wwPDB X-ray Structure Validation Summary Report (i)

Aug 15, 2023 – 04:12 AM EDT

PDB ID	:	1TWH
Title	:	RNA polymerase II complexed with 2'dATP
Authors	:	Westover, K.D.; Bushnell, D.A.; Kornberg, R.D.
Deposited on	:	2004-06-30
Resolution	:	3.40 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:

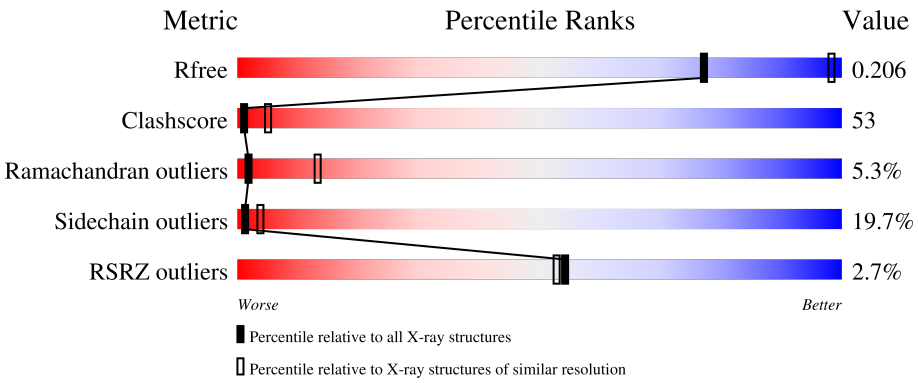
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 3.40 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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%; 7%, 32%, 28%, 11%, 22%
2	B	1224	4%; 6%, 37%, 34%, 12%, 11%
3	C	318	7%, 30%, 34%, 14%, 16%
4	E	215	2%; 5%, 35%, 40%, 20%
5	F	155	%; •, 22%, 23%, 5%, 46%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	H	146	5% • 19% 36% 34% 9%
7	I	122	3% 5% 40% 38% 16% •
8	J	70	7% 39% 31% 14% 9%
9	K	120	• 35% 45% 11% 5%
10	L	70	3% • 7% 30% 27% 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
12	ZN	C	3002	-	-	X	-
12	ZN	J	3001	-	-	X	-
13	ATP	A	3011	X	-	-	-

2 Entry composition (i)

There are 14 unique types of molecules in this entry. The entry contains 27728 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	1349	10606	6692	1839	2017	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
			Total	C	N	O	S			
6	H	133	1068	673	180	211	4	0	0	0

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

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

- 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
			Total	C	N	O	S			
8	J	64	525	334	92	93	6	0	0	0

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

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

- 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
			Total	C	N	O	S			
10	L	46	364	224	72	64	4	0	0	0

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

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

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

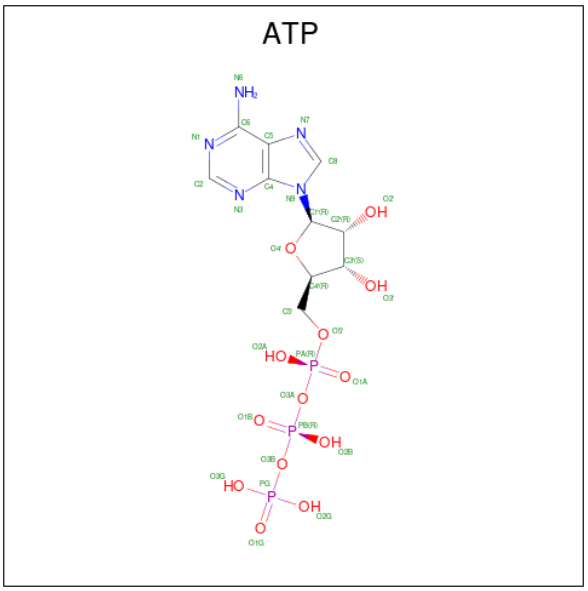
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
12	A	2	2	2	0	0
12	B	1	1	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
12	C	1	1	1	0	0
			Total	Zn		
12	I	2	2	2	0	0
			Total	Zn		
12	J	1	1	1	0	0
			Total	Zn		
12	L	1	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
			Total	C	N	O	P		
13	A	1	30	10	5	12	3	0	0

- Molecule 14 is water.

