



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 08:46 am GMT

PDB ID : 3H0G
Title : RNA Polymerase II from Schizosaccharomyces pombe
Authors : Spahr, H.; Calero, G.; Bushnell, D.A.; Kornberg, R.D.
Deposited on : 2009-04-09
Resolution : 3.65 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

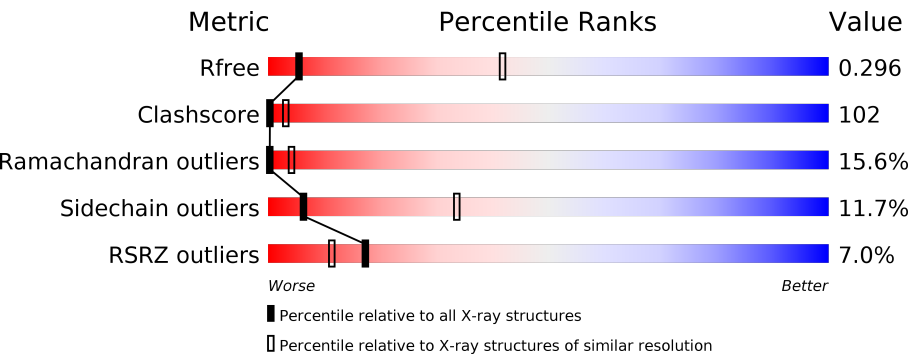
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.65 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}	100719	1134 (3.82-3.50)
Clashscore	112137	1267 (3.82-3.50)
Ramachandran outliers	110173	1219 (3.82-3.50)
Sidechain outliers	110143	1218 (3.82-3.50)
RSRZ outliers	101464	1160 (3.82-3.50)

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

Mol	Chain	Length	Quality of chain
1	A	1752	<div><div>5%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>19%45%18%•15%</div></div>
1	M	1752	<div><div>10%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>19%45%17%•16%</div></div>
2	B	1210	<div><div>2%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>25%54%14%•5%</div></div>
2	N	1210	<div><div>10%</div><div><div></div><div></div><div></div><div></div><div></div></div><div>25%54%14%•5%</div></div>
3	C	297	<div><div></div><div><div></div><div></div><div></div><div></div><div></div></div><div>30%48%10%•11%</div></div>
3	O	297	<div><div></div><div><div></div><div></div><div></div><div></div><div></div></div><div>30%49%9%•11%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	D	135	
4	P	135	
5	E	210	
5	Q	210	
6	F	142	
6	R	142	
7	G	172	
7	S	172	
8	H	125	
8	T	125	
9	I	113	
9	U	113	
10	J	71	
10	V	71	
11	K	123	
11	W	123	
12	L	63	
12	X	63	

2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 62870 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
1	A	1496	Total	C	N	O	S	0	0	0
			11802	7415	2071	2246	70			
1	M	1476	Total	C	N	O	S	0	0	0
			11666	7334	2047	2216	69			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1150	Total	C	N	O	S	0	0	0
			9180	5772	1630	1716	62			
2	N	1150	Total	C	N	O	S	0	0	0
			9180	5772	1630	1716	62			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	263	Total	C	N	O	S	0	0	0
			2088	1315	355	406	12			
3	O	263	Total	C	N	O	S	0	0	0
			2088	1315	355	406	12			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit rpb4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	130	Total	C	N	O	S	0	0	0
			1036	649	176	205	6			
4	P	130	Total	C	N	O	S	0	0	0
			1036	649	176	205	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	207	Total	C	N	O	S	0	0	0
			1663	1050	301	306	6			
5	Q	207	Total	C	N	O	S	0	0	0
			1663	1050	301	306	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	83	Total	C	N	O	S	0	0	0
			656	416	112	125	3			
6	R	83	Total	C	N	O	S	0	0	0
			656	416	112	125	3			

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit rpb7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	170	Total	C	N	O	S	0	0	0
			1330	860	217	247	6			
7	S	170	Total	C	N	O	S	0	0	0
			1330	860	217	247	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	124	Total	C	N	O	S	0	0	0
			996	631	167	195	3			
8	T	124	Total	C	N	O	S	0	0	0
			996	631	167	195	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	111	Total	C	N	O	S	0	0	0
			902	551	164	176	11			
9	U	111	Total	C	N	O	S	0	0	0
			902	551	164	176	11			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	64	Total	C	N	O	S	0	0	0
			518	330	87	94	7			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	V	64	Total	C	N	O	S	0	0	0
			518	330	87	94	7			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	119	Total	C	N	O	S	0	0	0
			955	608	159	182	6			
11	W	119	Total	C	N	O	S	0	0	0
			955	608	159	182	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	45	Total	C	N	O	S	0	0	0
			368	225	74	61	8			
12	X	45	Total	C	N	O	S	0	0	0
			368	225	74	61	8			

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	J	1	Total	Zn	0	0
			1	1		
13	B	1	Total	Zn	0	0
			1	1		
13	I	2	Total	Zn	0	0
			2	2		
13	C	1	Total	Zn	0	0
			1	1		
13	V	1	Total	Zn	0	0
			1	1		
13	A	2	Total	Zn	0	0
			2	2		
13	N	1	Total	Zn	0	0
			1	1		
13	U	2	Total	Zn	0	0
			2	2		
13	X	1	Total	Zn	0	0
			1	1		
13	O	1	Total	Zn	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	L	1	Total 1	Zn 1	0	0
13	M	2	Total 2	Zn 2	0	0

