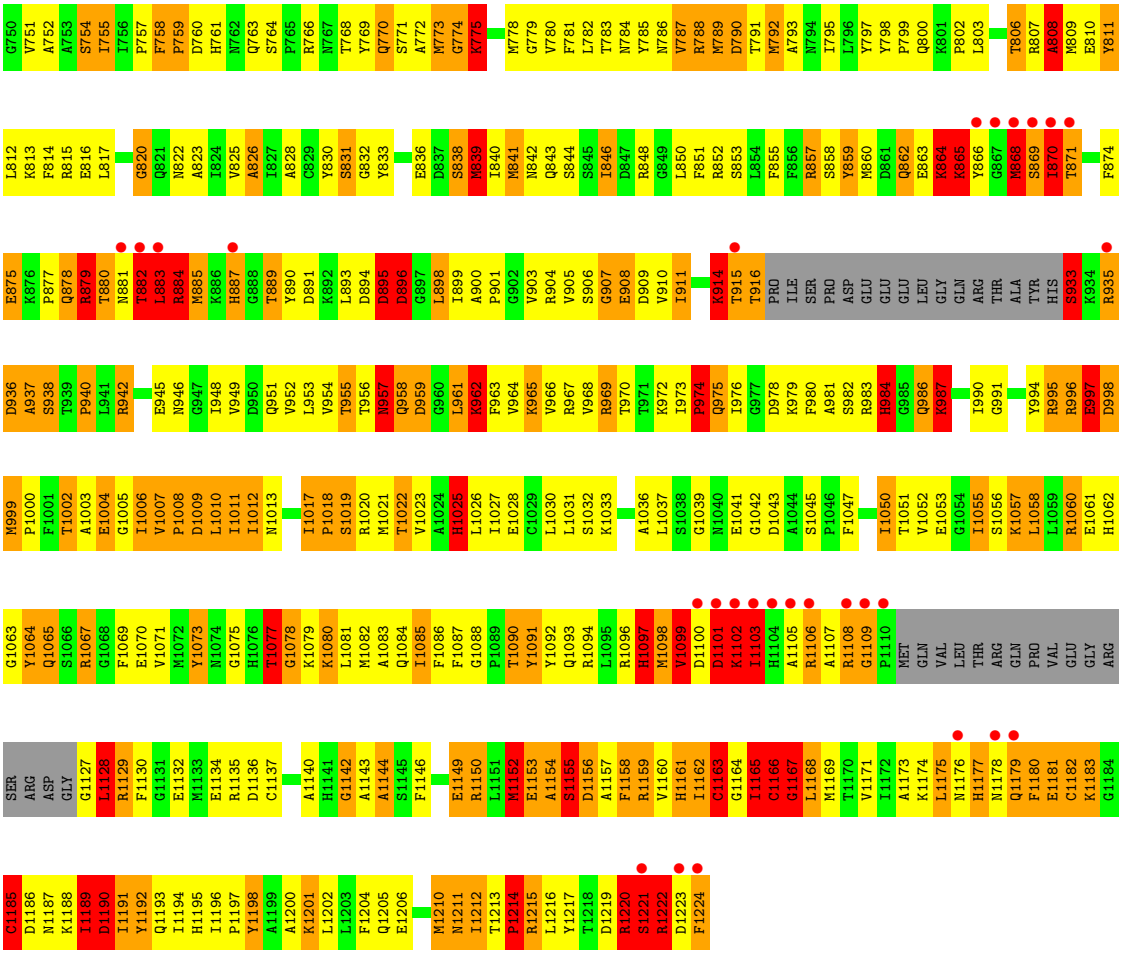

E681 T682 I683 A684 E685 A686 K687 K688 K689 V693 T694 K695 E696 A697 Q698 A699 N700 L701 L702 T703 A704 K705 H706 G707 M708 T709 L710 R711 E712 S713 F714 E715 V718 V719 R720 F721 L722 N723 E724 A725 R726 D727 K728 A729 G730 R731 L732 A733 E734 V735 N736 L737 K738 D739 L740 N741 N742
V743 K744 Q745 M746 V747 M748 A749 G750 S751 K752 G753 S754 A759 Q760 M761 S762 A763 C764 V765 G766 Q767 Q768 S769 V770 E771 G772 K773 R774 I775 A776 F777 G778 F779 V780 D781 R782 T783 L784 P785 H786 F787 S788 K789 D790 D791 Y792 S793 P794 E795 S796 K797 V800 E801 N802 S803 Y804 L805 R806
G807 L808 T809 P810 Q811 E812 F813 F814 F815 H816 R821 E822 I825 D826 V829 K830 E833 T834 G835 Y836 I837 Q838 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 E879 K880 Q881 S882 L883 D884 T885 I886 S889 D890 A891 A892 F893 E894 K895 R896 Y897 R898 V899 D900 L901 L902 N903 T904 D905 H906 T907 L908 D909 P910 S911 L912 L913 E914 S915 G916 S917 E918 I919 L920 G921 D922 L923 K924 L925 Q926 V927 L928 L929 D930 E931 E932 Y933 K934 Q935 L936
V937 K938 D939 R940 K941 F942 L943 R944 E945 V946 F947 V948 D949 G950 N953 W954 P955 L956 P957 V958 N959 I960 R961 R962 I963 I964 Q965 N966 A967 Q968 Q969 T970 F971 H972 I973 D974 H975 T976 K977 P978 S979 D980 L981 T982 I983 K984 D985 I986 V987 L988 G989 V990 K991 D992 L993 Q994 E995 L998
V999 L1000 R1001 G1002 K1003 N1004 E1005 I1006 A1010 Q1011 R1012 D1013 A1014 V1015 T1016 L1017 F1018 C1019 C1020 R1023 S1024 R1025 L1026 A1027 R1030 E1034 Y1035 R1036 L1037 T1038 K1039 F1042 D1043 W1044 V1045 L1046 S1047 N1048 I1049 E1050 A1051 Q1052 R1055 S1056 V1057 V1058 G1061 E1062 M1063 V1064 G1065 V1066 L1067
A1068 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 V1098 P1099 R1100 L1101 K1102 E1103 I1104 L1105 N1106 V1107 A1108 K1109 N1110 M1111 K1112 L1116 T1117 V1118 Y1119 L1120 E1121 P1122 G1123 A1126 D1127 Q1128 E1129 Q1130 A1131 K1132 L1133
