



# wwPDB X-ray Structure Validation Summary Report i

Feb 26, 2014 – 07:22 PM GMT

PDB ID : 2BE5  
Title : Crystal structure of the T. Thermophilus RNA polymerase holoenzyme in complex with inhibitor tagetitoxin  
Authors : Vassylyev, D.G.; Svetlov, V.; Vassylyeva, M.N.; Perederina, A.; Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Artsimovitch, I.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)  
Deposited on : 2005-10-22  
Resolution : 2.40 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

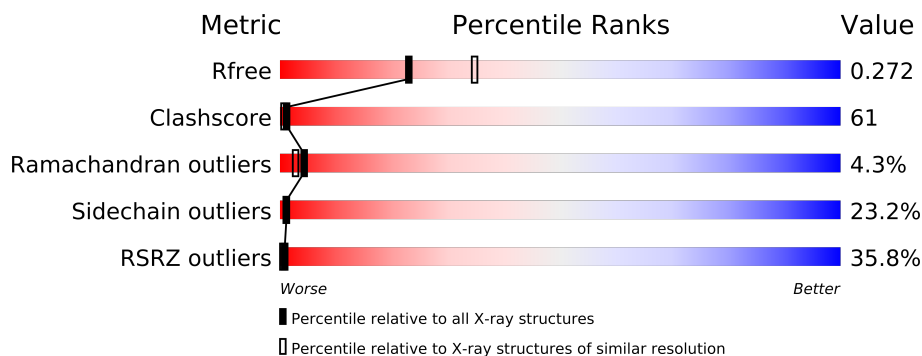
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.40 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))
$R_{free}$	66092	2207 (2.40-2.40)
Clashscore	79885	2789 (2.40-2.40)
Ramachandran outliers	78287	2736 (2.40-2.40)
Sidechain outliers	78261	2737 (2.40-2.40)
RSRZ outliers	66119	2210 (2.40-2.40)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	315	
1	B	315	
1	K	315	
1	L	315	
2	C	1119	
2	M	1119	
3	D	1524	
3	N	1524	
4	E	99	
4	O	99	
5	F	423	
5	P	423	

The following table lists non-polymeric compounds that are outliers for geometric or electron-

density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
8	TGT	N	9002	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 61800 atoms, of which 0 are hydrogen and 0 are deuterium.

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 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	B	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	K	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	L	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			

- Molecule 2 is a protein called DNA-directed RNA polymerase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			
2	M	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			

- Molecule 3 is a protein called DNA-directed RNA polymerase beta' chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1392	Total	C	N	O	S	0	0	0
			10797	6819	1925	2020	33			
3	N	1392	Total	C	N	O	S	0	0	0
			10797	6819	1925	2020	33			

- Molecule 4 is a protein called RNA polymerase omega chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	95	Total	C	N	O	S	0	0	0
			769	488	133	144	4			
4	O	95	Total	C	N	O	S	0	0	0
			769	488	133	144	4			

- Molecule 5 is a protein called RNA polymerase sigma factor rpoD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	345	Total	C	N	O	S	0	0	0
			2771	1744	504	519	4			
5	P	345	Total	C	N	O	S	0	0	0
			2771	1744	504	519	4			

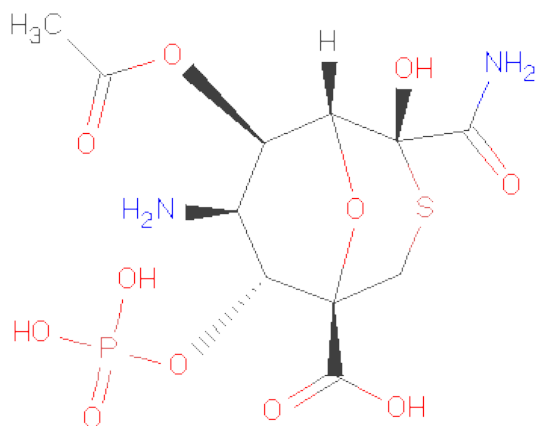
- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	1	Total	Mg	0	0
			1	1		
6	C	1	Total	Mg	0	0
			1	1		
6	N	2	Total	Mg	0	0
			2	2		

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	2	Total	Zn	0	0
			2	2		
7	N	2	Total	Zn	0	0
			2	2		

- Molecule 8 is TAGETITOXIN (three-letter code: TGT) (formula: C<sub>11</sub>H<sub>17</sub>N<sub>2</sub>O<sub>11</sub>PS).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
8	D	1	Total 26	C 11	N 2	O 11	P 1	S 1	0	0
8	N	1	Total 26	C 11	N 2	O 11	P 1	S 1	0	0

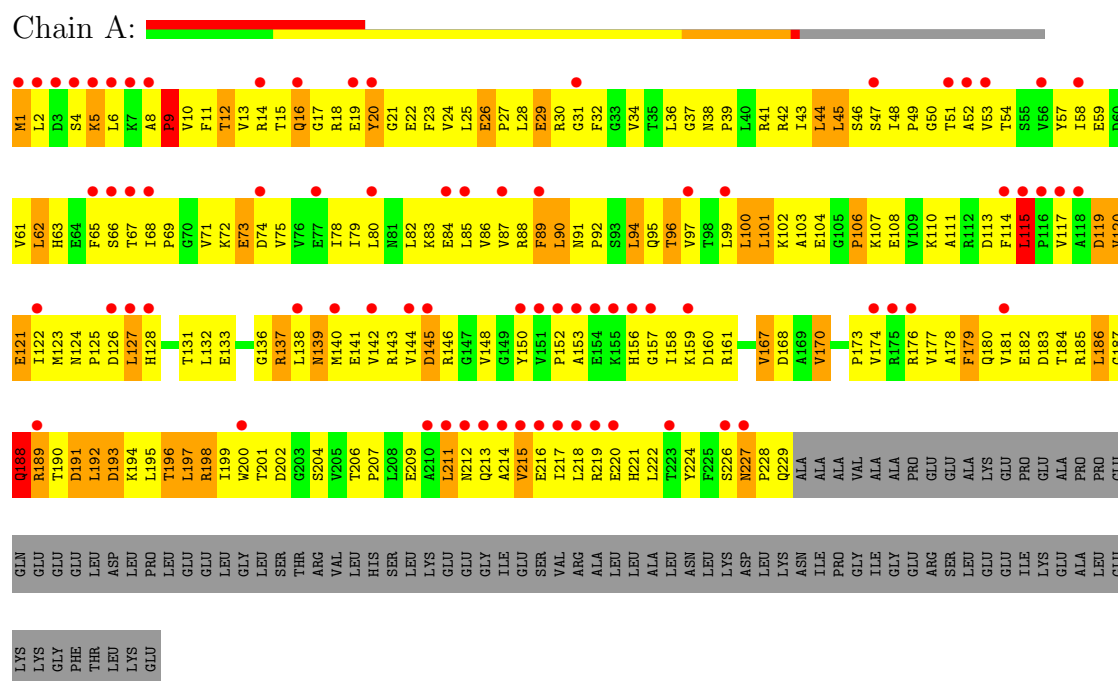
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	250	Total 250	O 250	0	0
9	B	329	Total 329	O 329	0	0
9	C	1321	Total 1321	O 1321	0	0
9	D	1655	Total 1655	O 1655	0	0
9	E	176	Total 176	O 176	0	0
9	F	519	Total 519	O 519	0	0
9	K	278	Total 278	O 278	0	0
9	L	309	Total 309	O 309	0	0
9	M	1236	Total 1236	O 1236	0	0
9	N	1552	Total 1552	O 1552	0	0
9	O	137	Total 137	O 137	0	0
9	P	422	Total 422	O 422	0	0

