



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:07 PM GMT

PDB ID : 4WQS  
Title : Thermus thermophilus RNA polymerase backtracked complex  
Authors : Murayama, Y.; Sekine, S.; Yokoyama, S.  
Deposited on : 2014-10-22  
Resolution : 4.31 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](mailto: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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

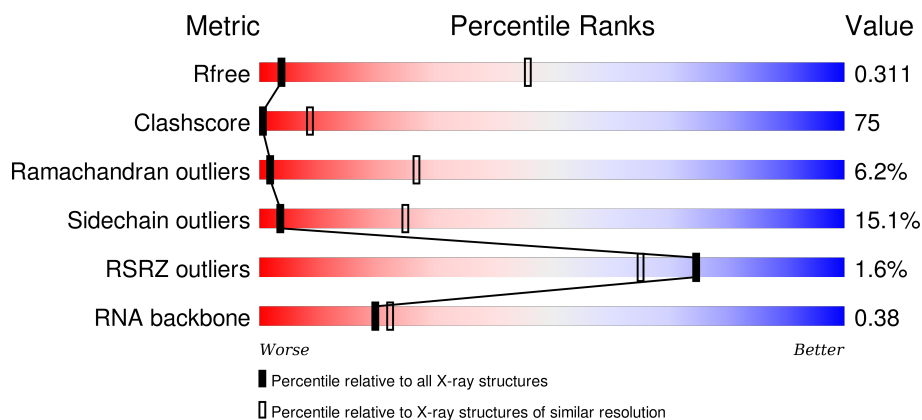
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.31 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))
$R_{free}$	91344	1059 (5.00-3.60)
Clashscore	102246	1166 (5.00-3.60)
Ramachandran outliers	100387	1106 (5.00-3.60)
Sidechain outliers	100360	1089 (5.00-3.60)
RSRZ outliers	91569	1062 (5.00-3.60)
RNA backbone	2183	1087 (5.60-2.80)

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. 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	315	
1	B	315	
1	K	315	
1	L	315	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	1119	
2	M	1119	
3	D	1524	
3	N	1524	
4	E	99	
4	O	99	
5	G	28	
5	X	28	
6	H	16	
6	Y	16	
7	I	21	
7	Z	21	

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 48166 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 subunit alpha.

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 subunit beta.

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 subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1151	Total	C	N	O	S	0	0	0
			9097	5753	1629	1682	33			
3	N	1288	Total	C	N	O	S	0	0	0
			10175	6441	1804	1899	31			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	95	Total	C	N	O	S	0	0	0
			770	491	133	142	4			
4	O	95	Total	C	N	O	S	0	0	0
			770	491	133	142	4			

- Molecule 5 is a DNA chain called DNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	G	27	Total	C	N	O	P	0	0	0
			548	259	104	158	27			
5	X	27	Total	C	N	O	P	0	0	0
			548	259	104	158	27			

- Molecule 6 is a RNA chain called RNA (5'-R(P\*CP\*CP\*AP\*GP\*CP\*CP\*GP\*GP\*CP\*GP\*CP\*UP\*CP\*GP\*CP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	16	Total	C	N	O	P	0	0	0
			340	151	61	112	16			
6	Y	15	Total	C	N	O	P	0	0	0
			318	141	56	106	15			

- Molecule 7 is a DNA chain called DNA (5'-D(P\*GP\*TP\*AP\*GP\*CP\*TP\*TP\*GP\*TP\*GP\*GP\*TP\*AP\*GP\*TP\*GP\*AP\*CP\*GP\*AP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	I	17	Total	C	N	O	P	0	0	0
			357	169	65	106	17			
7	Z	17	Total	C	N	O	P	0	0	0
			357	169	65	106	17			