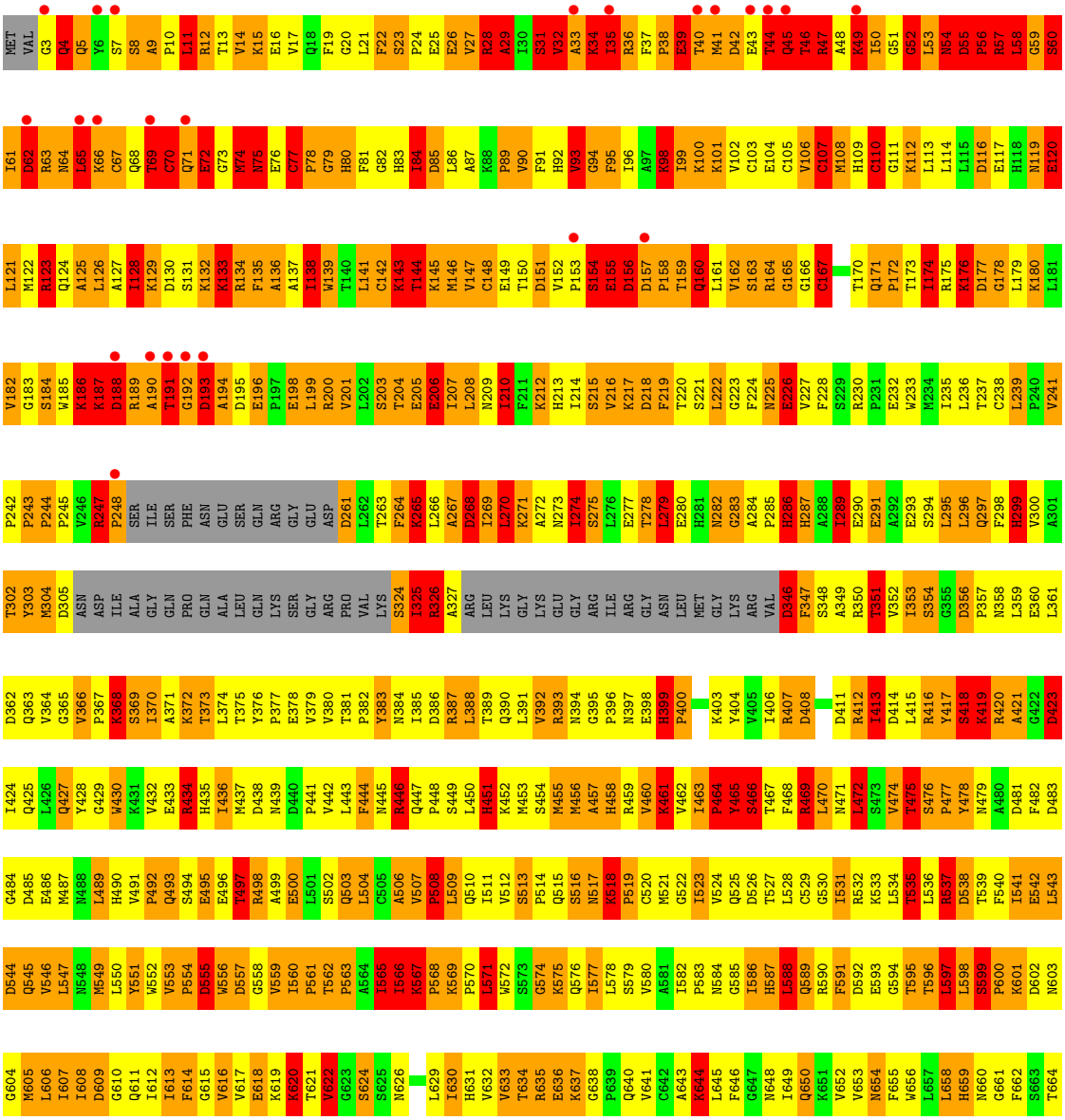
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
14	B	1	1	1	0	0

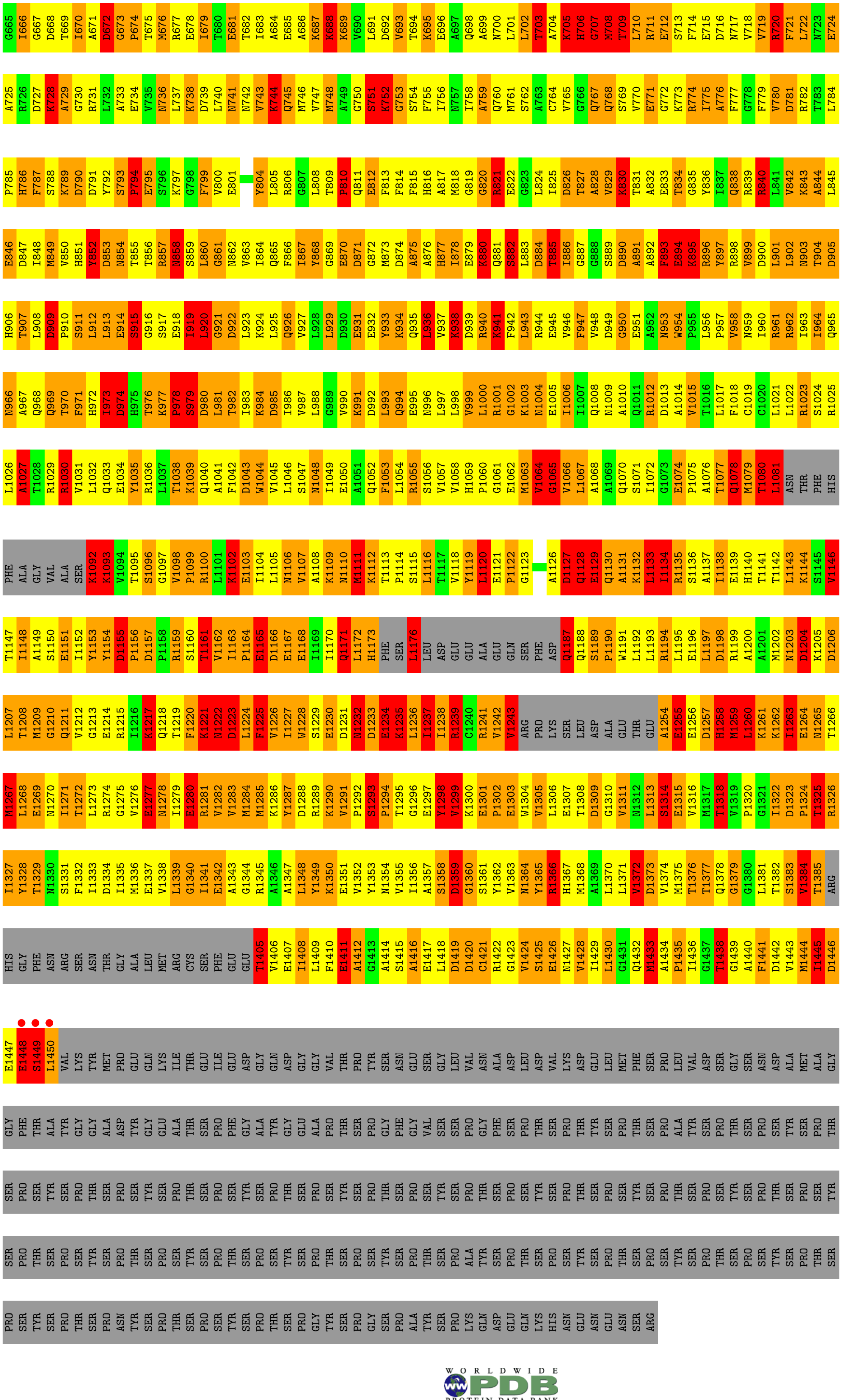
3 Residue-property plots (i)