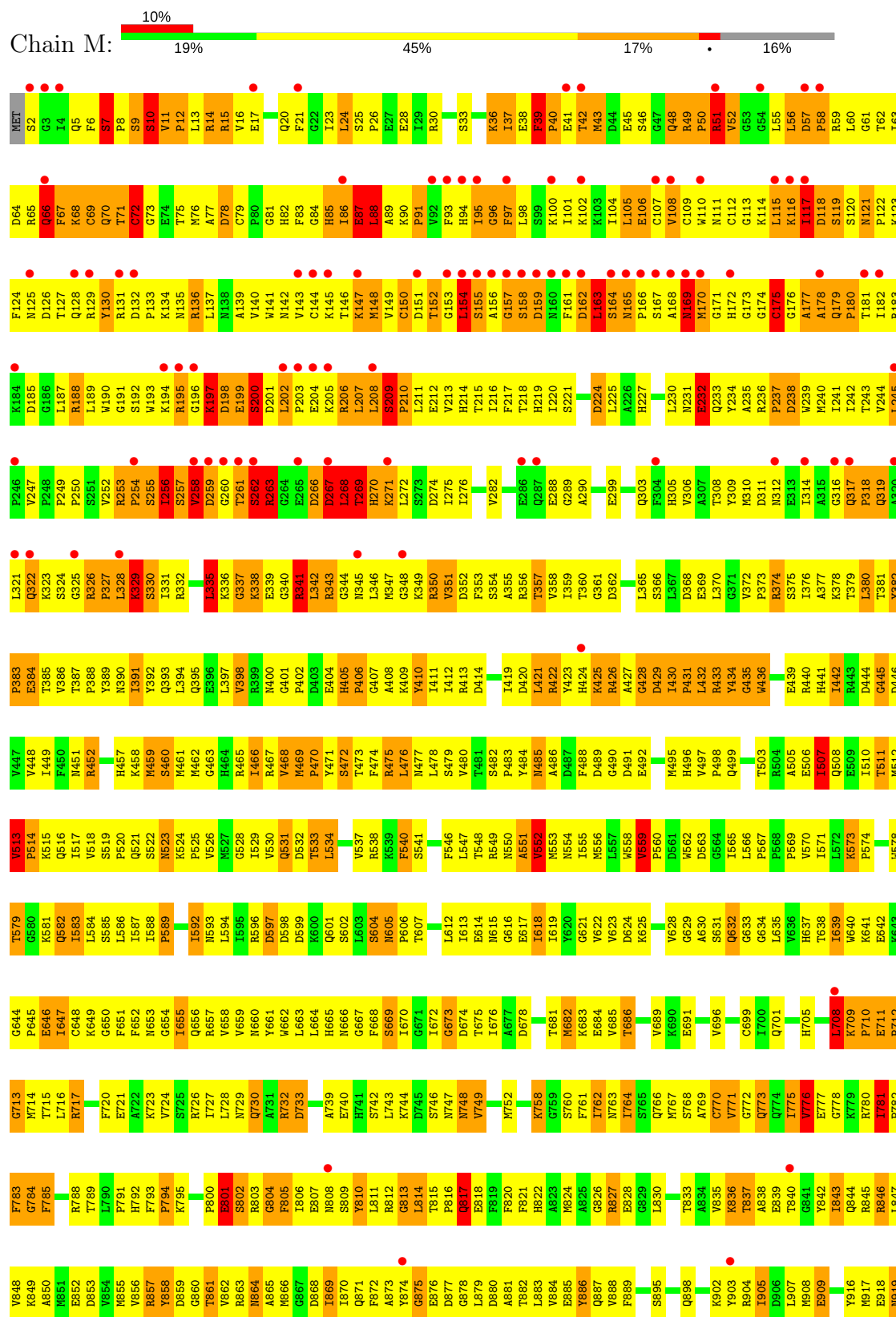
- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

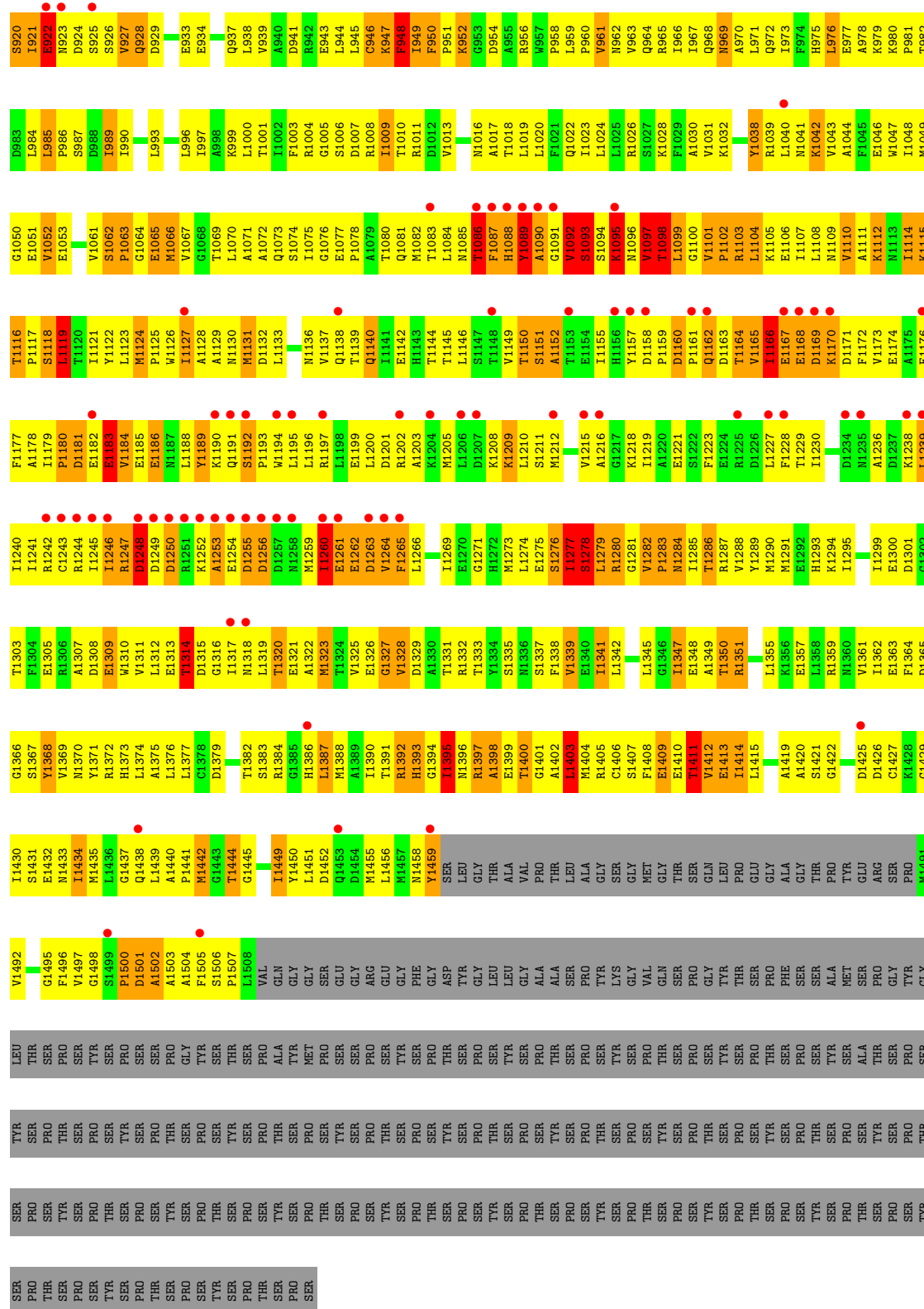
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	1	Total 1	Mg 1	0	0
14	M	1	Total 1	Mg 1	0	0