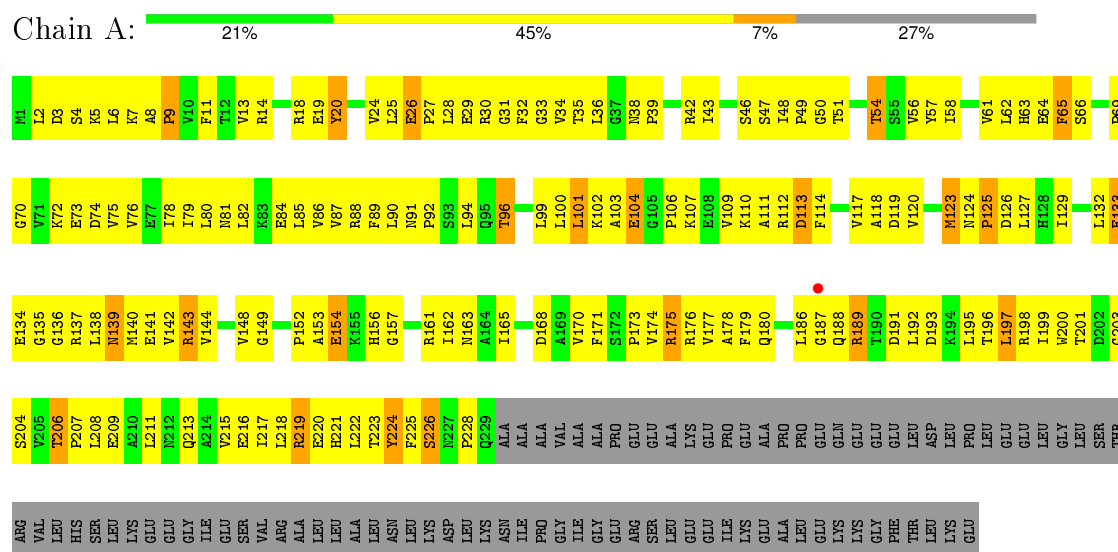
- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

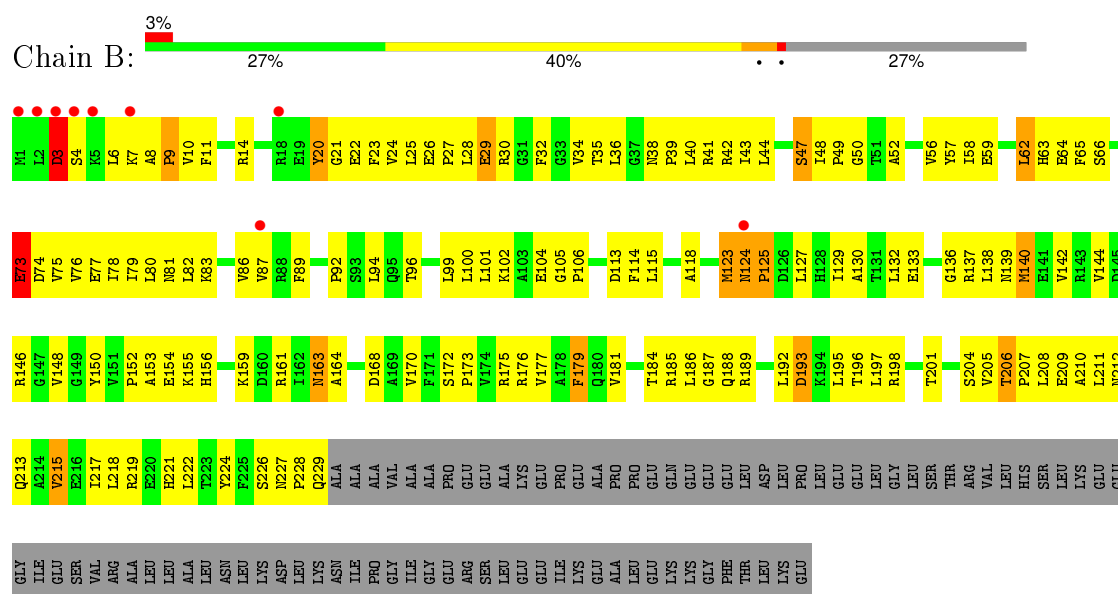
### 3 Residue-property plots

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.