### 3 Residue-property plots

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

#### • Molecule 1: DNA-directed RNA polymerase alpha chain

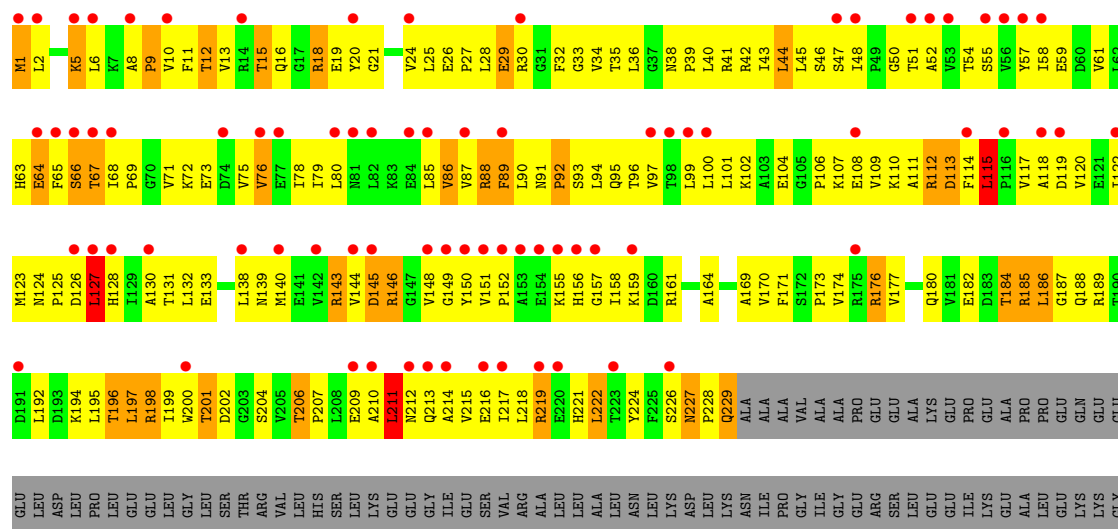


PRO LEU GLU GLU LEU LEU SER THR ARG VAL LEU HIS SER LEU LYS GLU GLU GLY ILE GLU SER VAL ARG ALA LEU LEU ALA LEU ASN LYS ASP LEU LYS LYS ASN ILE PRO GLY ILE GLY GLU ARG SER LEU GLU ILE LYS ALA LEU LYS GLY PHE THR LYS

GLU

• Molecule 1: DNA-directed RNA polymerase alpha chain

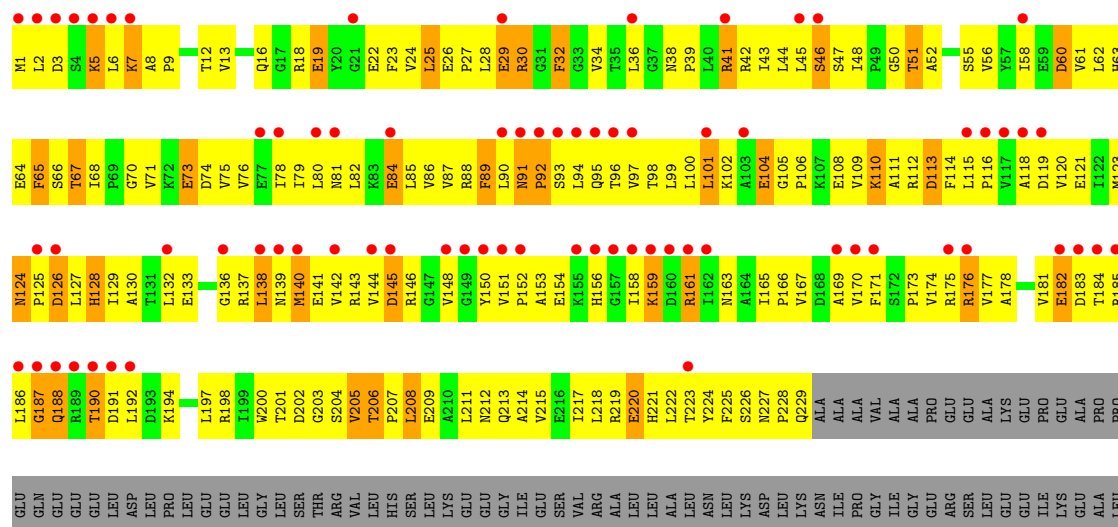
Chain K: 



PHE THR LEU LYS GLU

• Molecule 1: DNA-directed RNA polymerase alpha chain

Chain L: 