PRO	SER	SER	PRO	SER	PRO	ASP	M1435	H1373	E1309	R1244	A1175	K1115	M1049	T982	N919	R846	P782	P712
SER	PRO	SER	THR	ALA	VAL	ALA	L1436	L1374	W1310	I1245	F1176	T1116	G1050	D983	S920	L847	F783	G713
SER	THR	PRO	THR	ALA	VAL	ALA	Q1437	L1375	V1311	I1246	F1177	P1117	E1051	L984	I921	V848	G784	
PRO	PRO	PRO	THR	ALA	VAL	ALA	Q1438	L1376	L1312	I1247	I1178	S1118	E1052	L985	E922	K849	F785	T715
THR	SER	SER	THR	PHE	GLY	THR	L1439	L1377	E1313	D1248	P1179	L1119	E1053	P986	D924	K851	K786	L716
SER	PRO	PRO	THR	THR	LEU	PRO	C1378	D1378	T1314	D1249	I1180	T1120		S987	N924	M851	K787	R717
PRO	PRO	PRO	THR	THR	LEU	PRO	D1379	D1379	L1315	D1250	D1181	I1121	V1061	D988	S925	E852	R788	
PRO	PRO	PRO	THR	THR	LEU	PRO			G1316	R1251	D1182	Y1122	S1062	D989	S926	E853	F720	
SER	PRO	PRO	THR	ALA	GLN	THR	T1382	T1382	I1317	K1252	E1183	Y1123	P1063	D990	Y927	H854	L790	F721
THR	SER	THR	THR	ALA	GLN	THR	S1383	S1383	N1318	A1253	E1184	M1124	G1064		P791	M855	K722	
SER	SER	SER	THR	ALA	GLN	THR	L1384	L1384	L1319	E1254	E1185	P1125	E1065		D929	V856	K723	
PRO	PRO	PRO	THR	MET	GLY	GLY	G1385	G1385	T1320	D1255	E1186	W1126	M1066			R857	F724	
THR	SER	SER	THR	THR	GLY	GLY	H1386	H1386	E1321	D1256	N1187	I1127	V1067			X858	S725	
SER	PRO	SER	THR	THR	GLY	GLY	L1387	L1387	A1322	D1257	L1188	A1128				D859	K726	
PRO	PRO	PRO	THR	THR	GLY	GLY	M1388	M1388	M1323	A1258						G860	R727	
PRO	PRO	PRO	THR	THR	ARG	GLY	A1389	A1389	M1324	M1259	Y1189	A1129	T1069			T861	L728	
THR	SER	SER	THR	THR	GLY	GLY	T1390	T1390	V1325	E1260	K1190	M1130	L1070			R801	N729	
THR	PRO	PRO	THR	THR	GLY	GLY	L1391	L1391	E1326	I1261	S1191	M1131	A1071			S802	Q730	
THR	PRO	PRO	THR	THR	GLY	GLY	R1392	R1392	G1327	E1262	P1193	L1132	Q1073			R803	A731	
THR	PRO	PRO	THR	THR	GLY	GLY	H1393	H1393	V1328	D1263	W1194	A1134	S1074			G804	R732	
SER	THR	THR	THR	THR	GLY	GLY			D1329	F1264		K1135	R1004			P805	D733	
PRO	PRO	PRO	THR	THR	GLY	GLY	I1395	I1395	A1330	F1265	L1196	M1136	G1005			D868	R806	
PRO	PRO	PRO	THR	THR	GLY	GLY	G1462	G1462	T1331	F1266	R1197	W1137	S1006			I869	E807	
THR	SER	SER	THR	THR	LEU	LEU	T1463	T1463	R1332	L1266	L1198	Q1138	E1077			I870	N808	
SER	SER	SER	THR	THR	LEU	LEU	A1464	A1464	R1333		E1199	T1139	P1078			Q871	S809	
PRO	PRO	PRO	THR	THR	GLY	GLY	E1398	E1398	Y1334	I1269			A1078			Q872	N810	
THR	THR	THR	THR	THR	ALA	ALA	E1399	E1399	Y1335		L1200	Q1140	T1080			F872	Y810	
SER	THR	THR	THR	THR	ALA	ALA	P1466	P1466	S1335	M1273	D1201	T1141	Q1081			A873	L743	
PRO	PRO	PRO	THR	THR	ALA	ALA	T1467	T1467		L1274	D1202	E1142	M1082			Y874	K744	
PRO	PRO	PRO	THR	THR	ALA	ALA	L1468	L1468	F1338	E1275	S1211	H1143	T1083			G875	D745	
SER	PRO	PRO	THR	THR	ALA	ALA	Q1469	Q1469	V1339	S1276	K1203	H1144	L1084			E876	S746	
THR	THR	THR	THR	THR	ALA	ALA	L1403	L1403	O1340	I1277	M1205	T1145	N1085			D877	N747	
SER	SER	SER	THR	THR	ALA	ALA	M1404	M1404	E1340							G878	N748	
SER	PRO	PRO	THR	THR	ALA	ALA	G1470	G1470	I1341	S1278	K1209	L1146	T1086			C879	V749	
PRO	PRO	PRO	THR	THR	GLY	GLY	S1471	S1471	L1342	L1279			F1087			L879		
THR	SER	SER	THR	THR	GLN	GLN	G1472	G1472	L1345	R1280	K1210	T1148	M1088			D880		
THR	PRO	PRO	THR	THR	GLN	GLN	M1473	M1473		G1281	S1211	Y1149	R1089			A881	F819	
PRO	PRO	PRO	THR	THR	GLN	GLN	G1474	G1474	V1346	V1282	M1212	T1150	A1090			P820		
PRO	PRO	PRO	THR	THR	GLN	GLN	T1475	T1475	G1346	P1283		S1151	A1091			T882	F821	
SER	PRO	PRO	THR	THR	GLN	GLN	S1476	S1476	E1347	N1284	A1152	A1152	V1092			L883		
THR	SER	SER	THR	THR	GLN	GLN	Q1477	Q1477	E1348				Y1092			V884	H822	
SER	PRO	PRO	THR	THR	GLN	GLN	L1478	L1478		I1285	A1216	T1153	S1093			E885	H823	
PRO	PRO	PRO	THR	THR	GLN	GLN	P1479	P1479	T1350	T1286	G1217	E1154	S1094			Y886	M824	
THR	THR	THR	THR	THR	GLN	GLN	E1480	E1480	R1351	G1287	K1218	H1155	K1095			Q887	A825	
SER	PRO	PRO	THR	THR	GLN	GLN	G1481	G1481		V1288	I1219	H1156	N1096			G826	I762	
PRO	PRO	PRO	THR	THR	GLN	GLN	A1482	A1482	L1415	Y1289	E1220	Y1157	N1097			R827	N763	
SER	PRO	PRO	THR	THR	GLN	GLN			L1355	M1290	E1221	D1158	T1098			E828	I764	
THR	SER	SER	THR	THR	GLN	GLN	G1483	G1483	K1356	M1291	S1222	P1159	K1028			G829		
PRO	PRO	PRO	THR	THR	GLN	GLN	D1417	D1417	E1357	M1292	F1223	D1160	F1029			L830		
SER	PRO	PRO	THR	THR	GLN	GLN	A1416	A1416	L1358	H1293		P1161	A1030			Q898		
SER	PRO	PRO	THR	THR	GLN	GLN	A1419	A1419	L1359	K1294	F1227	Q1162	V1031			T833		
PRO	PRO	PRO	THR	THR	GLN	GLN	A1420	A1420	E1292	I1295	L1228	D1163	R1032			A834		
THR	THR	THR	THR	THR	GLN	GLN	S1421	S1421	H1294		T1229	Y1164	P1102			V835		
PRO	PRO	PRO	THR	THR	GLN	GLN	G1422	G1422	V1361		T1230	T1164	K1032			K836		
PRO	PRO	PRO	THR	THR	GLN	GLN			I1362							T837		
THR	THR	THR	THR	THR	GLN	GLN	D1425	D1425	E1363	I1299		Y1165	R1038			Y903		
THR	THR	THR	THR	THR	GLN	GLN	D1426	D1426	F1364	E1300		T1166	K1039			R904		
SER	PRO	PRO	THR	THR	GLN	GLN	C1427	C1427	D1365	D1301		E1167	L1040			T905		
PRO	PRO	PRO	THR	THR	GLN	GLN	K1428	K1428	G1366	G1302		E1168	N1041			D906		
THR	THR	THR	THR	THR	GLN	GLN	G1429	G1429	S1367	T1303		K1168	K1042			L907		
PRO	PRO	PRO	THR	THR	GLN	GLN			Y1367	F1304		N1109	V1043			N908		
SER	PRO	PRO	THR	THR	GLN	GLN			O1368	L1304		K1170	A1044			E977		
PRO	PRO	PRO	THR	THR	GLN	GLN	F1496	F1496	V1369	E1305		D1171	F1045			Y842		
PRO	PRO	PRO	THR	THR	GLN	GLN	V1497	V1497	V1370	R1306		T1172	E1046			Y916		
THR	THR	THR	THR	THR	GLN	GLN	GLY	GLY	A1307	L1241		V1173	M1047			Q843		
PRO	PRO	PRO	THR	THR	GLN	GLN	A1433	A1433	Y1371	C1242			K1112			N917		
SER	PRO	PRO	THR	THR	GLN	GLN			R1372	C1243			I1114			E918		