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: 2% 7% 32% 28% 11% 22%




G665 I666 G667 D668 T669 I670 A671 D672 G673 P674 T675 M676 R677 E678 I679 T680 E681 T682 I683 A684 E685 A686 K687 K688 K689 V690 L691 D692 V693 T694 K695 E696 A697 Q698 A699 N700 L701 L702 T703 A704 K705 H706 G707 M708 T709 L710 R711 E712 S713 F714 E715 D716 N717 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 F755 I756 N757 I758 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 G798 F799 V800 E801 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 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 E879 K880 Q881 S882 L883 D884 T885 I886 G887 G888 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 E951 A952 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 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 A1069 Q1070 S1071 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 V1098 P1099 R1100 L1101 K1102 E1103 I1104 L1105 N1106 V1107 A1108 K1109 N1110 M1111 K1112 T1113 P1114 S1115 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 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 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 PHE GLU GLU 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 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 TYR SER PRO THR SER PRO SER TYR 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 ALA TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO THR SER
PRO SER TYR SER PRO THR SER PRO ASN TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO GLY TYR SER PRO GLY SER PRO ALA TYR SER PRO LYS GLN ASP GLU GLN LYS HIS ASN GLU ASN GLU ASN SER ARG

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

Chain B: 4% 6% 37% 34% 12% 11%

MET SER ASP LEU ALA ASN SER GLU LYS TYR TYR ASP GLU ASP PRO TYR GLY F18 E19 D20 E21 P24 I25 T26 A27 E28 D29 S30 W31 A32 V33 I34 S35 A36 F37 F38 R39 E40 K41 G42 L43 V44 S45 Q46 Q47 L48 D49 S50 F51 F54 V55 D56 Y57 T58 L59 Q60 D61 I62

I63 C64 E65 D66 S67 T68 L69 I70 LEU GLU GLN LEU ALA GLN HIS THR THR GLU SER ASP ASN ILE SER ARG LYS TYR E89 I90 S91 F92 G93 K94 I95 Y96 V97 T98 R99 P100 M101 V102 N103 E104 S105 D106 G107 V108 T109 H110 A111 L112 Y113 P114 Q115 E116 A117 R118 L119 R120 N121 L122

T123 Y124 S125 S126 G127 L128 F129 V130 D131 V132 K133 K134 R135 T136 Y137 E138 ALA ILE ASP VAL PRO GLY ARG GLU LEU LYS TYR GLU LEU ILE ALA GLU GLU SER GLU ASP ASP SER GLU SER GLY K164 V165 F166 I167 G168 R169 L170 P171 I172 M173 L174 R175 S176 K177 N178 C179 Y180 L181 S182

E183 A184 T185 E186 S187 D188 L189 Y190 K191 L192 K193 E194 C195 P196 F197 D198 M199 G200 G201 Y202 F203 I204 I205 N206 G207 S208 E209 K210 V211 L212 I213 A214 Q215 E216 R217 S218 A219 G220 N221 I222 V223 Q224 V225 F226 K227 K228 A229 A230 P231 S232 P233 I234 S235 H236 V237 A238 E239 I240 R241 S242

A243 L244 E245 K246 G247 S248 R249 F250 I251 S252 T253 L254 Q255 V256 K257 L258 Y259 G260 R261 E262 G263 S264 S265 A266 R267 T268 I269 K270 A271 T272 L273 P274 Y275 I276 K277 Q278 D279 I280 P281 I282 V283 I284 I285 F286 R287 A288 L289 G290 I291 I292 P293 D294 G295 E296 I297 L298 E299 H300 I301 C302

Y303 D304 V305 N306 D307 W308 Q309 M310 L311 E312 M313 L314 K315 P316 C317 V318 E319 D320 G321 F322 V323 I324 Q325 D326 R327 E328 T329 A330 L331 D332 F333 R336 R337 G338 T339 A340 L341 G342 I343 K344 K345 E346 K347 R348 I349 Q350 Y351 A352 K353 D354 I355 L356 Q357 K358 E359 F360 L361 P362 H363

I364 T365 Q366 L367 E368 G369 F370 E371 S372 R373 K374 A375 F376 F377 L378 G379 Y380 M381 I382 N383 R384 L385 L386 L387 C388 A389 L390 D391 R392 K393 D394 Q395 D396 D397 R398 D399 H400 F401 G402 K403 K404 R405 L406 D407 L408 A409 G410 P411 L412 L413 A414 Q415 L416 F417 K418 T419 L420 F421 K422 K423

L424 T425 K426 D427 I428 F429 R430 Y431 M432 Q433 R434 T435 V436 E437 GLU ALA HIS ASP PHE ASN MET LYS LEU L446 A447 I448 N449 A450 K451 T452 I453 T454 S455 G456 L457 K458 Y459 A462 T463 G464 N465 W466 G467 GLU GLN LYS LYS ALA MET SER SER ARG A477 G478 V479 V482 L483 N484 R485

Y486 T487 Y488 S489 S490 T491 L492 L495 R496 R497 T498 N499 T500 P501 I502 GLY ARG ASP GLY LYS LEU A509 K510 P511 R512 Q513 L514 H515 N516 T517 H518 W519 G520 L521 V522 C523 P524 A525 E526 T527 P528 E529 G530 Q531 A532 C533 G534 L535 V536 K537 N538 L539 S540 L541 M542 S543 C544 I545 S546

V547 G548 T549 D550 P551 K552 P553 I554 I555 T556 R557 L558 S559 E560 W561 G562 M563 E564 P565 I566 E567 D568 Y569 V570 P571 H572 Q573 S574 P575 D576 A577 T578 R579 V580 F581 V582 N583 G584 V585 W586 H587 G588 V589 H590 R591 N592 P593 A594 R595 L596 V597 E598 T599 L600 R601 T602 L603 R604 R605 K606