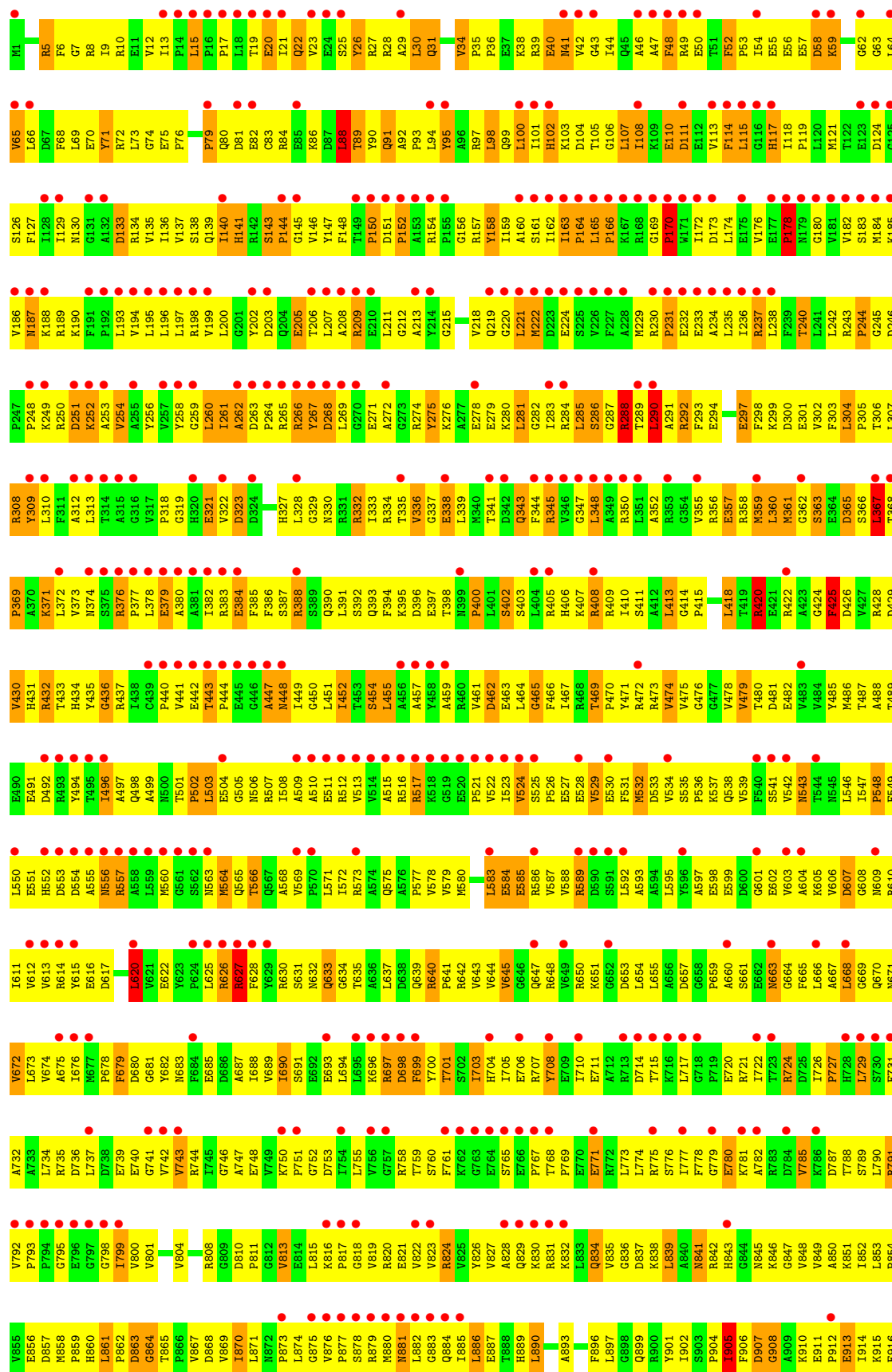


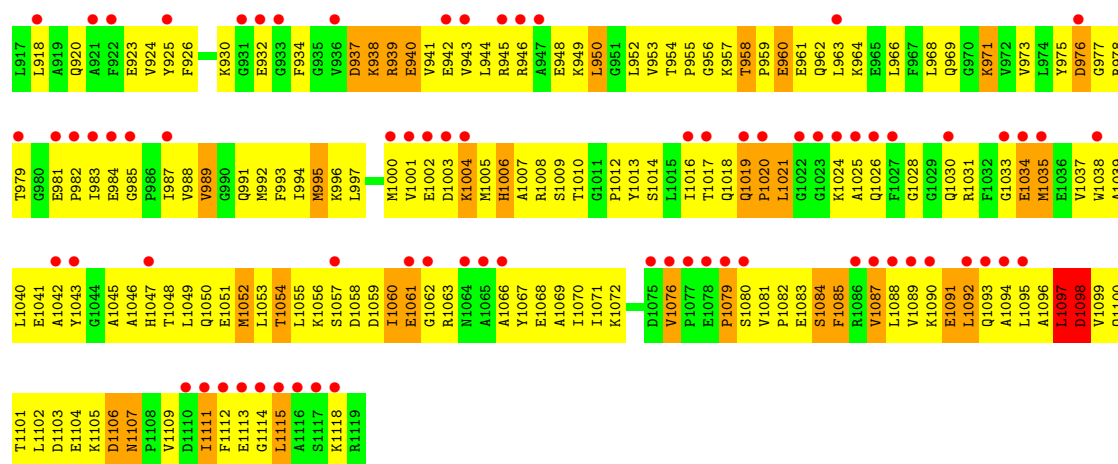
GLU LYS LYS GLY PHE THR LEU LYS GLU

• Molecule 2: DNA-directed RNA polymerase beta chain



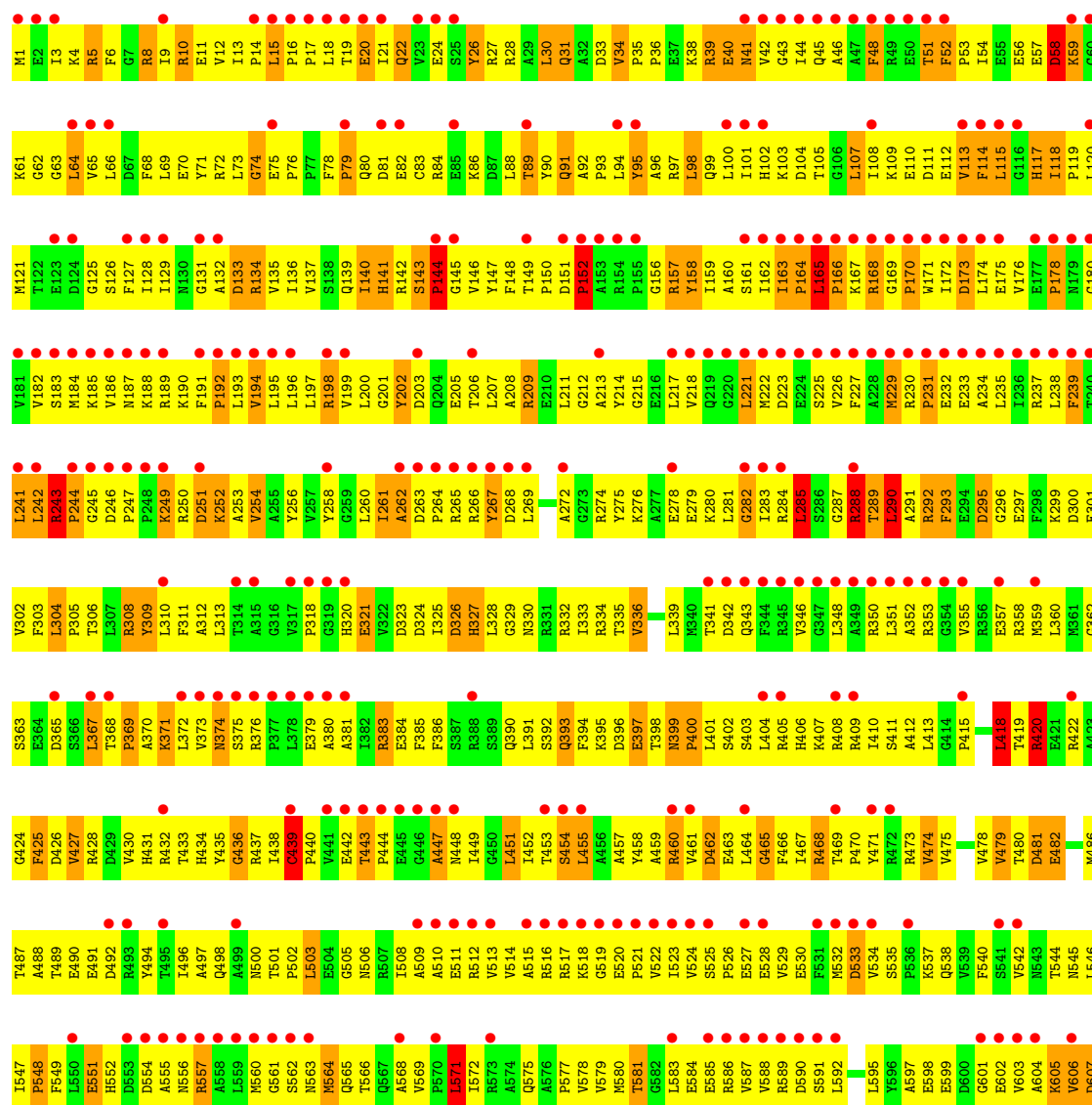
## Chain C:





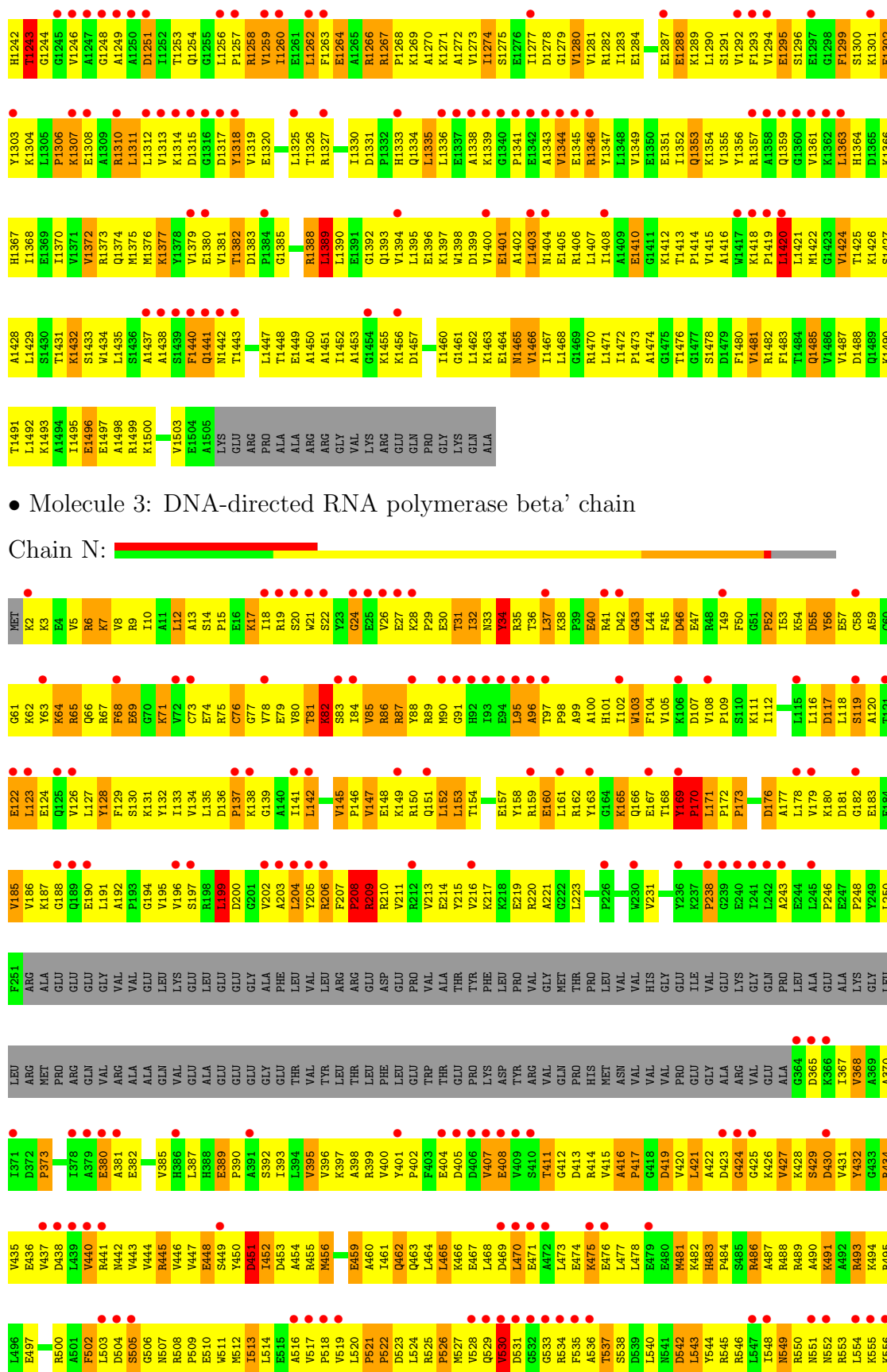
### • Molecule 2: DNA-directed RNA polymerase beta chain

Chain M:





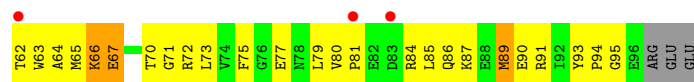
Y1117	P1056	I992	T927	V864	L804	G742	Q680	L618	L568	E497	E436	E374	ALA
I1118	V1057	L993	A928	T865	E805	D743	R681	L619	A559	V498	V437	V377	ALA
S1119	R1058	Q994	R929	T866	F806	Q744	D682	G620	A560	V499	D438	I378	ALA
V1120	S1059	L995	R930	R867	F806	Q745	I683	G621	A561	G561	D439	R379	ALA
P1121	S1060	W996	L931	R868	T808	W746	D685	V623	A562	F501	R440	A380	ALA
P1187	F1061			M869	P809	W747	D686	V624	P663	F502	R441	A381	ALA
F1123	R1062			G870	E810	W748	V687	D624	E564	L503	N442	A381	ALA
Q1124	I1063			K871	E811	W749	V688	S626	S505	S504	V443	E382	ALA
P1125	G1064			L872	A812	P750	W688	G627	S506	G506	R445	V385	ALA
D1126	L1065			L873	L813	L751	D689	R628	R567	N507	R446	H386	ALA
E1127	T1066			E874	A814	S752	A690	S629	R568	G507	V447	L387	ALA
V1128	Q1067			T875	A815	F753	L691	S630	N569	P509	E448	H388	ALA
T1129	L1068			S876	H816	W754	E692	H390	E570	E510	S449	E389	ALA
R1130	E1069			P877	E817	A755	E893	I631	K671	E510	S449	E389	ALA
S1131	Y1070			G878	R818	Q756	V694	V632	R572	W511	Y450	P390	ALA
L1132	F1071			R879	G819	A757	I695	V633	M573	M512	D451	A391	ALA
G1133	I1072			I880	E820	E758	H696	G634	L574	I513	I452	S392	ALA
R1135	S1073			L881	W821	A759	D697	P635	G697	A516	E459	K397	ALA
K1136	S1074			F882	A822	W760	L698	G636	E576	A517	A460	R399	ALA
K1201	H1075			A883	L823	I761	V699	L637	A577	V517	R455	V396	ALA
Q1202	G1076			R884	R824	Q762	V700	A638	V578	P518	M456	V396	ALA
K1203	A1077			I885	A825	W763	L701	H639	D579	V519	K397	A398	ALA
D1139	R1078			V886	P826	L764	L702	H640	A580	L520	I452	S392	ALA
E1141	K1079				R827	A765	N703	Q641	L581	P521	L394	I393	ALA
				V890	K828	A766	R704	G642	L582	Q522	A460	R399	ALA
				E893	W829	H767	A705	G643	D583	D523	Q462	Y401	ALA
				E894	A830	W768	P706	L644	N584	L524	Q463	F402	ALA
				V895	G831	L769	T707	P645	G585	R525	L464	F403	ALA
				R896	R832	L770	L708	K646	R586	P526	L465	E404	ALA
				R897	E833	S771	H709	R647	R587	M527	K466	D405	ALA
				E898	T834	P772	R710	M648	G588	V528	E467	D406	ALA
				E899	S835	A773	L711	A649	A889	Q529	L468	Y407	ALA
				R899	R836	S774	L715	L650	P590	V530	E469	E408	ALA
				I900	G837	G775	F716	B651	V591	D531	L470	V409	ALA
				Q901	R838	E776	Q717	L652	T592	G533	E471	S410	ALA
				L902	L839	P777	P718	P653	N593	G533	L472	T411	ALA
				D903	K840	L778	P719	P655	P594	R534	L473	G412	ALA
				V904	Y841		V719	P655	G595	F535	E474	D413	ALA
				P905	F842		L720	F656	S596	A536	K475	R414	ALA
				Q906	F843		V721	L657	D597	T537	E476	V415	ALA
				K908	R844		E722	L658	R598	S538	L477	A416	ALA
				S909	R845		G723	K659	P599	D539	L478	P417	ALA
				L910	D846		Q724	M600	L600	N541	E479	G418	ALA
				L911	E847		S725	M601	R601	N541	E480	D419	ALA
				K912	E848		I726	M602	S602	D542	M481	V420	ALA
				D913	A849		Q727	E663	L603	L543	K482	L421	ALA
				L914	L851		H729		T604	Y544	H483	A422	ALA
				V915	A852		P730	I666	D605	R546	P484	D423	ALA
				Y916	R853		L731	A667	L606	R546	S485	G424	ALA
				Q917	E854		V732	P668	L607	L547	R486	G425	ALA
				F919	R855		C733	M669	S608	I548	A487	K426	ALA
				L920	G856		E734	V670	G609	N549	R488	V427	ALA
				R921	R857		A735	K671	K610	R550	R489	K428	ALA
				T984	V858		K796	A672	Q611	N551	R490	S429	ALA
				Y984	V859		K797		G612	R552	K491	D430	ALA
				E987	D859		W737	R675	R613	R553	A492	V431	ALA
				R988	L860		K799	M676	F614	L564	R493	Y432	ALA
				Y989	Q861		R800	L677	R615	K555	K494	G433	ALA
					D862		F740	E678	Q616	K566	R494	A370	ALA
					V863		G803	R679	N617	L567	L496	V435	ALA