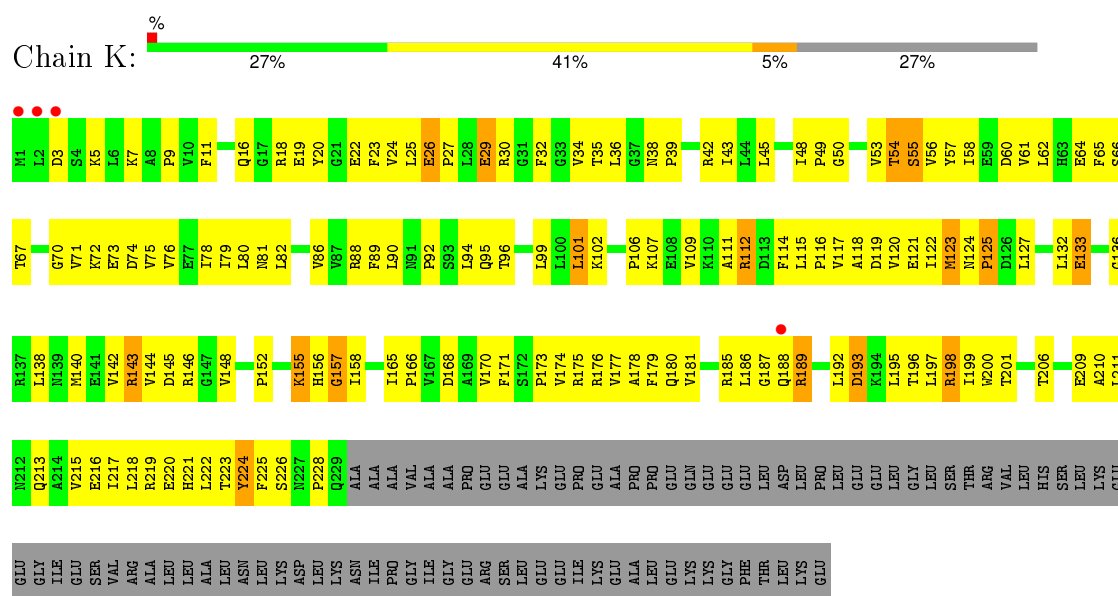
#### • Molecule 1: DNA-directed RNA polymerase subunit alpha