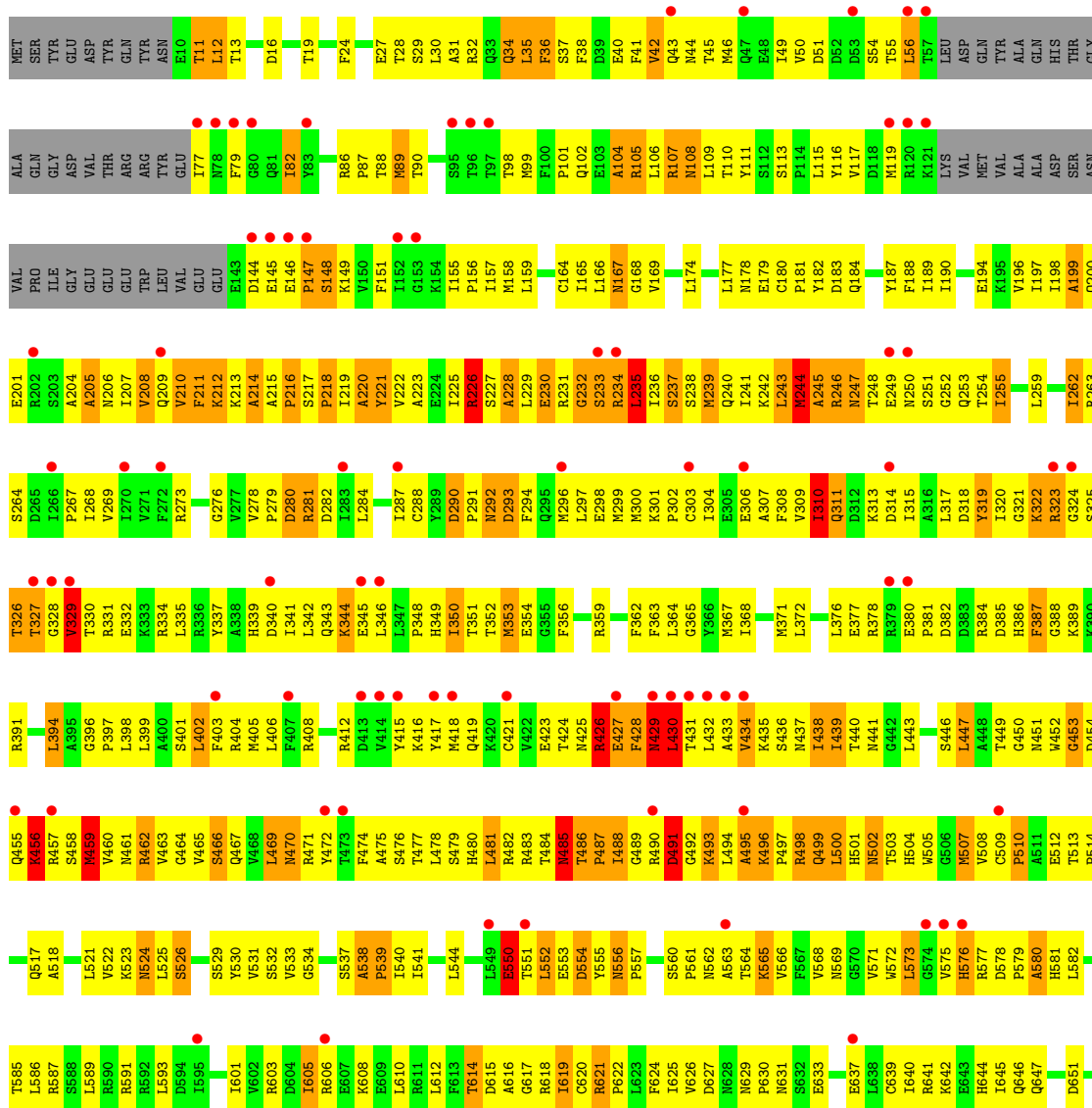
PRO
THR
SER
PRO
SER
TYR
SER
PRO
THR
SER
PRO
SER