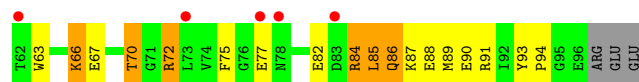
A1358	Q1359	G1360	V1361	K1362	L1363	H1364	D1365	K1366	G1367	I1368	A1369	I1370	V1371	V1372	R1373	R1374	Q1375	K1376	K1377	Y1378	V1379	E1380	V1381	T1382	D1383	P1384	G1385	S1386	S1387	R1388	L1389	L1390	E1391	Q1392	Q1393	V1394	L1395	E1396	K1397	D1398	V1400	E1401	A1402	L1403	N1404	E1405	R1406	L1407	L1408	A1409	E1410	G1411	G1412	T1413	P1414	V1415	A1416	W1417			
Q1235	L1236	T1237	M1238	R1239	K1240	F1241	H1242	T1243	G1244	E1245	V1246	A1247	G1248	A1249	A1250	P1251	I1252	K1253	Q1254	P1257	R1258	E1259	I1260	E1261	L1262	F1263	A1264	L1265	R1266	R1267	P1268	K1269	L1270	S1271	K1272	M1273	A1274	E1275	E1276	I1277	D1278	G1279	V1280	R1281	R1282	E1283	E1284	E1285	T1286	E1287	K1288	E1289	G1290	E1291	V1292	T1293	V1294	E1295			
S1296	E1297	G1298	L1299	K1300	K1301	E1302	Y1303	K1304	P1305	E1306	K1307	E1308	A1309	R1310	L1311	L1312	V1313	K1314	D1315	Q1316	D1317	Y1318	E1319	E1320	Q1321	P1322	L1323	T1324	L1325	R1326	R1327	T1328	P1329	Q1330	D1331	Q1332	Q1333	Q1334	L1335	L1336	E1337	A1338	K1339	G1340	P1341	E1342	A1343	V1344	E1345	R1346	Y1347	L1348	E1349	E1350	E1351	L1352	Q1353	K1354	V1355	R1356	L1357
C1112	G1113	T1114	T1115	N1116	N1117	T1118	S1119	V1120	P1121	L1122	F1123	A1124	P1125	D1126	E1127	V1128	T1129	R1130	R1133	L1134	R1135	K1136	L1137	A1138	L1139	T1140	E1141	A1142	G1143	L1144	Y1145	G1146	R1147	V1148	L1149	A1150	R1151	E1152	V1153	E1154	V1155	L1156	G1157	V1158	R1159	L1160	E1161	E1162	G1163	R1164	V1165	L1166	S1167	L1168	D1169	P1170	A1171	H1172			
L1173	L1174	K1175	L1176	A1177	A1178	E1182	T1183	Q1184	E1185	V1186	P1187	V1188	R1189	S1190	P1191	L1192	T1193	Q1194	Q1195	T1196	R1197	Y1198	G1199	V1200	C1201	Q1202	K1203	C1204	Y1205	G1206	Y1207	L1208	L1209	S1210	M1211	A1212	R1213	P1214	S1215	S1216	I1217	G1218	E1219	A1220	V1221	G1222	I1223	V1224	A1225	A1226	K1227	S1228	I1229	G1230	E1231	P1232	G1233	T1234			
Q1235	L1236	M1237	M1238	R1239	K1240	F1241	H1242	T1243	G1244	E1245	V1246	A1247	G1248	A1249	A1250	P1251	I1252	K1253	Q1254	P1257	R1258	E1259	I1260	E1261	L1262	F1263	A1264	L1265	R1266	R1267	P1268	K1269	L1270	S1271	K1272	M1273	A1274	E1275	E1276	I1277	D1278	G1279	V1280	R1281	R1282	E1283	E1284	E1285	T1286	E1287	K1288	E1289	G1290	E1291	V1292	T1293	V1294	E1295			
S1296	E1297	G1298	L1299	K1300	K1301	E1302	Y1303	K1304	P1305	E1306	K1307	E1308	A1309	R1310	L1311	L1312	V1313	K1314	D1315	Q1316	D1317	Y1318	E1319	E1320	Q1321	P1322	L1323	T1324	L1325	R1326	R1327	T1328	P1329	Q1330	D1331	Q1332	Q1333	Q1334	L1335	L1336	E1337	A1338	K1339	G1340	P1341	E1342	A1343	V1344	E1345	R1346	Y1347	L1348	E1349	E1350	E1351	L1352	Q1353	K1354	V1355	R1356	L1357
A1618	L619	G620	K621	R622	P623	E624	Y625	R628	S629	R630	E631	V632	K631	V632	V633	G634	P635	K636	L637	K638	L639	L702	N703	Q641	C642	G643	L644	P645	E646	E647	R648	L711	L712	L713	Q714	P715	F716	Q717	P718	V719	K659	K660	P599	L600	R601	S602	L603	T604	D605	I606	V607	S608	G609	K610	Q611	G612	F614	R615	Q616	Q690	
R681	D682	K683	K684	E685	E686	V687	W688	D689	A690	L691	E692	V693	L694	V695	V696	G697	K698	V699	V700	L701	L702	N703	Q641	C642	G643	L644	P645	E646	E647	R648	L711	L712	L713	Q714	P715	F716	Q717	P718	V719	K659	K660	P599	L600	R601	S602	L603	T604	D605	I606	V607	S608	G609	K610	Q611	G612	F614	R615	Q616	Q690		
D743	Q744	M745	A746	V747	H748	V749	P750	L751	S752	S753	F754	Q755	Q756	A757	E758	A759	R760	L761	Q762	M763	L764	S765	A766	H767	N768	L769	L770	P771	S772	A773	G774	G775	L776	K780	S781	S782	R783	D784	L785	L786	L787	G788	L789	P790	S791	L792	T793	Q794	V795	R796	D797	K798	L799	G800	A802						
G803	L804	F806	R807	T808	P809	E810	E811	L812	L813	E817	E820	V821	A822	L823	R824	A825	P826	L827	K828	G829	V829	A830	G831	R832	E833	T834	S835	V836	C837	R838	L839	K840	H841	V842	F843	K844	R845	P846	D847	R848	R849	L850	L851	A852	V853	H854	H855	G856	L857	V858	R859	L860	R861	D862	V863	M864	T865				
V866	R867	Y868	M869	G870	K871	L872	L873	E874	S876	R877	G878	R879	L880	L881	F882	R883	R884	L885	R886	A887	E888	A889	V890	E891	D892	E893	K894	V895	P896	L897	Q901	L902	D903	P904	P905	Q906	E907	K908	N909	S910	L911	K912	L913	D913	L914	V915	Y916	Q917	A918	F919	L920	R921	L922	G923	M924	E925					
K926	T927	A928	R929	L930	L931	D932	A933	Y937	Y936	E937	T940	F941	S942	T943	T944	S945	G946	T947	T948	I949	G950	I951	D952	D953	A954	A955	I956	P957	E958	E959	R960	L964	E965	E966	A967	D968	R969	K970	L971	L972	L973	Q974	E975	Q976	A977	Y978	E979	M980	G981	F982	D985	R986	E987	R988							
Y989	D990	Q991	E992	L993	Q994	L995	W996	T997	Y997	E998	T999	K1000	K1002	V1003	T1004	K1005	A1006	A1007	F1008	K1009	N1010	F1011	E1012	E1013	N1014	F1017	N1018	P1019	L1020	Y1021	V1022	M1023	S1026	G1027	A1028	K1029	G1030	N1031	P1032	Q1033	L1034	I1035	R1036	Q1037	L1038	G1039	L1040	L1041	R1042	G1043	L1044	M1045	Q1046	K1047	P1048	S1049	G1050				
E1051	T1052	F1053	E1054	V1055	P1056	V1057	L1058	S1059	S1060	F1061	R1062	E1063	G1064	L1065	T1066	V1067	L1068	E1069	Y1070	F1071	I1072	S1073	S1074	H1075	G1076	A1077	G1080	E1081	P1082	L1083	T1084	A1085	L1086	R1087	T1088	A1089	D1090	S1091	G1092	Y1093	L1094	T1095	A1096	E1097	L1098	V1099	E1100	V1101	T1102	H1103	E1104	L1105	V1106	V1107	A1108	E1109	A1110	D1111			
G1112	G1113	T1114	T1115	N1116	N1117	T1118	S1119	V1120	P1121	L1122	F1123	A1124	P1125	D1126	E1127	V1128	T1129	R1130	R1133	L1134	R1135	K1136	L1137	A1138	L1139	T1140	E1141	A1142	G1143	L1144	Y1145	G1146	R1147	V1148	L1149	A1150	R1151	E1152	V1153	E1154	V1155	L1156	G1157	V1158	R1159	L1160	E1161	E1162	G1163	R1164	V1165	L1166	S1167	L1168	D1169	P1170	A1171	H1172			
L1173	L1174	K1175	L1176	A1177	A1178	E1182	T1183	Q1184	E1185	V1186	P1187	V1188	R1189	S1190	P1191	L1192	T1193	Q1194	Q1195	T1196	R1197	Y1198	G1199	V1200	C1201	Q1202	K1203	C1204	Y1205	G1206	Y1207	L1208	L1209	S1210	M1211	A1212	R1213	P1214	S1215	S1216	I1217	G1218	E1219	A1220	V1221	G1222	I1223	V1224	A1225	A1226	K1227	S1228	I1229	G1230	E1231	P1232	G1233	T1234			
Q1235	L1236	M1237	M1238	R1239	K1240	F1241	H1242	T1243	G1244	E1245	V1246	A1247	G1248	A1249	A1250	P1251	I1252	K1253	Q1254	P1257	R1258	E1259	I1260	E1261	L1262	F1263	A1264	L1265	R1266	R1267	P1268	K1269	L1270	S1271	K1272	M1273	A1274	E1275	E1276	I1277	D1278	G1279	V1280	R1281	R1282	E1283	E1284	E1285	T1286	E1287	K1288	E1289	G1290	E1291	V1292	T1293	V1294	E1295			
S1296	E1297	G1298	L1299	K1300	K1301	E1302	Y1303	K1304	P1305	E1306	K1307	E1308	A1309	R1310	L1311	L1312	V1313	K1314	D1315	Q1316	D1317	Y1318	E1319	E1320	Q1321	P1322	L1323	T1324	L1325	R1326	R1327	T1328	P1329	Q1330	D1331	Q1332	Q1333	Q1334	L1335	L1336	E1337	A1338	K1339	G1340	P1341	E1342	A1343	V1344	E1345	R1346	Y1347	L1348	E1349	E1350	E1351	L1352	Q1353	K1354	V1355	R1356	L1357