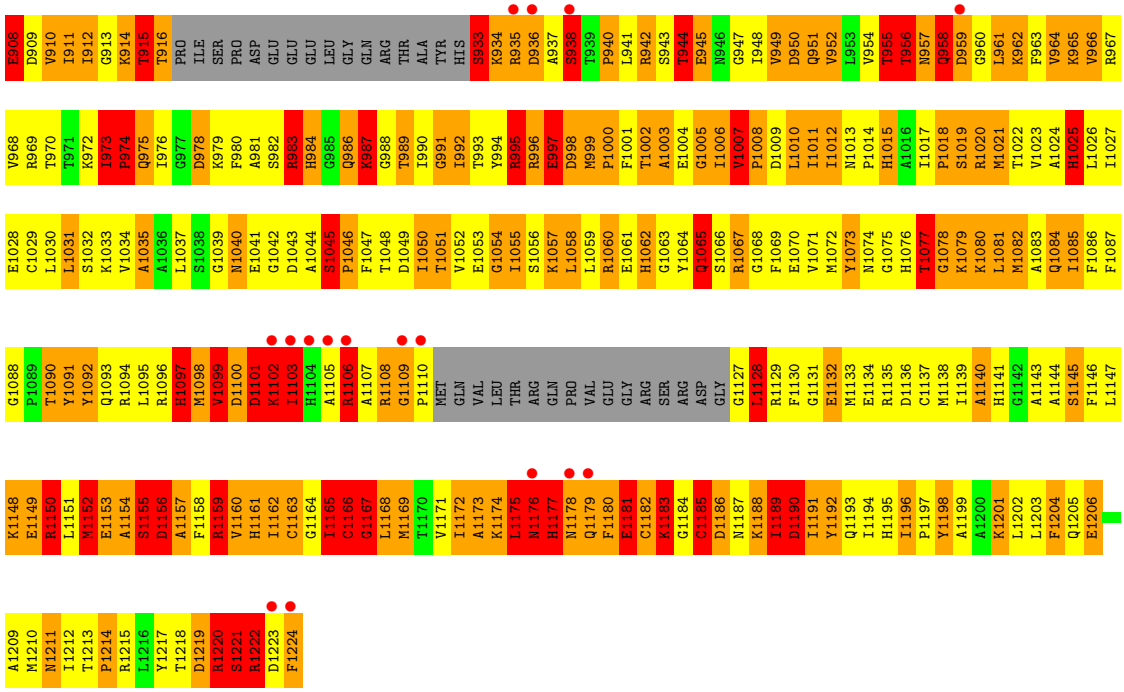
G607 D608 I609 N610 P611 E612 V613 S614 M615 I616 R617 D618 I619 R620 E621 K622 E623 L624 K625 I626 F627 T628 D629 A630 G631 R632 V633 Y634 R635 P636 L637 F638 I639 V640 E641 D642 D643 E644 S645 L646 G647 H648 K649 E650 L651 K652 V653 R654 K655 G656 H657 I658 A659 K660 L661 M662 A663 T664 E665 Y666

Q667 D668 ILE GLU GLY GLY PHE GLU ASP VAL GLU E678 Y679 T680 W681 S682 S683 L684 L685 N686 E687 G688 L689 V690 E691 Y692 I693 D694 A695 E696 E697 E698 E699 S700 I701 L702 I703 A704 M705 Q706 P707 E708 D709 L710 E711 P712 ALA GLU ALA ASN GLU GLU ASN ASP LEU D722 V723 D724 P725 A726

K727 R728 I729 R730 V731 S732 H733 H734 A735 T736 T737 F738 T739 H740 C741 E742 I743 H744 P745 S746 M747 G750 V751 A752 A753 S754 I755 I756 P757 F758 P759 D760 H761 N762 Q763 S764 P765 R766 N767 T768 Y769 Q770 S771 A772 M773 G774 K775 Q776 A777 M778 G779 V780 F781 L782 T783 N784 Y785 N786 V787

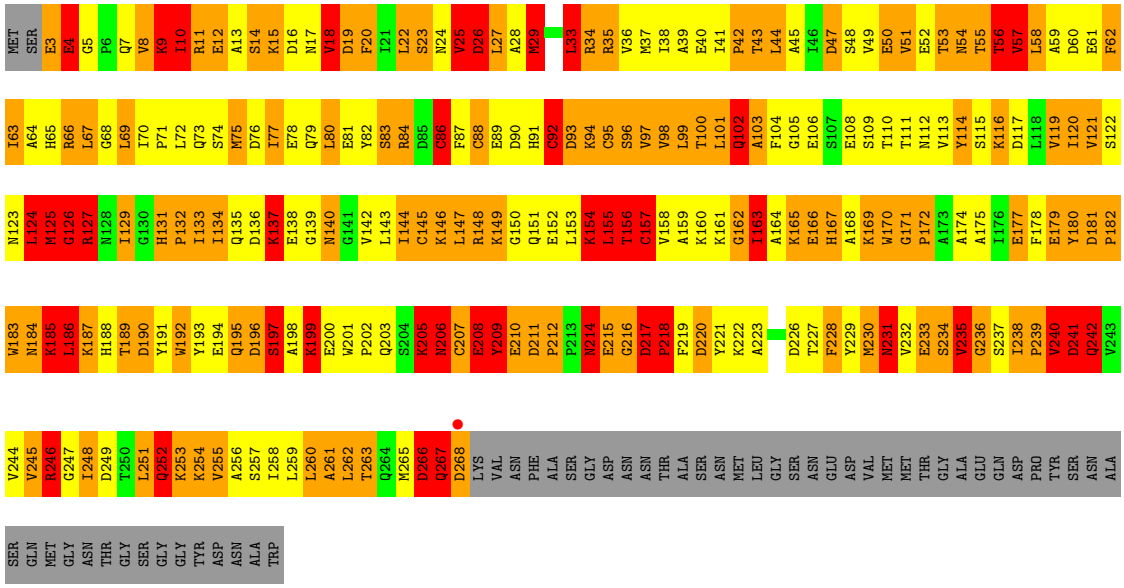
R788 M789 D790 T791 M792 A793 N794 I795 L796 Y797 Y798 P799 Q800 K801 P802 L803 G804 T805 T806 R807 A808 M809 E810 Y811 L812 K813 F814 R815 E816 L817 P818 A819 G820 Q821 N822 A823 I824 V825 A826 I827 A828 C829 Y830 S831 G832 Y833 N834 Q835 E836 D837 S838 M839 I840 M841 N842 Q843 S844 S845 I846 D847