• Molecule 1: DNA-directed RNA polymerase II subunit rpb1

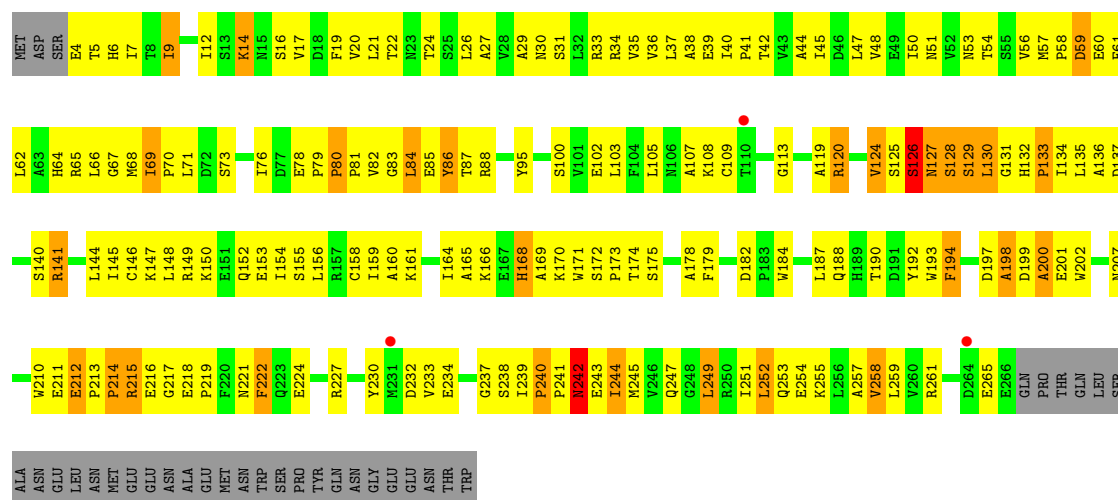




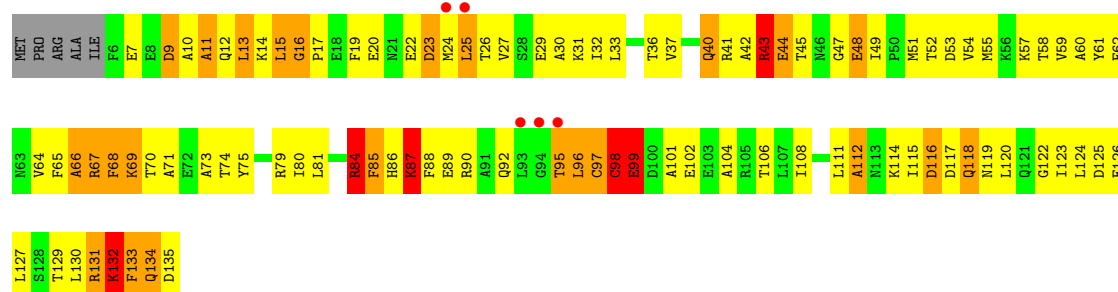




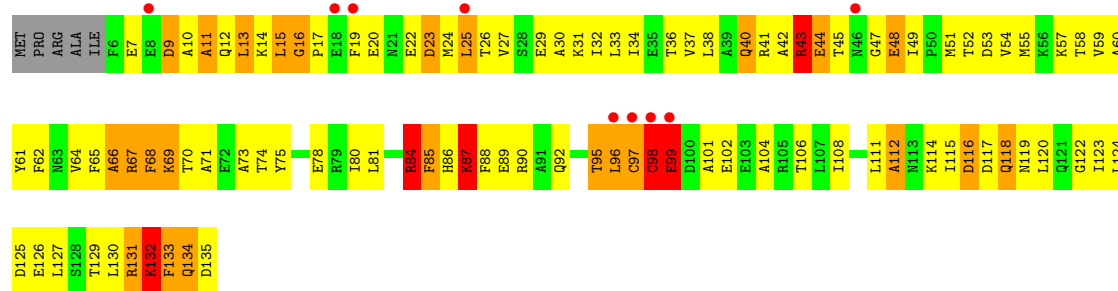




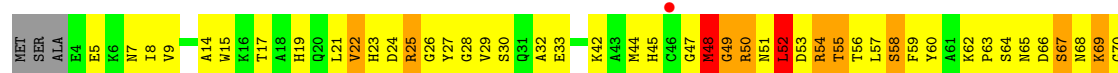
• Molecule 4: DNA-directed RNA polymerase II subunit rpb4

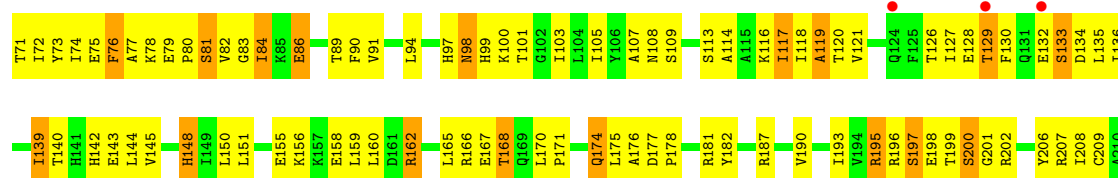


• Molecule 4: DNA-directed RNA polymerase II subunit rpb4

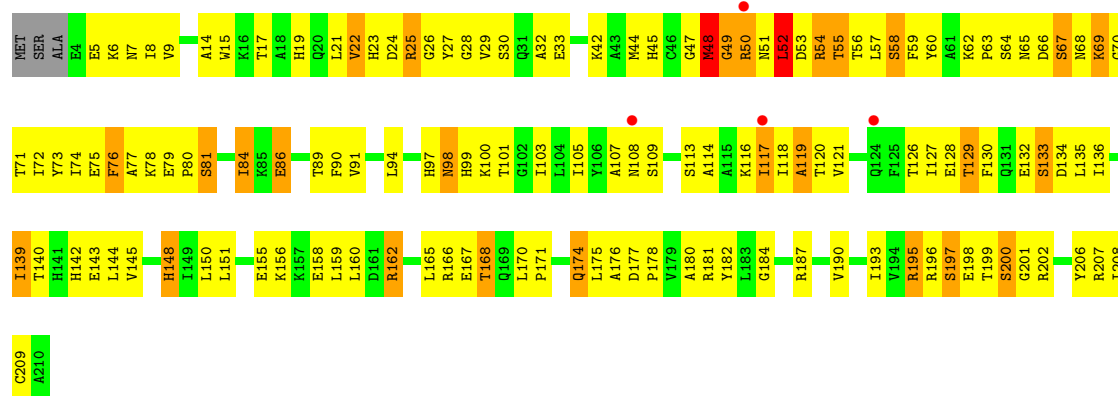


• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

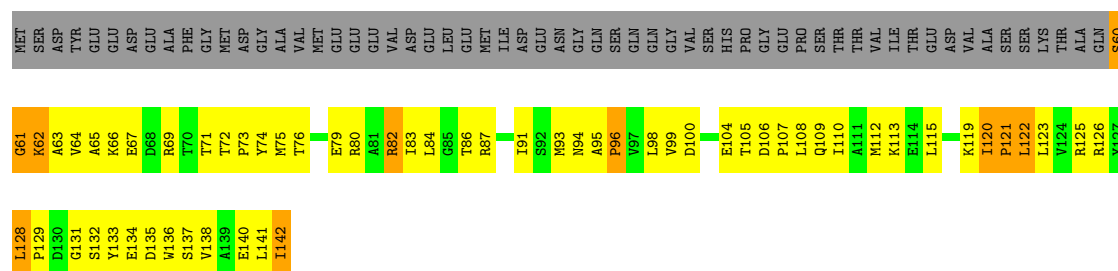




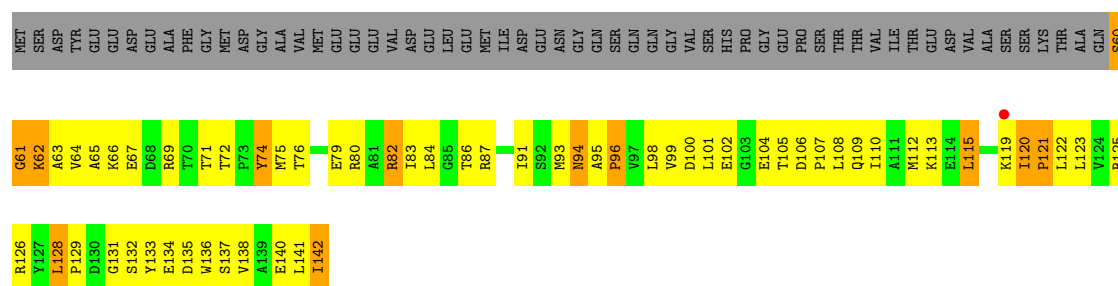
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1