I1134 R1135 S1136 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 E1168 I1169 I1170 Q1171 L1172 H1173 PHE SER L1176 LEU ASP GLU GLU ALA GLU GLN SER PHE ASP Q1187 Q1188 S1189 P1190 W1191 L1192 L1193
R1194 L1195 E1196 L1197 D1198 R1199 A1200 A1201 M1202 N1203 D1204 K1205 D1206 L1207 T1208 M1209 G1210 Q1211 V1212 G1213 E1214 R1215 I1216 K1217 Q1218 T1219 F1220 K1221 N1222 D1223 L1224 F1225 V1226 I1227 W1228 S1229 E1230 D1231 N1232 D1233 E1234 K1235 L1236 I1237 I1238 R1239 C1240 R1241 V1242 V1243 ARG PRO LYS SER LEU ASP ALA GLU THR GLU
A1254 E1255 E1256 D1257 H1258 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 M1285 K1286 Y1287 D1288 R1289 K1290 V1291 P1292 S1293 P1294 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 Y1353 N1354 V1355 I1356 A1357 S1358 D1359 G1360 S1361 Y1362 V1363 N1364 Y1365 R1366 H1367 M1368 A1369 L1370 L1371 V1372 D1373 V1374 M1375
T1376 T1377 Q1378 G1379 G1380 L1381 T1382 S1383 V1384 T1385 ARG HIS GLY PHE ASN ARG SER ASN THR GLY ALA LEU MET ARG CYS SER F1402 E1403 E1404 T1405 V1406 E1407 I1408 L1409 F1410 E1411 A1414 S1415 A1416 E1417 L1418 D1419 D1420 C1421 R1422 G1423 V1424 S1425 E1426 N1427 V1428 Q1432 M1433 A1434 P1435 I1436 G1437 T1438
G1439 A1440 F1441 D1442 V1443 M1444 I1445 D1446 E1447 E1448 S1449 L1450 VAL LYS TYR MET PRO GLU GLN LYS ILE THR GLU ILE GLU ASP GLY GLN ASP GLY GLY VAL THR PRO TYR SER ASN GLU SER GLY LEU VAL ASN ALA ASP LEU ASP VAL LYS ASP GLU LEU MET PHE SER PRO LEU VAL ASP SER
GLY SER ASN ASP ALA MET ALA GLY GLY PHE THR ALA TYR GLY GLY ALA ASP TYR GLY GLU ALA THR SER PRO PHE GLY ALA TYR GLY GLU ALA PRO THR SER PRO GLY PHE GLY VAL SER SER PRO GLY PHE SER PRO THR SER PRO THR TYR SER PRO THR SER PRO ALA 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



• Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide

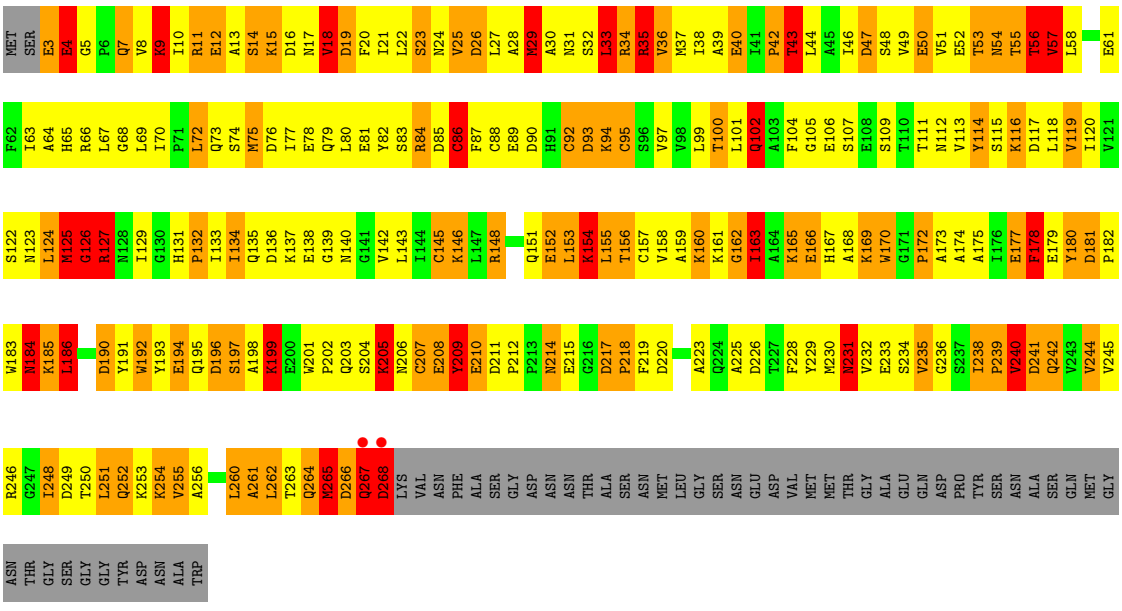
Chain B:





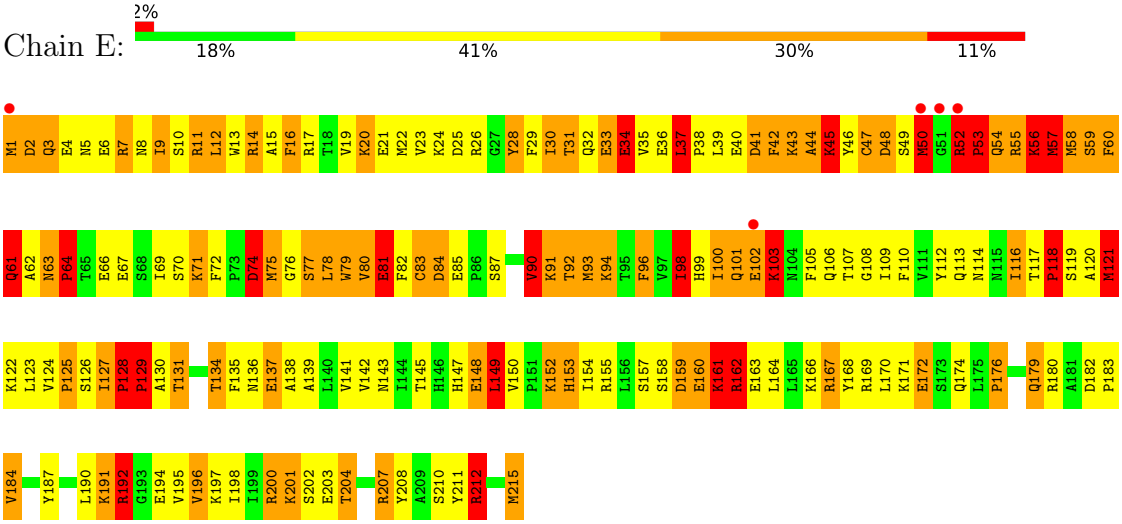
• Molecule 3: DNA-directed RNA polymerase II 45 kDa polypeptide

Chain C:



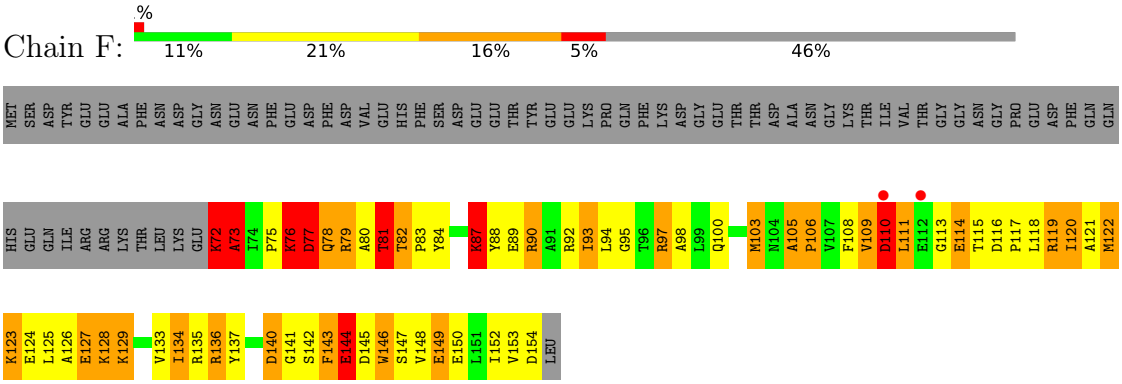
• Molecule 4: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

Chain E:



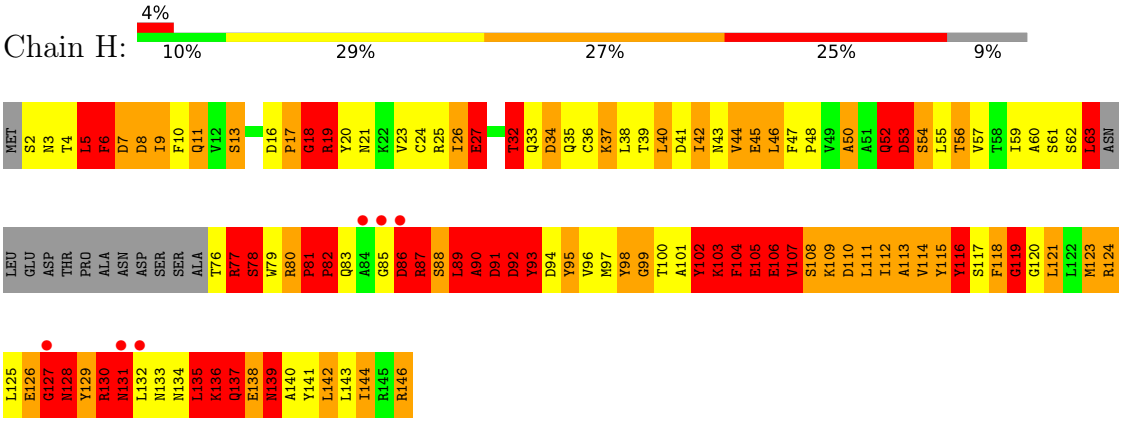
• Molecule 5: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

Chain F:



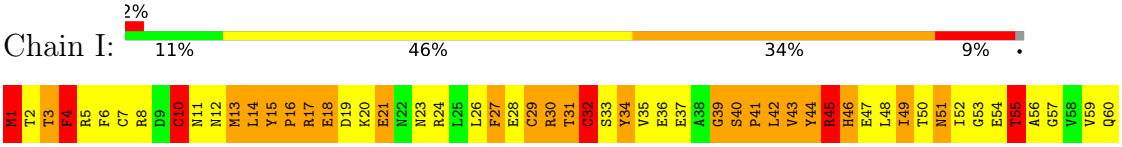
• Molecule 6: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

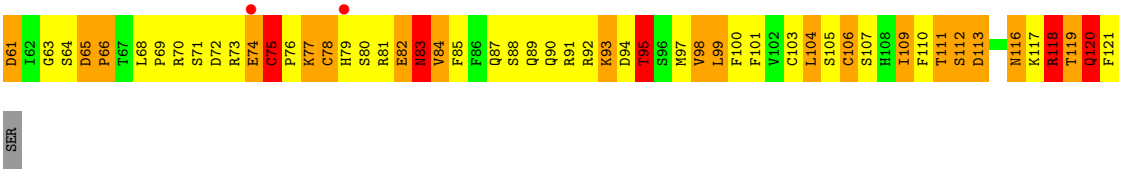
Chain H:



• Molecule 7: DNA-directed RNA polymerase II 14.2 kDa polypeptide

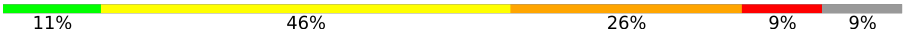
Chain I:

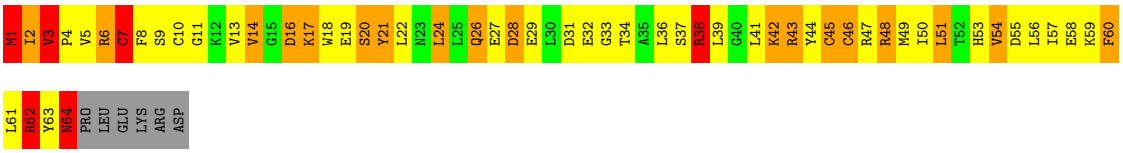




• Molecule 8: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide

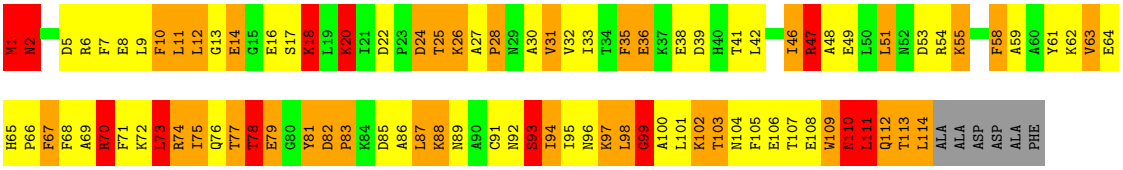
Chain J:





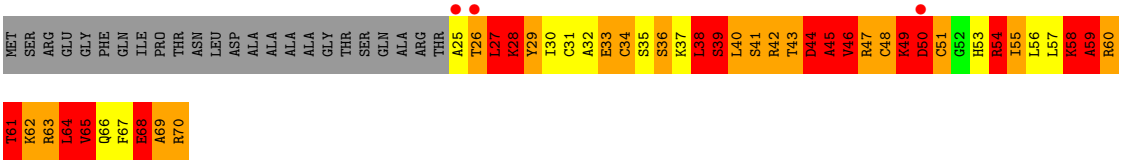
• Molecule 9: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K:



• Molecule 10: 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	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	123.00Å 223.00Å 374.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.20 39.69 – 3.10	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.20) 97.5 (39.69-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.55 (at 3.12Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.227 , 0.246 0.195 , 0.238	Depositor DCC
R_{free} test set	2779 reflections (2.97%)	wwPDB-VP
Wilson B-factor (Å^2)	55.9	Xtriage
Anisotropy	0.349	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 42.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	27757	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.89% 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: ATP, MN, ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	3.53	1385/10822 (12.8%)	2.62	715/14641 (4.9%)
2	B	3.54	1164/8860 (13.1%)	2.56	552/11945 (4.6%)
3	C	3.57	288/2133 (13.5%)	2.62	126/2891 (4.4%)
4	E	3.50	210/1796 (11.7%)	2.40	107/2416 (4.4%)
5	F	3.12	78/682 (11.4%)	2.30	35/922 (3.8%)
6	H	3.37	138/1086 (12.7%)	2.46	68/1470 (4.6%)
7	I	3.73	145/1009 (14.4%)	2.69	86/1357 (6.3%)
8	J	3.41	75/533 (14.1%)	3.10	53/715 (7.4%)
9	K	3.36	117/937 (12.5%)	2.65	61/1265 (4.8%)
10	L	3.99	58/366 (15.8%)	2.90	43/485 (8.9%)
All	All	3.52	3658/28224 (13.0%)	2.59	1846/38107 (4.8%)

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

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	41
2	B	0	40
3	C	0	13
4	E	1	6
5	F	0	2
6	H	0	12
7	I	0	6
8	J	0	1
9	K	0	1
10	L	0	2
All	All	1	124

The worst 5 of 3658 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	552	MET	CG-SD	26.10	2.49	1.81
1	A	728	LYS	CD-CE	24.04	2.11	1.51
2	B	598	GLU	CG-CD	23.97	1.88	1.51
1	A	771	GLU	CD-OE2	23.53	1.51	1.25
3	C	165	LYS	CE-NZ	23.14	2.06	1.49

The worst 5 of 1846 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	774	ARG	NE-CZ-NH1	34.28	137.44	120.30
1	A	469	ARG	NE-CZ-NH2	-32.36	104.12	120.30
1	A	1366	ARG	NE-CZ-NH1	31.49	136.04	120.30
3	C	35	ARG	NE-CZ-NH2	-28.02	106.29	120.30
3	C	34	ARG	NE-CZ-NH2	-27.97	106.32	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	E	204	THR	CB

5 of 124 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	158	PRO	Peptide
1	A	165	GLY	Peptide
1	A	44	THR	Peptide
1	A	60	SER	Peptide
1	A	74	MET	Peptide

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	10635	0	10688	689	0
2	B	8690	0	8715	513	0
3	C	2095	0	2053	155	0
4	E	1760	0	1788	130	0
5	F	670	0	689	44	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	H	1068	0	1040	164	0
7	I	990	0	949	71	0
8	J	525	0	537	44	0
9	K	919	0	929	73	0
10	L	364	0	388	65	0
11	A	2	0	0	0	0
12	A	2	0	0	1	0
12	B	1	0	0	1	0
12	C	1	0	0	2	0
12	I	2	0	0	0	0
12	J	1	0	0	2	0
12	L	1	0	0	1	0
13	A	31	0	12	0	0
All	All	27757	0	27788	1823	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 33.