R848 G849 L850 F851 R852 S853 L854 F855 F856 R857 S858 Y859 M860 D861 Q862 E863 K864 K865 Y866 G867 M868 S869 I870 T871 E872 T873 F874 E875 K876 P877 Q878 R879 T880 N881 T882 L883 R884 M885 K886 H887 G888 T889 Y890 D891 K892 L893 D894 D895 D896 G897 L898 I899 A900 P901 G902 V903 R904 V905 S906 G907



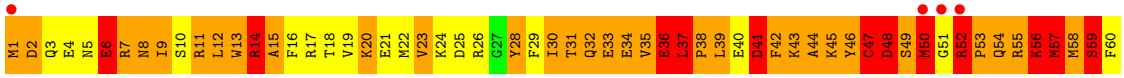
- Molecule 3: DNA-directed RNA polymerase II 45 kDa polypeptide

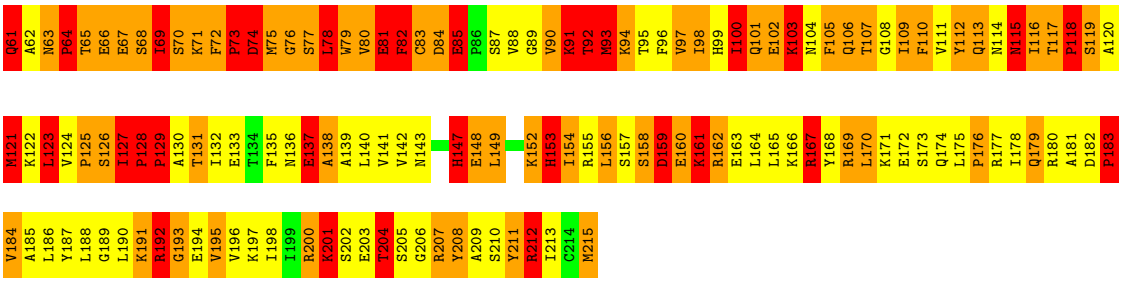
Chain C: 7% 30% 34% 14% 16%



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

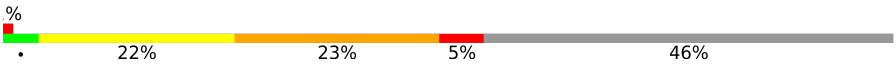
Chain E: 2% 5% 35% 40% 20%

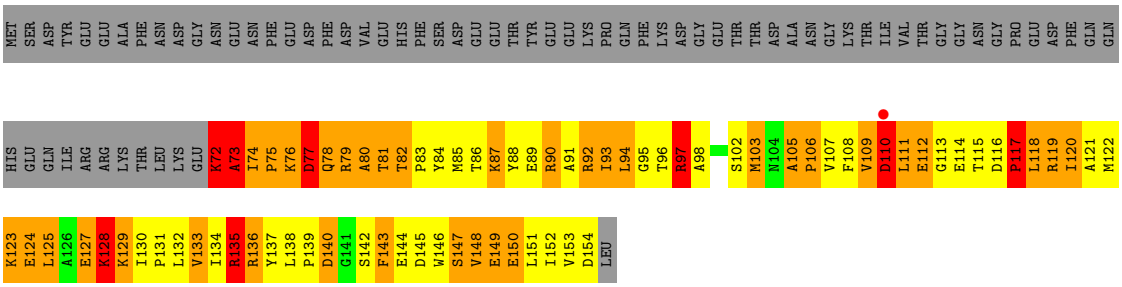




• Molecule 5: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

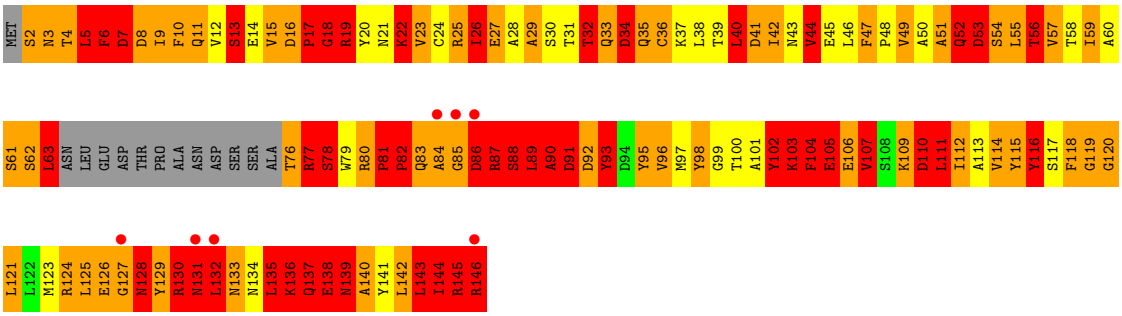
Chain F:





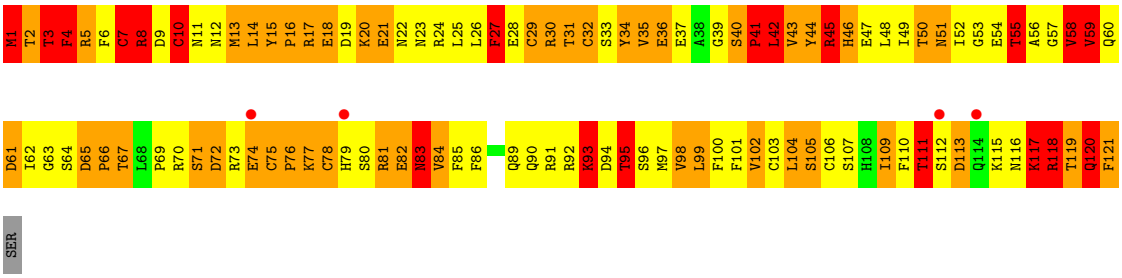
• Molecule 6: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

Chain H:



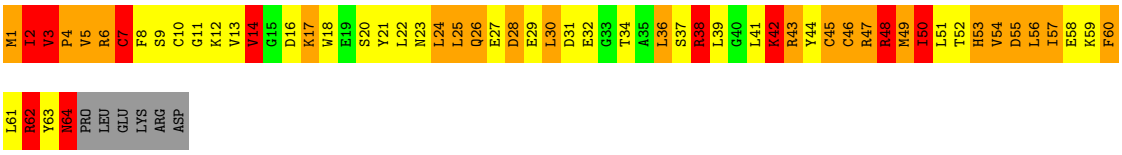
• Molecule 7: DNA-directed RNA polymerase II 14.2 kDa polypeptide

Chain I:



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

Chain J:

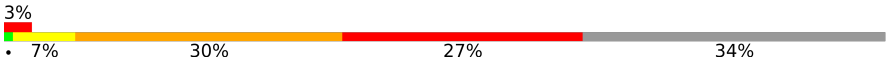


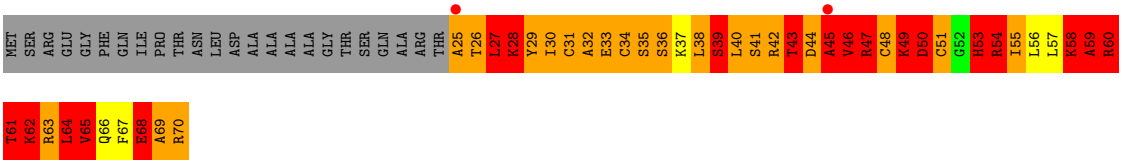
• Molecule 9: DNA-directed RNA polymerase II 13.6 kDa polypeptide

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 – 3.40 39.69 – 3.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.40) 97.3 (39.69-3.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.00 (at 3.32Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.222 , 0.262 0.191 , 0.206	Depositor DCC
R_{free} test set	2319 reflections (2.99%)	wwPDB-VP
Wilson B-factor (Å^2)	57.9	Xtriage
Anisotropy	0.559	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 50.1	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	0.91	EDS
Total number of atoms	27728	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.85% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ATP, ZN, MN

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	4.62	2331/10792 (21.6%)	3.06	1145/14601 (7.8%)
2	B	4.57	1918/8860 (21.6%)	3.01	930/11945 (7.8%)
3	C	4.57	476/2133 (22.3%)	2.99	227/2891 (7.9%)
4	E	4.64	405/1796 (22.6%)	2.92	193/2416 (8.0%)
5	F	4.01	117/682 (17.2%)	3.01	56/922 (6.1%)
6	H	4.54	246/1086 (22.7%)	2.94	113/1470 (7.7%)
7	I	5.00	250/1009 (24.8%)	3.19	124/1357 (9.1%)
8	J	4.42	116/533 (21.8%)	3.66	79/715 (11.0%)
9	K	4.36	210/937 (22.4%)	3.12	108/1265 (8.5%)
10	L	5.55	101/366 (27.6%)	3.72	71/485 (14.6%)
All	All	4.60	6170/28194 (21.9%)	3.05	3046/38067 (8.0%)

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
1	A	0	57
2	B	0	55
3	C	0	15
4	E	1	9
5	F	0	1
6	H	0	12
7	I	0	6
9	K	0	2
10	L	0	2
All	All	1	159

The worst 5 of 6170 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1064	VAL	C-O	33.98	1.88	1.23
10	L	68	GLU	CG-CD	32.08	2.00	1.51
1	A	1064	VAL	CA-C	31.22	2.34	1.52
1	A	734	GLU	CD-OE2	30.34	1.59	1.25
2	B	552	MET	CG-SD	28.64	2.55	1.81

The worst 5 of 3046 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1064	VAL	CA-C-O	-35.84	44.85	120.10
5	F	136	ARG	NE-CZ-NH1	-34.13	103.24	120.30
1	A	774	ARG	NE-CZ-NH1	32.04	136.32	120.30
1	A	1064	VAL	O-C-N	-31.34	69.93	123.20
1	A	821	ARG	NE-CZ-NH2	-29.52	105.54	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	E	204	THR	CB

5 of 159 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	31	SER	Peptide
1	A	44	THR	Peptide
1	A	52	GLY	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	10606	0	10662	1054	0
2	B	8690	0	8707	918	0
3	C	2095	0	2054	226	0
4	E	1760	0	1788	222	0
5	F	670	0	689	58	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	220	0
7	I	990	0	948	101	0
8	J	525	0	537	61	0
9	K	919	0	928	108	0
10	L	364	0	387	92	0
11	A	2	0	0	0	0
12	A	2	0	0	1	0
12	B	1	0	0	0	0
12	C	1	0	0	3	0
12	I	2	0	0	0	0
12	J	1	0	0	2	0
12	L	1	0	0	1	0
13	A	30	0	9	0	0
14	B	1	0	0	0	0
All	All	27728	0	27749	2914	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 53.