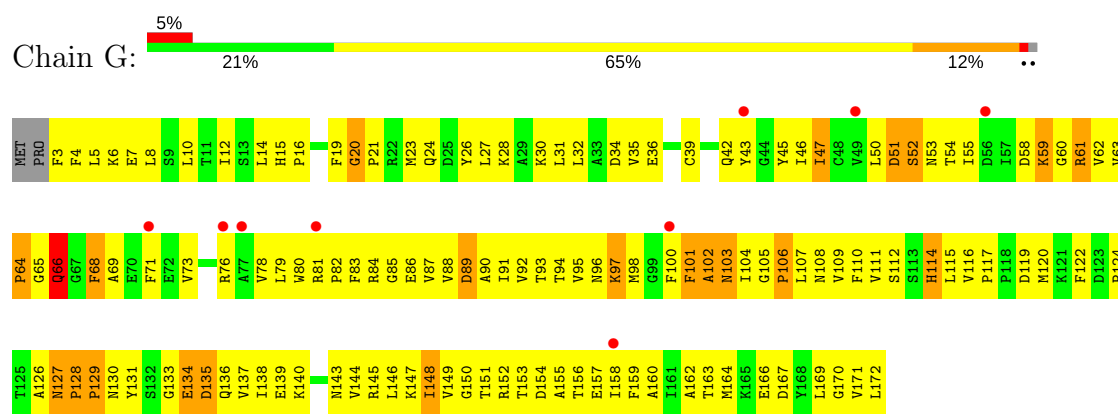
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2



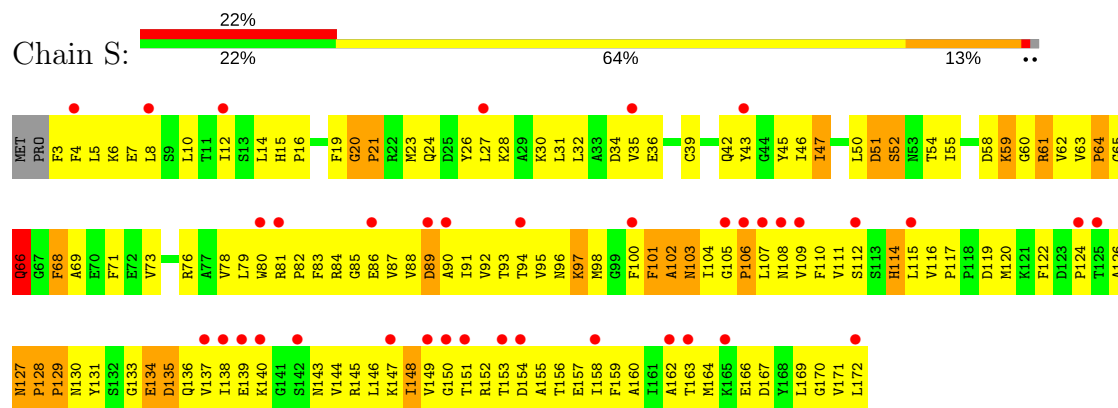
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2



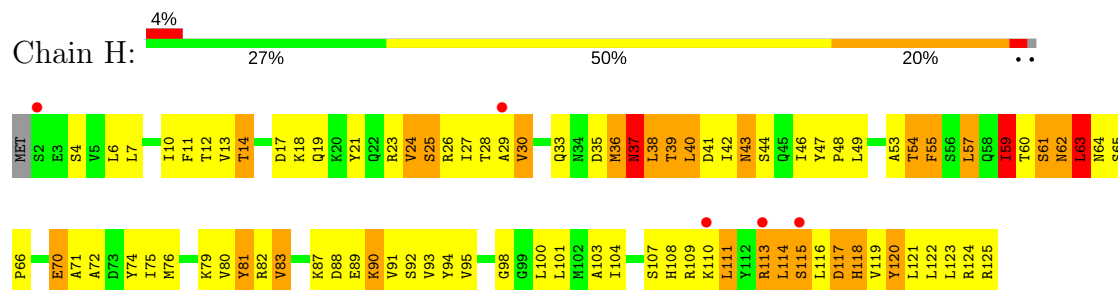
• Molecule 7: DNA-directed RNA polymerase II subunit rpb7

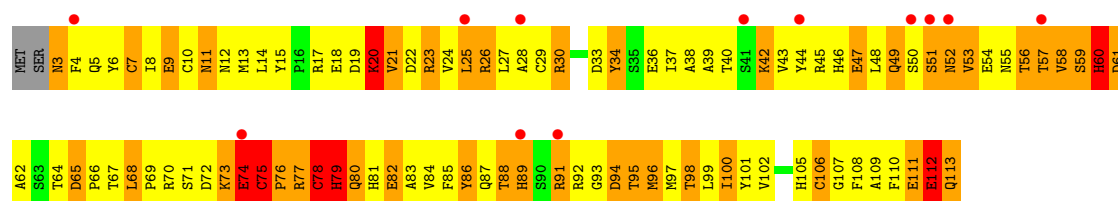


• Molecule 7: DNA-directed RNA polymerase II subunit rpb7

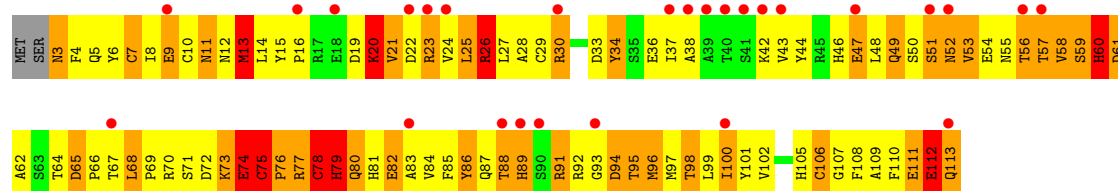


• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

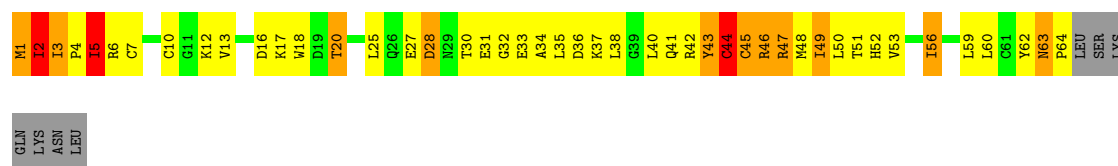




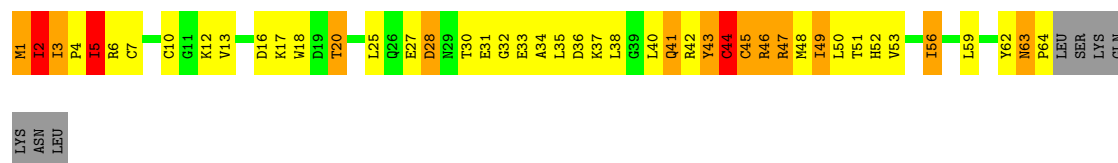
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9



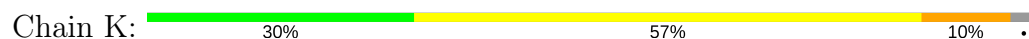
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