The worst 5 of 1823 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)
2:B:1057:LYS:CE	2:B:1057:LYS:CD	1.75	1.65
1:A:368:LYS:CD	1:A:368:LYS:CE	1.75	1.65
1:A:919:ILE:CD1	1:A:919:ILE:CG1	1.75	1.64
4:E:37:LEU:CD1	4:E:37:LEU:CG	1.75	1.64
1:A:1112:LYS:CD	1:A:1112:LYS:CE	1.74	1.64

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	1335/1733 (77%)	1187 (89%)	91 (7%)	57 (4%)	2 20
2	B	1071/1224 (88%)	924 (86%)	104 (10%)	43 (4%)	3 21
3	C	264/318 (83%)	231 (88%)	23 (9%)	10 (4%)	3 22
4	E	213/215 (99%)	183 (86%)	16 (8%)	14 (7%)	1 9
5	F	81/155 (52%)	73 (90%)	6 (7%)	2 (2%)	5 32
6	H	129/146 (88%)	89 (69%)	19 (15%)	21 (16%)	0 0
7	I	119/122 (98%)	110 (92%)	8 (7%)	1 (1%)	19 58
8	J	62/70 (89%)	58 (94%)	4 (6%)	0	100 100
9	K	112/120 (93%)	98 (88%)	12 (11%)	2 (2%)	8 41
10	L	44/70 (63%)	24 (54%)	13 (30%)	7 (16%)	0 0
All	All	3430/4173 (82%)	2977 (87%)	296 (9%)	157 (5%)	2 18

5 of 157 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	ILE
1	A	45	GLN
1	A	47	ARG
1	A	59	GLY
1	A	60	SER

5.3.2 Protein sidechains (i)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	1184/1520 (78%)	998 (84%)	186 (16%)	2 12
2	B	947/1061 (89%)	806 (85%)	141 (15%)	3 14
3	C	234/274 (85%)	194 (83%)	40 (17%)	2 10
4	E	197/197 (100%)	152 (77%)	45 (23%)	1 4
5	F	73/137 (53%)	64 (88%)	9 (12%)	4 21
6	H	117/128 (91%)	73 (62%)	44 (38%)	0 0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
7	I	115/116 (99%)	91 (79%)	24 (21%)	1 6
8	J	59/65 (91%)	49 (83%)	10 (17%)	2 10
9	K	99/102 (97%)	80 (81%)	19 (19%)	1 8
10	L	40/57 (70%)	21 (52%)	19 (48%)	0 0
All	All	3065/3657 (84%)	2528 (82%)	537 (18%)	2 9

5 of 537 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	H	144	ILE
7	I	61	ASP
6	H	137	GLN
10	L	33	GLU
2	B	89	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 71 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	167	HIS
3	C	252	GLN
7	I	46	HIS
1	A	1364	ASN
1	A	1173	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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 11 ligands modelled in this entry, 10 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
13	ATP	A	3011	11	26,33,33	1.14	2 (7%)	31,52,52	1.57	4 (12%)

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
13	ATP	A	3011	11	-	2/18/38/38	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	A	3011	ATP	C2-N3	4.13	1.38	1.32
13	A	3011	ATP	C2-N1	2.67	1.38	1.33

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	A	3011	ATP	N3-C2-N1	-5.48	120.11	128.68
13	A	3011	ATP	PB-O3B-PG	-3.45	120.99	132.83
13	A	3011	ATP	PA-O3A-PB	-3.40	121.17	132.83
13	A	3011	ATP	C3'-C2'-C1'	2.89	105.32	100.98

There are no chirality outliers.

All (2) torsion outliers are listed below:

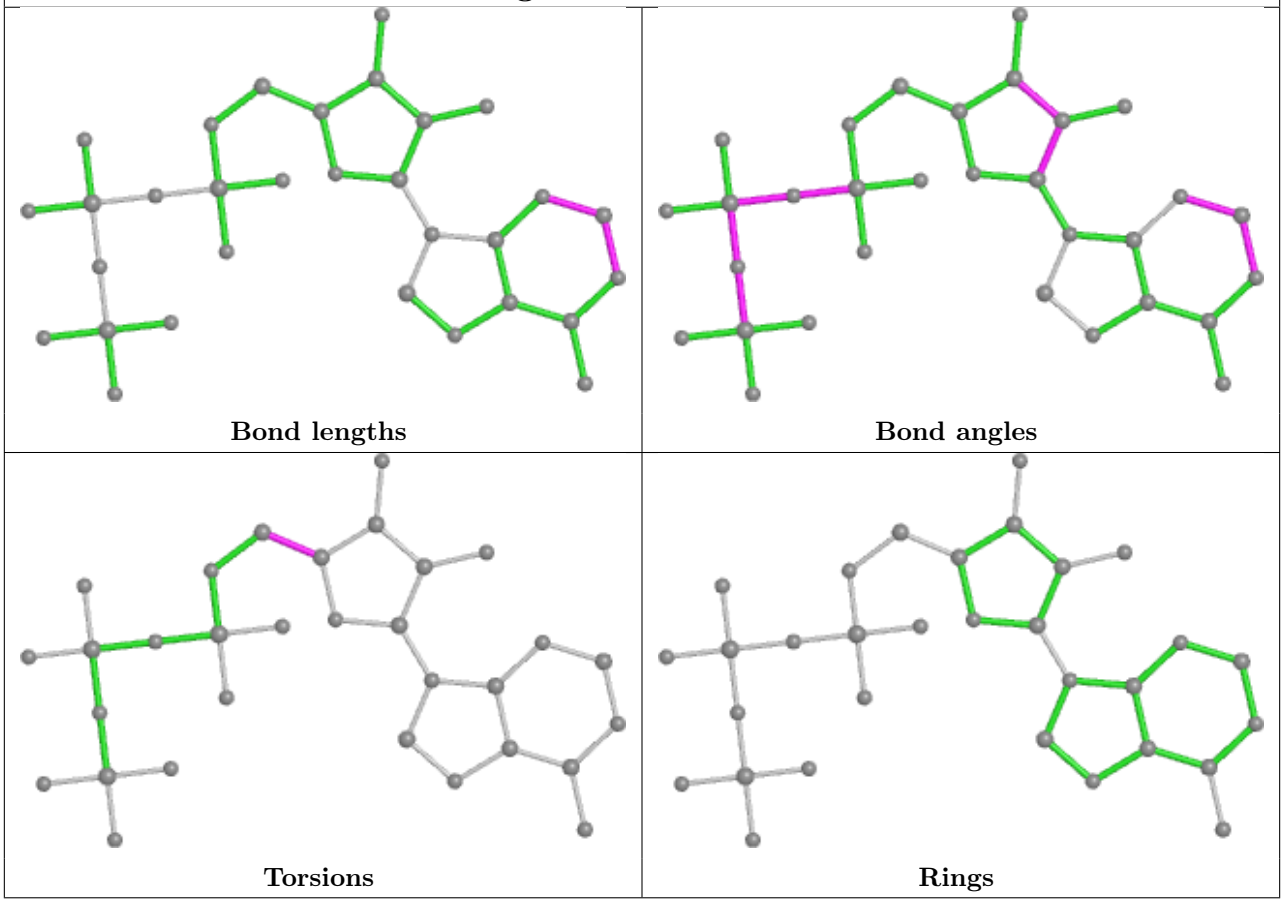
Mol	Chain	Res	Type	Atoms
13	A	3011	ATP	O4'-C4'-C5'-O5'
13	A	3011	ATP	C3'-C4'-C5'-O5'

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

Ligand ATP A 3011



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
1	A	5
2	B	1

The worst 5 of 6 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	686:ALA	C	687:LYS	N	1.20
1	A	1064:VAL	C	1065:GLY	N	1.20
1	B	632:ARG	C	633:VAL	N	1.20
1	A	366:VAL	C	367:PRO	N	1.19
1	A	1195:LEU	C	1196:GLU	N	1.19

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	1352/1733 (78%)	-0.36	27 (1%) 65 51	9, 38, 105, 151	0
2	B	1091/1224 (89%)	-0.26	50 (4%) 32 20	10, 36, 111, 138	0
3	C	266/318 (83%)	-0.45	2 (0%) 86 78	21, 42, 72, 123	0
4	E	215/215 (100%)	-0.25	5 (2%) 60 47	13, 51, 102, 134	0
5	F	83/155 (53%)	-0.37	2 (2%) 59 44	18, 36, 64, 73	0
6	H	133/146 (91%)	0.21	6 (4%) 33 21	41, 79, 121, 132	0
7	I	121/122 (99%)	-0.21	2 (1%) 70 57	23, 44, 78, 108	0
8	J	64/70 (91%)	-0.55	0 100 100	21, 35, 61, 76	0
9	K	114/120 (95%)	-0.30	0 100 100	21, 50, 72, 81	0
10	L	46/70 (65%)	0.26	3 (6%) 18 11	39, 89, 118, 121	0
All	All	3485/4173 (83%)	-0.30	97 (2%) 53 37	9, 41, 107, 151	0