The worst 5 of 2914 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:137:TYR:CB	2:B:137:TYR:CG	1.77	1.66
9:K:26:LYS:CD	9:K:26:LYS:CE	1.74	1.66
1:A:1225:PHE:CB	1:A:1225:PHE:CG	1.78	1.65
1:A:977:LYS:CG	1:A:977:LYS:CD	1.74	1.64
2:B:884:ARG:CD	2:B:884:ARG:CG	1.75	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	1332/1733 (77%)	1173 (88%)	93 (7%)	66 (5%)	2	14
2	B	1071/1224 (88%)	914 (85%)	110 (10%)	47 (4%)	2	16
3	C	264/318 (83%)	230 (87%)	22 (8%)	12 (4%)	2	16
4	E	213/215 (99%)	182 (85%)	20 (9%)	11 (5%)	2	13
5	F	81/155 (52%)	72 (89%)	7 (9%)	2 (2%)	5	26
6	H	129/146 (88%)	82 (64%)	21 (16%)	26 (20%)	0	0
7	I	119/122 (98%)	107 (90%)	11 (9%)	1 (1%)	19	51
8	J	62/70 (89%)	58 (94%)	3 (5%)	1 (2%)	9	34
9	K	112/120 (93%)	99 (88%)	9 (8%)	4 (4%)	3	21
10	L	44/70 (63%)	24 (54%)	9 (20%)	11 (25%)	0	0
All	All	3427/4173 (82%)	2941 (86%)	305 (9%)	181 (5%)	2	13

5 of 181 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	ILE
1	A	38	PRO
1	A	45	GLN
1	A	47	ARG
1	A	59	GLY

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	1181/1520 (78%)	971 (82%)	210 (18%)	2	6
2	B	947/1061 (89%)	788 (83%)	159 (17%)	2	8
3	C	234/274 (85%)	194 (83%)	40 (17%)	2	8
4	E	197/197 (100%)	145 (74%)	52 (26%)	0	1
5	F	73/137 (53%)	62 (85%)	11 (15%)	3	12
6	H	117/128 (91%)	71 (61%)	46 (39%)	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
7	I	115/116 (99%)	82 (71%)	33 (29%)	0 1
8	J	59/65 (91%)	45 (76%)	14 (24%)	1 2
9	K	99/102 (97%)	81 (82%)	18 (18%)	1 6
10	L	40/57 (70%)	21 (52%)	19 (48%)	0 0
All	All	3062/3657 (84%)	2460 (80%)	602 (20%)	1 4

5 of 602 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	F	81	THR
9	K	54	ARG
6	H	13	SER
5	F	77	ASP
7	I	4	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	1193	GLN
7	I	12	ASN
3	C	112	ASN
4	E	104	ASN
9	K	2	ASN

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	26,32,33	1.05	2 (7%)	30,50,52	1.51	6 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
13	ATP	A	3011	11	1/1/6/7	5/18/34/38	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	C4-N3	3.60	1.40	1.35
13	A	3011	ATP	C8-N7	-2.60	1.30	1.34

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
13	A	3011	ATP	O4'-C1'-C2'	3.62	113.08	106.25
13	A	3011	ATP	PA-O3A-PB	-3.55	120.63	132.83
13	A	3011	ATP	PB-O3B-PG	-3.43	121.06	132.83
13	A	3011	ATP	C2'-C1'-N9	2.67	120.42	114.27
13	A	3011	ATP	C4-C5-N7	2.59	112.10	109.40

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
13	A	3011	ATP	C1'

All (5) torsion outliers are listed below:

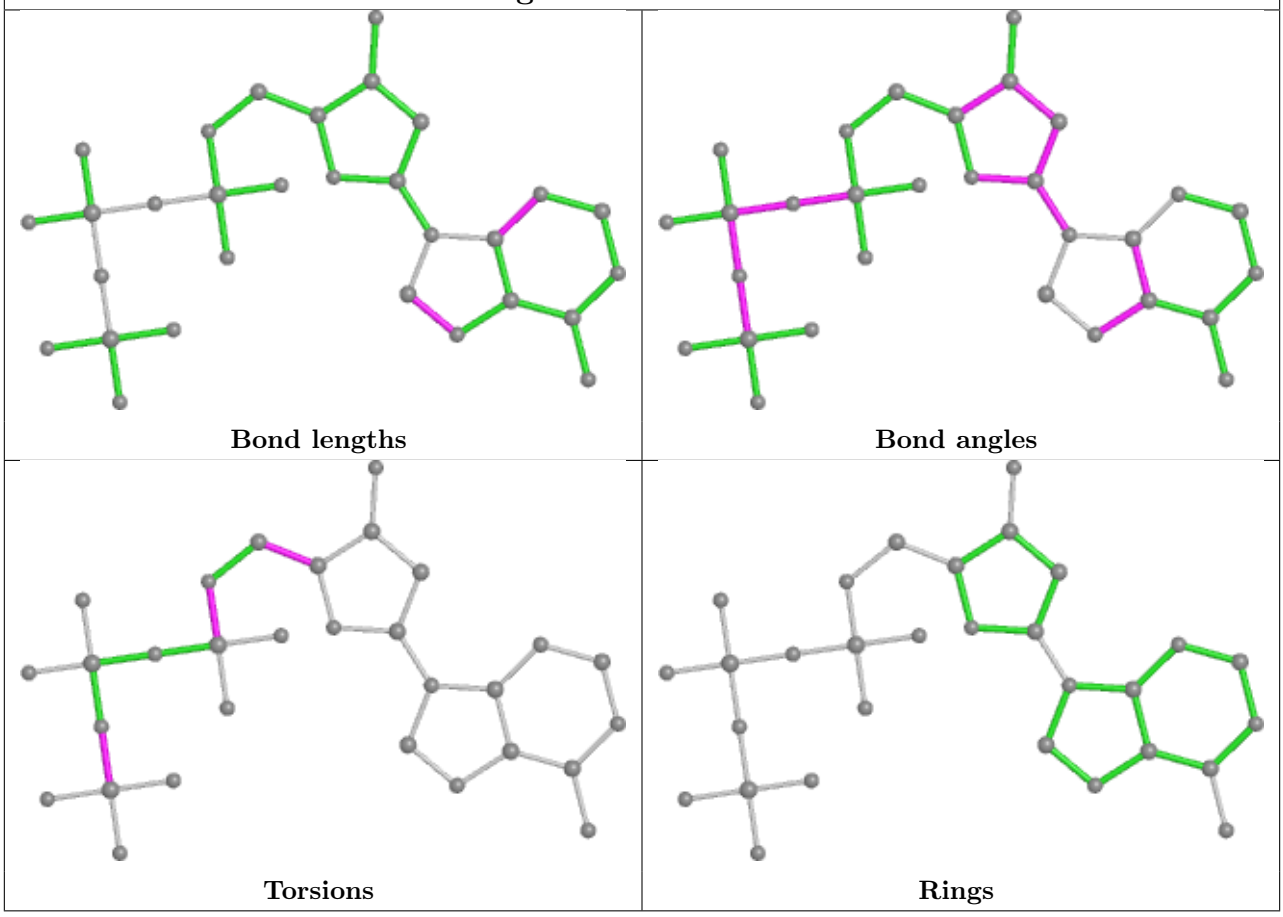
Mol	Chain	Res	Type	Atoms
13	A	3011	ATP	O4'-C4'-C5'-O5'
13	A	3011	ATP	C3'-C4'-C5'-O5'
13	A	3011	ATP	PB-O3B-PG-O2G
13	A	3011	ATP	PB-O3B-PG-O3G
13	A	3011	ATP	C5'-O5'-PA-O3A