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



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

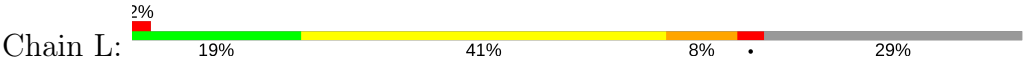


• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

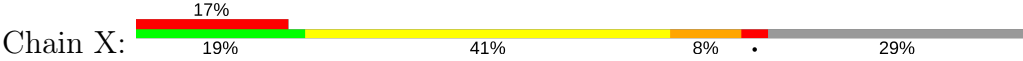




• Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4



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



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	163.03Å 202.68Å 391.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.70 – 3.65 48.69 – 3.65	Depositor EDS
% Data completeness (in resolution range)	92.8 (48.70-3.65) 96.2 (48.69-3.65)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.66 (at 3.67Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.297 , 0.321 0.289 , 0.296	Depositor DCC
R_{free} test set	6923 reflections (4.99%)	DCC
Wilson B-factor (Å ²)	125.6	Xtriage
Anisotropy	0.296	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 139.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	62870	wwPDB-VP
Average B, all atoms (Å ²)	203.0	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.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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

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	0.41	0/12026	0.57	1/16260 (0.0%)
1	M	0.42	0/11887	0.56	1/16069 (0.0%)
2	B	0.44	0/9360	0.60	5/12643 (0.0%)
2	N	0.43	0/9360	0.59	5/12643 (0.0%)
3	C	0.44	0/2135	0.60	0/2904
3	O	0.44	0/2135	0.59	0/2904
4	D	0.24	0/1049	0.38	0/1412
4	P	0.24	0/1049	0.38	0/1412
5	E	0.38	0/1695	0.60	0/2287
5	Q	0.39	0/1695	0.60	0/2287
6	F	0.50	0/666	0.67	0/901
6	R	0.50	0/666	0.67	0/901
7	G	0.26	0/1361	0.57	3/1847 (0.2%)
7	S	0.26	0/1361	0.57	3/1847 (0.2%)
8	H	0.42	0/1010	0.65	0/1363
8	T	0.42	0/1010	0.65	0/1363
9	I	0.22	0/921	0.37	0/1246
9	U	0.24	0/921	0.37	0/1246
10	J	0.57	0/526	0.77	0/709
10	V	0.57	0/526	0.76	0/709
11	K	0.47	0/972	0.61	0/1317
11	W	0.47	0/972	0.61	0/1317
12	L	0.36	0/371	0.57	0/491
12	X	0.37	0/371	0.57	0/491
All	All	0.41	0/64045	0.58	18/86569 (0.0%)

There are no bond length outliers.

The worst 5 of 18 bond angle outliers are listed below:

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	992	ALA	N-CA-C	-5.71	95.59	111.00
2	N	992	ALA	N-CA-C	-5.69	95.65	111.00
7	G	167	ASP	CB-CG-OD2	5.38	123.14	118.30
7	S	167	ASP	CB-CG-OD2	5.35	123.12	118.30
2	N	711	ASP	CB-CG-OD2	5.30	123.07	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11802	0	11784	3283	0
1	M	11666	0	11647	3212	0
2	B	9180	0	9163	1630	0
2	N	9180	0	9164	1653	0
3	C	2088	0	2045	267	0
3	O	2088	0	2045	269	0
4	D	1036	0	1025	349	0
4	P	1036	0	1025	318	0
5	E	1663	0	1684	205	0
5	Q	1663	0	1684	209	0
6	F	656	0	679	75	0
6	R	656	0	679	82	0
7	G	1330	0	1329	424	0
7	S	1330	0	1329	425	0
8	H	996	0	1006	168	0
8	T	996	0	1006	178	0
9	I	902	0	840	282	0
9	U	902	0	839	268	0
10	J	518	0	529	90	0
10	V	518	0	529	90	0
11	K	955	0	968	120	0
11	W	955	0	968	113	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	L	368	0	380	38	0
12	X	368	0	380	38	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
13	M	2	0	0	0	0
13	N	1	0	0	0	0
13	O	1	0	0	0	0
13	U	2	0	0	0	0
13	V	1	0	0	0	0
13	X	1	0	0	0	0
14	A	1	0	0	0	0
14	M	1	0	0	0	0
All	All	62870	0	62727	12749	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 102.

The worst 5 of 12749 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)
1:M:1161:PRO:HG2	1:M:1190:LYS:CG	1.29	1.62
1:A:1161:PRO:HG2	1:A:1190:LYS:CG	1.29	1.58
1:M:1161:PRO:CG	1:M:1190:LYS:HG2	1.33	1.57
1:M:1091:GLY:HA3	1:M:1092:VAL:CG1	1.35	1.56
1:M:267:ASP:CB	1:M:268:LEU:HB2	1.35	1.54

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	1494/1752 (85%)	939 (63%)	280 (19%)	275 (18%)	0	2
1	M	1472/1752 (84%)	932 (63%)	281 (19%)	259 (18%)	0	3
2	B	1142/1210 (94%)	735 (64%)	245 (22%)	162 (14%)	0	5
2	N	1142/1210 (94%)	738 (65%)	243 (21%)	161 (14%)	0	5
3	C	261/297 (88%)	178 (68%)	56 (22%)	27 (10%)	0	10
3	O	261/297 (88%)	178 (68%)	56 (22%)	27 (10%)	0	10
4	D	128/135 (95%)	86 (67%)	20 (16%)	22 (17%)	0	3
4	P	128/135 (95%)	86 (67%)	20 (16%)	22 (17%)	0	3
5	E	205/210 (98%)	137 (67%)	43 (21%)	25 (12%)	0	7
5	Q	205/210 (98%)	137 (67%)	43 (21%)	25 (12%)	0	7
6	F	81/142 (57%)	58 (72%)	15 (18%)	8 (10%)	1	11
6	R	81/142 (57%)	57 (70%)	16 (20%)	8 (10%)	1	11
7	G	168/172 (98%)	128 (76%)	24 (14%)	16 (10%)	1	12
7	S	168/172 (98%)	128 (76%)	24 (14%)	16 (10%)	1	12
8	H	122/125 (98%)	73 (60%)	27 (22%)	22 (18%)	0	2
8	T	122/125 (98%)	73 (60%)	27 (22%)	22 (18%)	0	2
9	I	109/113 (96%)	46 (42%)	27 (25%)	36 (33%)	0	0
9	U	109/113 (96%)	45 (41%)	25 (23%)	39 (36%)	0	0
10	J	62/71 (87%)	41 (66%)	15 (24%)	6 (10%)	1	12
10	V	62/71 (87%)	41 (66%)	15 (24%)	6 (10%)	1	12
11	K	117/123 (95%)	80 (68%)	24 (20%)	13 (11%)	0	8
11	W	117/123 (95%)	80 (68%)	24 (20%)	13 (11%)	0	8
12	L	43/63 (68%)	23 (54%)	12 (28%)	8 (19%)	0	2
12	X	43/63 (68%)	23 (54%)	12 (28%)	8 (19%)	0	2
All	All	7842/8826 (89%)	5042 (64%)	1574 (20%)	1226 (16%)	0	4