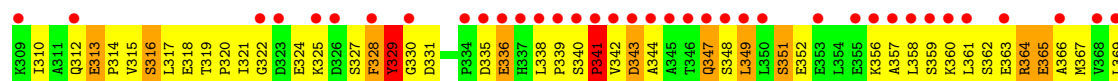
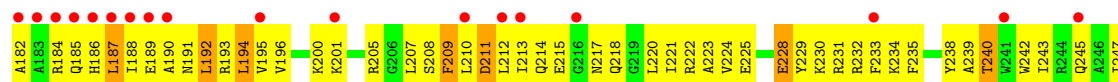
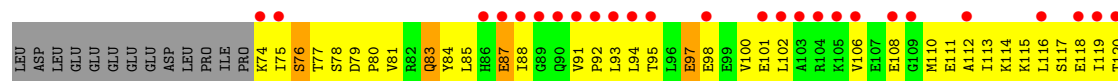
Chain E:

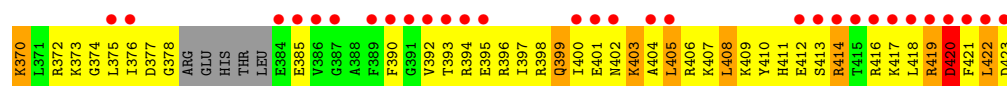


Chain Q:



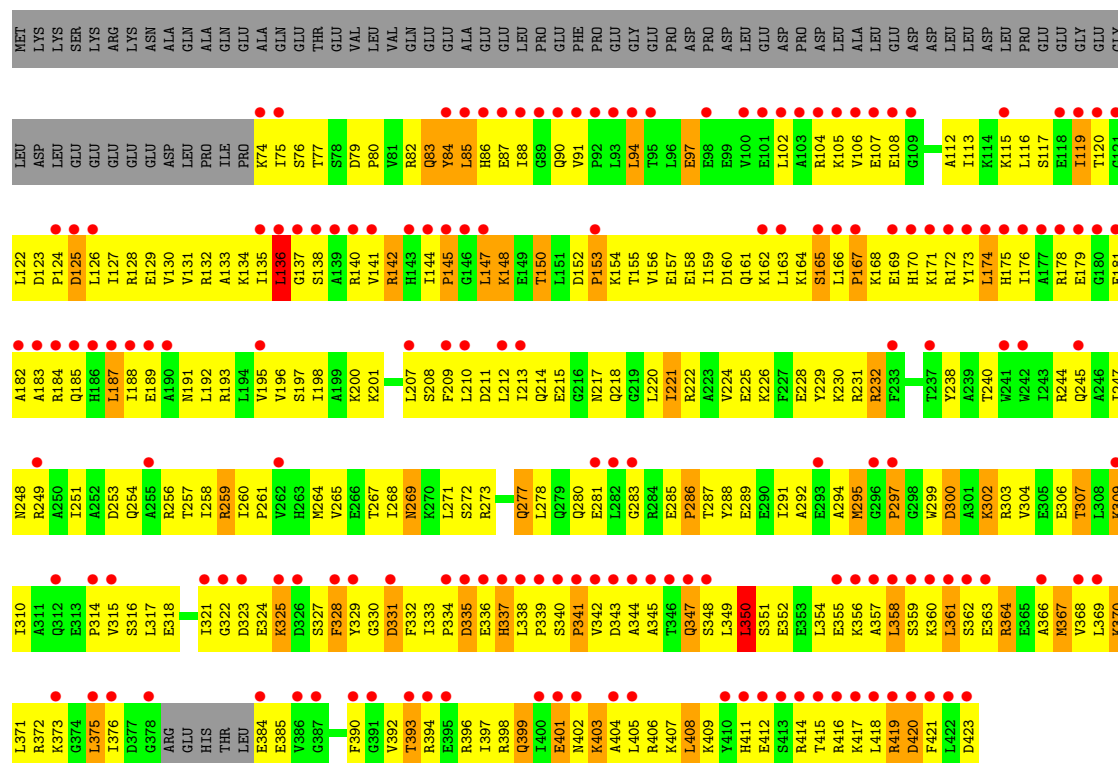
Chain F: 





● Molecule 5: RNA polymerase sigma factor rpoD

Chain P:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	239.50Å 239.50Å 253.10Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.00 – 2.40 36.81 – 2.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.40) 95.2 (36.81-2.40)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.65 (at 2.39Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.237 , 0.274 0.235 , 0.272	Depositor DCC
$R_{free}$ test set	34795 reflections (6.11%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.7	Xtriage
Anisotropy	0.106	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.07 , -20.0	EDS
Estimated twinning fraction	0.500 for H, K, L 0.500 for -H, -K, L 0.499 for -h,-k,l 0.065 for h,-h-k,-l 0.065 for -k,-h,-l	Xtriage
Reported twinning fraction	0.500 for H, K, L 0.500 for -H, -K, L	Depositor
L-test for twinning	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	0 of 604645 reflections	Xtriage
$F_o, F_c$ correlation	0.69	EDS
Total number of atoms	61800	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.81% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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, TGT

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	0.77	0/1838	0.86	3/2498 (0.1%)
1	B	0.70	0/1838	0.83	4/2498 (0.2%)
1	K	0.76	0/1838	0.85	4/2498 (0.2%)
1	L	0.73	0/1838	0.76	0/2498
2	C	0.81	0/8997	0.89	8/12164 (0.1%)
2	M	0.80	2/8997 (0.0%)	0.89	12/12164 (0.1%)
3	D	0.82	0/10975	0.92	21/14836 (0.1%)
3	N	0.80	1/10975 (0.0%)	0.92	17/14836 (0.1%)
4	E	0.80	0/783	0.94	0/1054
4	O	0.81	0/783	0.92	0/1054
5	F	0.71	0/2812	0.81	1/3781 (0.0%)
5	P	0.72	0/2812	0.78	2/3781 (0.1%)
All	All	0.79	3/54486 (0.0%)	0.89	72/73662 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	N	733	CYS	CB-SG	-5.54	1.72	1.81
2	M	202	TYR	CD2-CE2	5.05	1.47	1.39
2	M	682	TYR	CD2-CE2	5.02	1.46	1.39

The worst 5 of 72 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	138	LEU	CA-CB-CG	10.11	138.56	115.30
3	N	199	LEU	CA-CB-CG	-8.78	95.11	115.30
2	M	557	ARG	NE-CZ-NH2	7.73	124.17	120.30
3	D	199	LEU	CA-CB-CG	-7.64	97.72	115.30
3	N	1389	LEU	CA-CB-CG	7.54	132.65	115.30

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1806	0	1861	232	0
1	B	1806	0	1861	217	0
1	K	1806	0	1861	195	0
1	L	1806	0	1861	216	0
2	C	8829	0	8933	1248	0
2	M	8829	0	8933	1139	0
3	D	10797	0	10873	1481	0
3	N	10797	0	10873	1398	0
4	E	769	0	775	101	0
4	O	769	0	775	98	0
5	F	2771	0	2844	350	0
5	P	2771	0	2844	345	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	N	2	0	0	0	0
7	D	2	0	0	0	0
7	N	2	0	0	0	0
8	D	26	0	15	3	0
8	N	26	0	14	1	0
9	A	250	0	0	46	0
9	B	329	0	0	67	0
9	C	1321	0	0	266	0
9	D	1655	0	0	324	0
9	E	176	0	0	32	0
9	F	519	0	0	103	0
9	K	278	0	0	43	0
9	L	309	0	0	68	0
9	M	1236	0	0	259	0
9	N	1552	0	0	306	0
9	O	137	0	0	23	0
9	P	422	0	0	84	0
All	All	61800	0	54323	6611	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 61.