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
2	B	12
1	A	8
3	C	2
7	I	2
5	F	1
9	K	1
8	J	1

The worst 5 of 27 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	1014:ALA	C	1015:VAL	N	1.20
1	A	1378:GLN	C	1379:GLY	N	1.20
1	B	49:ASP	C	50:SER	N	1.20
1	B	586:TRP	C	587:HIS	N	1.20
1	B	724:ASP	C	725:PRO	N	1.20

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	1349/1733 (77%)	-0.34	27 (2%)	65	64	1, 22, 99, 165	0
2	B	1091/1224 (89%)	-0.31	47 (4%)	35	35	1, 22, 107, 155	0
3	C	266/318 (83%)	-0.53	1 (0%)	92	92	1, 22, 72, 145	0
4	E	215/215 (100%)	-0.30	4 (1%)	66	65	1, 40, 102, 138	0
5	F	83/155 (53%)	-0.33	1 (1%)	79	77	1, 21, 63, 88	0
6	H	133/146 (91%)	0.17	7 (5%)	26	27	14, 65, 120, 167	0
7	I	121/122 (99%)	-0.17	4 (3%)	46	45	1, 30, 79, 120	0
8	J	64/70 (91%)	-0.50	0	100	100	3, 18, 64, 91	0
9	K	114/120 (95%)	-0.32	0	100	100	1, 32, 72, 103	0
10	L	46/70 (65%)	0.10	2 (4%)	35	35	16, 72, 123, 140	0
All	All	3482/4173 (83%)	-0.31	93 (2%)	54	53	1, 25, 103, 167	0

The worst 5 of 93 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	69	THR	7.8
2	B	882	THR	6.9
2	B	1109	GLY	6.6
1	A	1450	LEU	6.2
1	A	1449	SER	6.2

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)

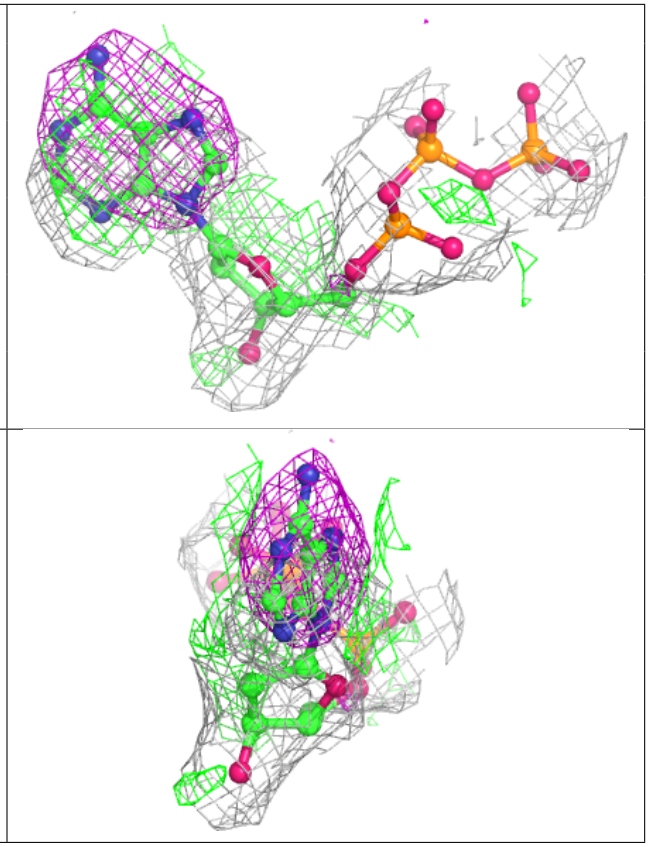
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
13	ATP	A	3011	30/31	0.80	0.30	5,46,82,85	0
12	ZN	I	3004	1/1	0.94	0.08	66,66,66,66	0
11	MN	A	3009	1/1	0.94	0.09	4,4,4,4	0
12	ZN	A	3006	1/1	0.95	0.05	47,47,47,47	0
12	ZN	J	3001	1/1	0.95	0.06	34,34,34,34	0
12	ZN	A	3008	1/1	0.95	0.05	91,91,91,91	0
12	ZN	L	3005	1/1	0.96	0.11	90,90,90,90	0
11	MN	A	3010	1/1	0.97	0.05	12,12,12,12	0
12	ZN	C	3002	1/1	0.98	0.07	42,42,42,42	0
12	ZN	I	3003	1/1	0.98	0.03	46,46,46,46	0
12	ZN	B	3007	1/1	0.98	0.06	57,57,57,57	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.