5 of 1226 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	10	SER
1	A	12	PRO
1	A	37	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	42	THR
1	A	50	PRO

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1301/1536 (85%)	1113 (86%)	188 (14%)	4	24
1	M	1286/1536 (84%)	1097 (85%)	189 (15%)	3	24
2	B	1012/1064 (95%)	914 (90%)	98 (10%)	9	42
2	N	1012/1064 (95%)	914 (90%)	98 (10%)	9	42
3	C	236/267 (88%)	220 (93%)	16 (7%)	18	58
3	O	236/267 (88%)	220 (93%)	16 (7%)	18	58
4	D	111/115 (96%)	94 (85%)	17 (15%)	3	22
4	P	111/115 (96%)	94 (85%)	17 (15%)	3	22
5	E	182/184 (99%)	169 (93%)	13 (7%)	17	56
5	Q	182/184 (99%)	169 (93%)	13 (7%)	17	56
6	F	71/121 (59%)	64 (90%)	7 (10%)	9	41
6	R	71/121 (59%)	64 (90%)	7 (10%)	9	41
7	G	146/148 (99%)	139 (95%)	7 (5%)	30	67
7	S	146/148 (99%)	139 (95%)	7 (5%)	30	67
8	H	113/114 (99%)	99 (88%)	14 (12%)	5	31
8	T	113/114 (99%)	99 (88%)	14 (12%)	5	31
9	I	103/105 (98%)	80 (78%)	23 (22%)	1	7
9	U	103/105 (98%)	80 (78%)	23 (22%)	1	7
10	J	59/66 (89%)	46 (78%)	13 (22%)	1	8
10	V	59/66 (89%)	46 (78%)	13 (22%)	1	8
11	K	109/113 (96%)	104 (95%)	5 (5%)	31	68
11	W	109/113 (96%)	103 (94%)	6 (6%)	25	64

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
12	L	39/53 (74%)	34 (87%)	5 (13%)	5	29
12	X	39/53 (74%)	34 (87%)	5 (13%)	5	29
All	All	6949/7772 (89%)	6135 (88%)	814 (12%)	6	33

5 of 814 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	I	47	GLU
1	M	262	SER
7	S	114	HIS
9	I	94	ASP
1	M	48	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 225 such sidechains are listed below:

Mol	Chain	Res	Type
8	H	33	GLN
1	M	395	GLN
6	R	94	ASN
9	I	55	ASN
1	M	48	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 18 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, 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	1496/1752 (85%)	0.16	87 (5%) 24 16	51, 150, 385, 628	0
1	M	1476/1752 (84%)	0.53	174 (11%) 5 5	83, 199, 449, 672	0
2	B	1150/1210 (95%)	0.05	26 (2%) 61 46	69, 143, 314, 580	0
2	N	1150/1210 (95%)	0.55	119 (10%) 7 6	85, 222, 419, 618	0
3	C	263/297 (88%)	-0.18	0 100 100	80, 126, 252, 477	0
3	O	263/297 (88%)	0.09	3 (1%) 80 67	125, 190, 350, 511	0
4	D	130/135 (96%)	0.33	5 (3%) 41 29	128, 240, 360, 596	0
4	P	130/135 (96%)	0.55	9 (6%) 18 11	152, 287, 414, 545	0
5	E	207/210 (98%)	-0.15	4 (1%) 67 52	70, 164, 293, 516	0
5	Q	207/210 (98%)	-0.04	4 (1%) 67 52	119, 181, 373, 501	0
6	F	83/142 (58%)	-0.49	0 100 100	70, 97, 170, 232	0
6	R	83/142 (58%)	-0.34	1 (1%) 79 66	111, 126, 200, 335	0
7	G	170/172 (98%)	0.36	9 (5%) 27 18	92, 199, 323, 514	0
7	S	170/172 (98%)	1.08	38 (22%) 1 1	122, 240, 388, 494	0
8	H	124/125 (99%)	0.14	5 (4%) 39 26	82, 138, 298, 408	0
8	T	124/125 (99%)	0.60	15 (12%) 5 4	115, 197, 344, 429	0
9	I	111/113 (98%)	0.51	12 (10%) 6 5	98, 239, 367, 549	0
9	U	111/113 (98%)	1.64	27 (24%) 1 1	151, 330, 482, 585	0
10	J	64/71 (90%)	-0.20	0 100 100	85, 108, 220, 304	0
10	V	64/71 (90%)	0.05	0 100 100	137, 203, 326, 373	0
11	K	119/123 (96%)	-0.33	0 100 100	60, 126, 221, 319	0
11	W	119/123 (96%)	0.15	3 (2%) 58 43	76, 168, 295, 470	0
12	L	45/63 (71%)	-0.01	1 (2%) 62 47	93, 177, 299, 481	0
12	X	45/63 (71%)	1.01	11 (24%) 1 1	153, 277, 425, 580	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
All	All	7904/8826 (89%)	0.29	553 (6%) 17 11	51, 181, 397, 672	0

The worst 5 of 553 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
9	U	41	SER	21.7
9	U	40	THR	16.1
9	U	39	ALA	14.8
11	W	119	GLU	14.4
1	M	1239	LEU	12.2

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 carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 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	LLDF	B-factors(Å ²)	Q<0.9
13	ZN	A	2457	1/1	0.94	0.20	-0.36	122,122,122,122	0
13	ZN	N	2225	1/1	0.80	0.22	-0.76	180,180,180,180	0
13	ZN	M	2456	1/1	0.64	0.25	-0.85	432,432,432,432	0
13	ZN	B	2225	1/1	0.94	0.17	-0.94	149,149,149,149	0
13	ZN	I	1122	1/1	0.83	0.10	-1.26	126,126,126,126	0
13	ZN	C	1269	1/1	0.91	0.07	-1.32	122,122,122,122	0
13	ZN	L	1071	1/1	0.87	0.09	-1.38	123,123,123,123	0
13	ZN	M	2457	1/1	0.82	0.14	-1.39	172,172,172,172	0
13	ZN	J	1066	1/1	0.97	0.23	-1.44	87,87,87,87	0
13	ZN	X	1071	1/1	0.84	0.05	-1.81	195,195,195,195	0
13	ZN	A	2456	1/1	0.68	0.06	-1.92	245,245,245,245	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
13	ZN	V	1066	1/1	0.95	0.20	-1.98	155,155,155,155	0
13	ZN	I	1121	1/1	0.83	0.06	-2.19	160,160,160,160	0
13	ZN	O	1269	1/1	0.69	0.11	-2.22	189,189,189,189	0
13	ZN	U	1121	1/1	0.89	0.06	-2.58	284,284,284,284	0
13	ZN	U	1122	1/1	0.80	0.05	-3.16	240,240,240,240	0
14	MG	A	2458	1/1	0.55	1.80	-	109,109,109,109	0
14	MG	M	2458	1/1	0.28	0.39	-	197,197,197,197	0

6.5 Other polymers [i](#)

There are no such residues in this entry.