The worst 5 of 6611 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:M:409:ARG:HA	2:M:454:SER:HA	1.20	1.15
3:D:1045:MET:HG2	3:D:1073:SER:HA	1.33	1.10
3:D:119:SER:HB2	3:D:123:LEU:H	1.23	1.04
2:C:987:ILE:HG23	3:D:948:THR:HG21	1.41	1.02
2:C:457:ALA:HB3	2:C:538:GLN:HA	1.43	1.01

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	227/315 (72%)	200 (88%)	22 (10%)	5 (2%)	10	11
1	B	227/315 (72%)	200 (88%)	22 (10%)	5 (2%)	10	11
1	K	227/315 (72%)	200 (88%)	23 (10%)	4 (2%)	13	15
1	L	227/315 (72%)	200 (88%)	23 (10%)	4 (2%)	13	15
2	C	1117/1119 (100%)	927 (83%)	138 (12%)	52 (5%)	4	2
2	M	1117/1119 (100%)	926 (83%)	142 (13%)	49 (4%)	4	2
3	D	1388/1524 (91%)	1155 (83%)	168 (12%)	65 (5%)	4	2
3	N	1388/1524 (91%)	1133 (82%)	187 (14%)	68 (5%)	3	2
4	E	93/99 (94%)	76 (82%)	13 (14%)	4 (4%)	4	3
4	O	93/99 (94%)	76 (82%)	13 (14%)	4 (4%)	4	3
5	F	341/423 (81%)	290 (85%)	35 (10%)	16 (5%)	4	2
5	P	341/423 (81%)	288 (84%)	38 (11%)	15 (4%)	4	2
All	All	6786/7590 (89%)	5671 (84%)	824 (12%)	291 (4%)	4	3

5 of 291 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	29	GLU
1	B	29	GLU
1	B	48	ILE
2	C	152	PRO
2	C	231	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	202/273 (74%)	149 (74%)	53 (26%)	1	0
1	B	202/273 (74%)	167 (83%)	35 (17%)	3	3
1	K	202/273 (74%)	154 (76%)	48 (24%)	1	1
1	L	202/273 (74%)	152 (75%)	50 (25%)	1	1
2	C	941/941 (100%)	722 (77%)	219 (23%)	1	1
2	M	941/941 (100%)	731 (78%)	210 (22%)	1	1
3	D	1123/1279 (88%)	861 (77%)	262 (23%)	1	1
3	N	1123/1279 (88%)	832 (74%)	291 (26%)	1	1
4	E	83/87 (95%)	65 (78%)	18 (22%)	1	1
4	O	83/87 (95%)	61 (74%)	22 (26%)	1	0
5	F	295/370 (80%)	234 (79%)	61 (21%)	2	2
5	P	295/370 (80%)	242 (82%)	53 (18%)	2	3
All	All	5692/6446 (88%)	4370 (77%)	1322 (23%)	1	1

5 of 1322 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	F	136	LEU
1	L	197	LEU
3	N	1401	GLU
5	F	249	ARG
1	K	112	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 167 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1441	GLN
1	K	229	GLN
3	N	1465	ASN
4	E	28	GLN
5	F	218	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 5.6 Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 8 are monoatomic - leaving 2 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	$\# Z  > 2$
8	TGT	D	9001	6	27,27,27	4.00	20 (74%)	43,44,44	3.05	12 (27%)
8	TGT	N	9002	6	27,27,27	4.25	17 (62%)	43,44,44	2.88	13 (30%)

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
8	TGT	D	9001	6	-	1/16/57/57	0/0/2/2
8	TGT	N	9002	6	-	1/16/57/57	0/0/2/2

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	D	9001	TGT	O11-C10	9.11	1.55	1.20
8	N	9002	TGT	O11-C10	8.41	1.53	1.20
8	N	9002	TGT	C3-C2	8.40	1.64	1.52
8	N	9002	TGT	O3-C8	7.94	1.41	1.23
8	D	9001	TGT	C3-C2	7.22	1.62	1.52

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	D	9001	TGT	O2-C5-C8	-10.74	107.91	110.46
8	N	9002	TGT	O2-C5-C8	-8.54	108.44	110.46
8	N	9002	TGT	C8-C5-S1	8.49	110.77	103.69
8	D	9001	TGT	C8-C5-S1	8.05	110.41	103.69
8	D	9001	TGT	O10-C10-C11	6.93	124.18	111.12

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
8	N	9002	TGT	P1-O6-C1-C7
8	D	9001	TGT	P1-O6-C1-C7

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	229/315 (72%)	2.24	75 (32%) 1 1	18, 47, 72, 88	0
1	B	229/315 (72%)	3.05	85 (37%) 1 0	34, 66, 82, 88	0
1	K	229/315 (72%)	1.58	78 (34%) 1 0	21, 43, 70, 92	0
1	L	229/315 (72%)	2.23	74 (32%) 1 1	34, 62, 82, 95	0
2	C	1119/1119 (100%)	3.35	446 (39%) 1 0	15, 58, 81, 94	0
2	M	1119/1119 (100%)	3.49	464 (41%) 1 0	15, 55, 81, 97	0
3	D	1392/1524 (91%)	2.25	422 (30%) 1 1	15, 49, 82, 97	0
3	N	1392/1524 (91%)	2.28	431 (30%) 1 1	16, 48, 83, 105	0
4	E	95/99 (95%)	1.59	24 (25%) 1 1	30, 59, 82, 103	0
4	O	95/99 (95%)	2.04	26 (27%) 1 1	22, 59, 77, 87	0
5	F	345/423 (81%)	4.56	155 (44%) 1 0	38, 63, 83, 97	0
5	P	345/423 (81%)	4.65	163 (47%) 1 0	41, 64, 85, 92	0
All	All	6818/7590 (89%)	2.87	2443 (35%) 1 0	15, 54, 82, 105	0

The worst 5 of 2443 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	N	1248	GLY	61.7
3	N	532	GLY	57.0
3	N	533	GLY	56.2
3	D	854	ALA	55.7
3	D	1248	GLY	55.5

### 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 ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
8	TGT	N	9002	26/26	0.62	0.53	41,47,51,52	0
8	TGT	D	9001	26/26	0.49	0.48	44,47,50,52	0
6	MG	N	9006	1/1	0.04	-0.93	4,4,4,4	0
6	MG	C	9004	1/1	0.08	-1.03	17,17,17,17	0
7	ZN	D	9112	1/1	0.08	-1.39	50,50,50,50	0
7	ZN	N	9113	1/1	0.11	-1.56	41,41,41,41	0
6	MG	N	9005	1/1	0.04	-1.64	13,13,13,13	0
6	MG	D	9003	1/1	0.06	-1.88	17,17,17,17	0
7	ZN	N	9059	1/1	0.06	-1.99	42,42,42,42	0
7	ZN	D	9058	1/1	0.19	-5.04	56,56,56,56	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.