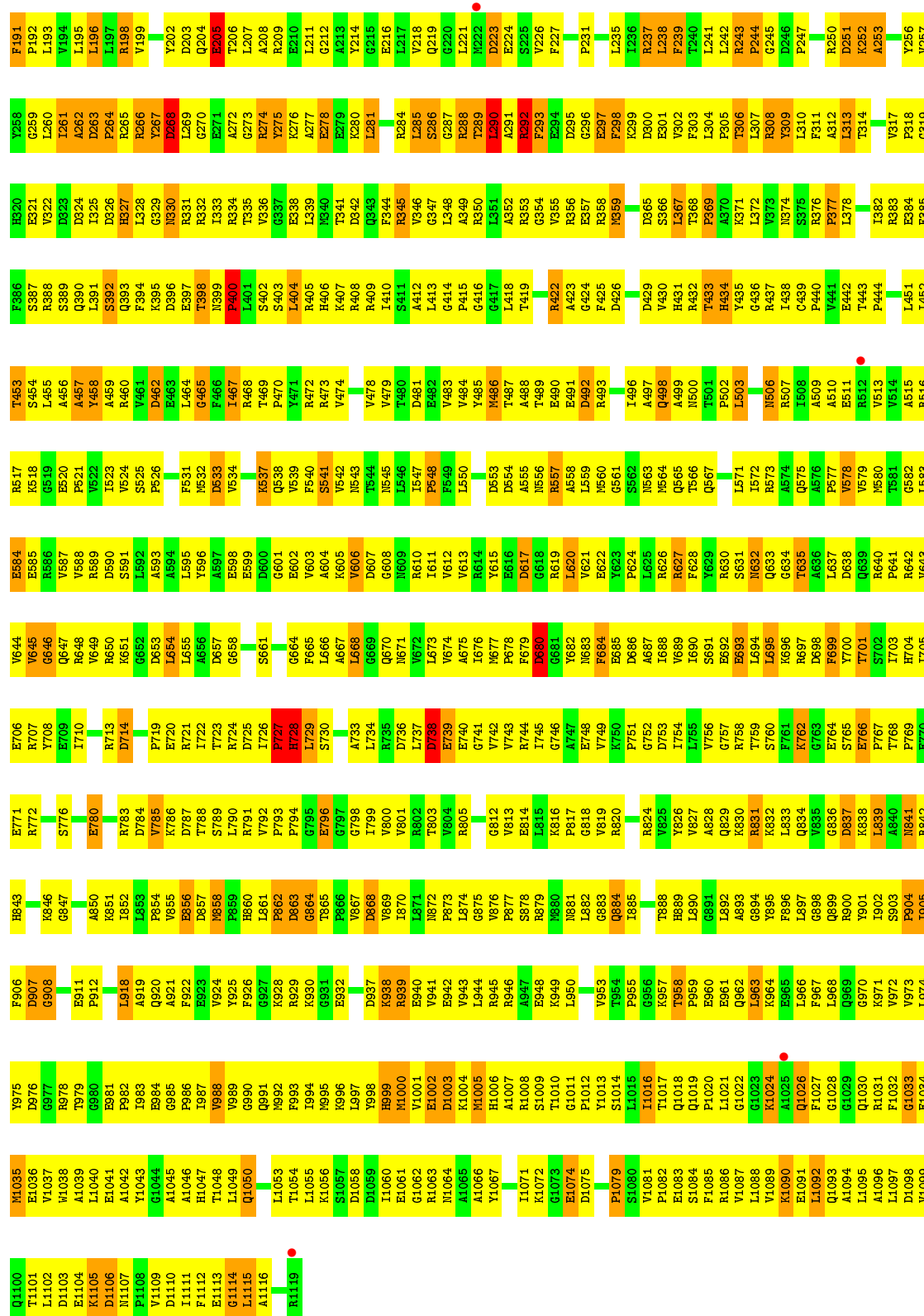


#### • Molecule 1: DNA-directed RNA polymerase subunit alpha



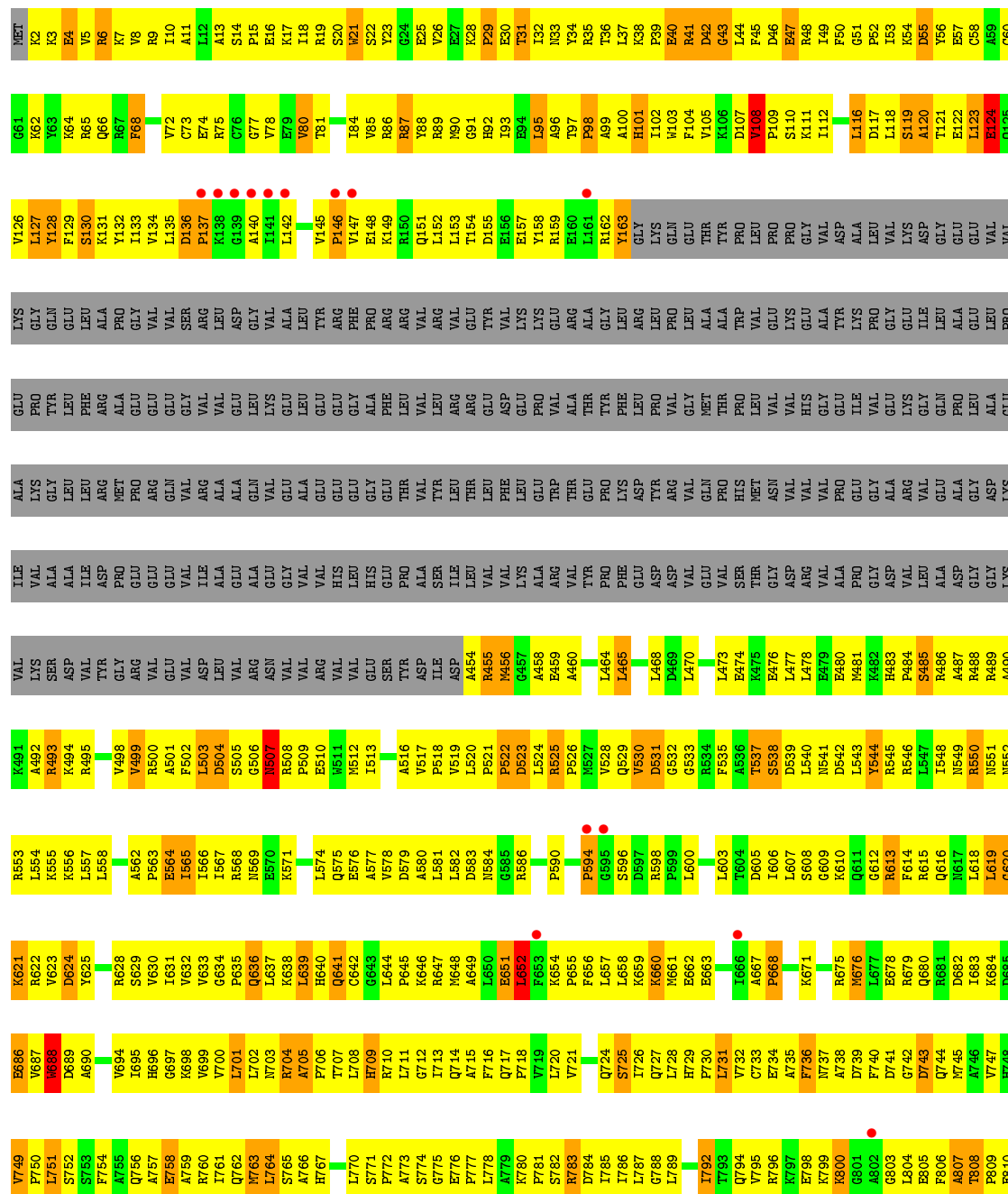
#### • Molecule 1: DNA-directed RNA polymerase subunit alpha