The worst 5 of 97 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1109	GLY	9.4
1	A	69	THR	7.7
2	B	1110	PRO	7.4
1	A	248	PRO	7.1
6	H	85	GLY	6.8

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

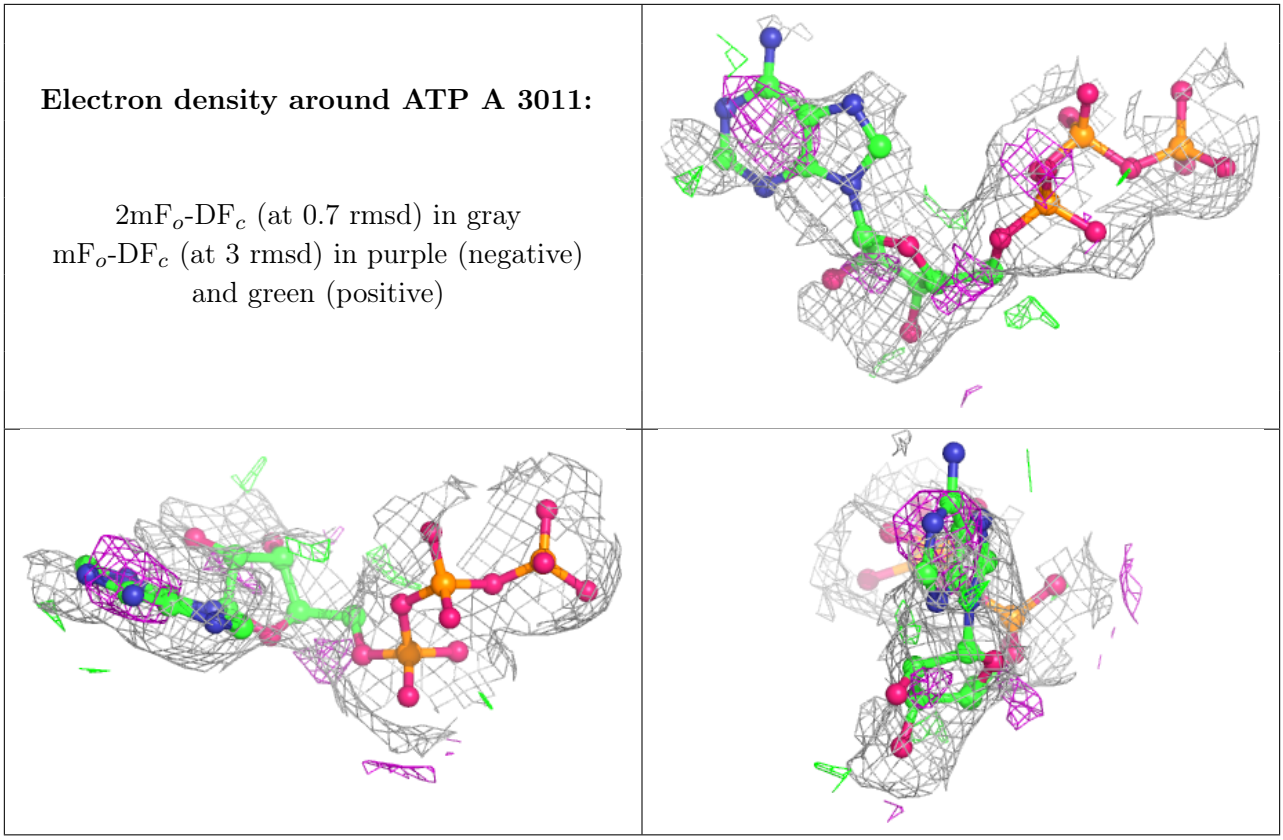
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
12	ZN	L	3005	1/1	0.69	0.06	139,139,139,139	0
12	ZN	I	3003	1/1	0.87	0.06	144,144,144,144	0
13	ATP	A	3011	31/31	0.89	0.23	69,74,83,83	0
12	ZN	B	3007	1/1	0.93	0.06	108,108,108,108	0
12	ZN	A	3006	1/1	0.96	0.04	124,124,124,124	0
12	ZN	I	3004	1/1	0.97	0.05	144,144,144,144	0
11	MN	A	3010	1/1	0.97	0.12	29,29,29,29	0
12	ZN	A	3008	1/1	0.97	0.03	96,96,96,96	0
11	MN	A	3009	1/1	0.98	0.12	24,24,24,24	0
12	ZN	C	3002	1/1	0.98	0.06	78,78,78,78	0
12	ZN	J	3001	1/1	0.99	0.08	84,84,84,84	0

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

Electron density around ATP A 3011:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.