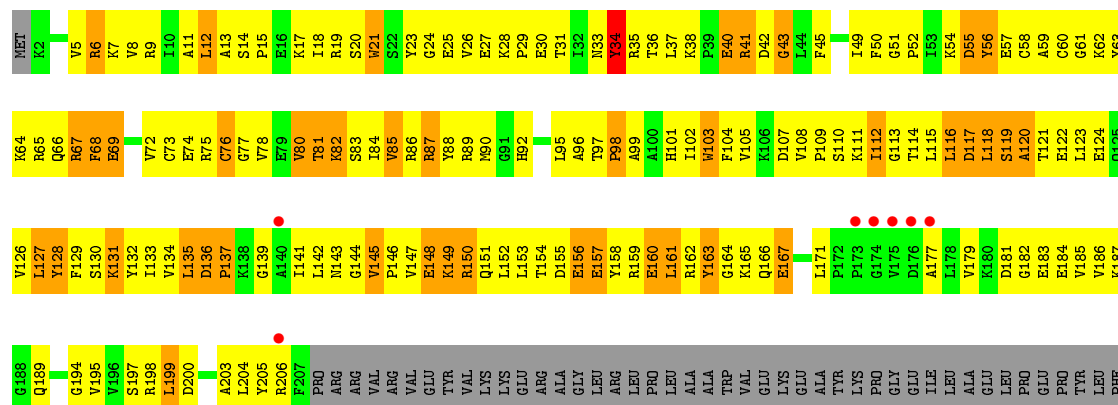




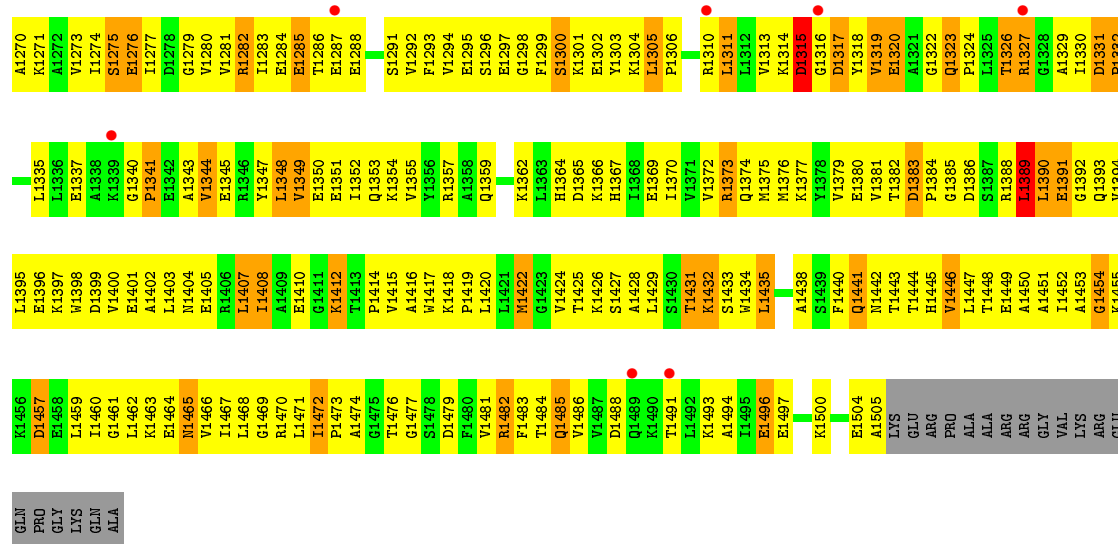


G977	F906	B843	E771	T705	V643	G582	K518	T453	Q390	G329	R265	L193		I128	D67
T979	D907	G844	R772	E706	V644	L583	G519	S454	L391	R330	R266	V194		I129	F68
G980	G908	B845	L773	R707	V645	E584	E520	L455	S392	R331	E267	L195		N130	L69
E981	A909	K346	L774	Y708	G646	E585	P521	A457	Q393	R332	D268	L196		G131	
G982	K910	G347	R775	E709	Q647	E586	V522	A458	F394	L333	L269	L197		V71	Y71
E983	E911	V848	S776	L710	R645	E587	L523	V458	K395	R334	G270	K198		R134	R72
E984	E912	B849		E711	V649	E588	V524	A459	D396	T335	E271	V199		V135	
G985	E913	A850		A712	R650	E589	S525	A460	S526	A272	E272	L73		G74	R8
G986	I914	R351		R713	K651	D590	P526	V461	T398	G337	C273	G201		V137	
E987	L918	L852	K781	D714	G652		E527	D462	K399	E338	R274	Y202		P76	R10
E988		L853	A782	T715	D653	A593	E528	E463	P400	L339	Y275	Q139		P77	E11
E989	A921	V855	R784	G718	L654	A594	F531	G465	L401	K340	K276	Q204		F78	V12
G990	F922	E356	P719	R719	L655	L595	M532	R472	S402	T341	A277	P79		A79	I13
Q991	E923	D857	K786	E720	D657	E596	D533	L467	L404	Q343	E278	Q80		Q80	P14
M992	V924	B858	R787	R721		E598	V534	L468	R405	Q343	E279	L207		D81	L15
F993	Y925		T788	L722	S661	E599	E535		H406	F344	K280			P144	P16
E994	F926	L861	S789	T723	E662	D600	P536	Y471	K407	V346		E210		C83	P17
M995	G927	P862	R790	R724	B663	G601	K537	R472	R408	E85		L211		R84	L18
K996	K928	D863	R791	R725	B664	E602	L538	R473	R409	S286		G212		K86	T19
L997	R929	B864	V792	L726	L666	B603	V539	V474	L410	A349	G287	L217		F148	E20
Y998	K930	T865	P793	P727	L666	A604	F540		S411	R350	L218	V218		D151	I31
B999		R866	P794	H728	A667	K605	S541	V478	A412	L351	T289	T89		P152	Q22
M1000	V936	V867	L799	L729	L668	B606	V542	V479	L413	A352	L290	Y90			S25
E1002	K937	D868	T799	S730	E669	D607	N543	V483	G414	R353	A291	Q91		G156	Y26
D1003	R939	T870	R805	E732	N671	G608	T544	V484	P415	G354	R292	D223		R157	L30
K1004	E940	L871		A733	V672	R610	N545	Y485	G417	R356	E294	L94		Y158	Q31
M1005	V941	N872	R808	L734	L673	I611	P548	M486	L418	F357	D295	Y95		A160	
E1006	E942	P873	G909	R735	V674	V612	F549	T487	T419	R358	G296	A96		S161	V34
A1007	V943	L874	D810	L736	A675	B613	L550	A488	R420	R359	E297	R230		I162	P35
R1008	L944	G875	P811	L737	L676	R614	E551	T489	E421	L360	F298	P231		I163	P36
S1009	R945	B876	G812	D738	N677	Y615	H552	E490	R422	N361	K299	Q99		P164	E37
T1010	R946	P877	V813	E739	P878	B616	D553	E491	A423	G362	D900	L235		L165	K38
G1011	A947	S878	E814	E740	R679	D617	D554	D492	L400	S363	E301	L236		P166	E38
P1012	E948	R879	L815	G741	D680	G618	A555	R493	F425	E364	V302	R237		H101	R39
Y1013	K949	R880	K816	V742	G681	R619	M556	Y494	D426	D385	F303	R237		H102	E40
S1014	L950	N881	P817	L743	V682	L620	R657	T495	E427	L238	K103	R237		K103	M41
L1015	G951	L882	G818	R744	N683	G621	A558	L496	R428	S366	D104	F239		P170	M42
L1016	L952	R883	R919		F684	E622	L559	A497	D429	T306	G106	T240		W171	G43
T1017		O884	R820	V749	B685	Y623	M560	D498	V430	P369	L307	L241		I172	I44
Q1018	T958	L885	E821	K750	D686	P624	G561	A499	R431	A370	R308	R243		D173	Q45
Q1019	P959	L886	V822	P751	R687	L625		N500	R432	R371	Y309	R244		L174	Q46
P1020	E960	E877	V823	G752	L688	R626	M564	T501	T433	L372	E110	G245		E175	A46
L1021	E961	R889	R824	D753	V689	R627	Q565	P502	H434	F311	G245	P244		P176	A47
G1022	Q962	T889	V825	L754	L690	F628	T566	L503	Y435	N374	F311	D111		E177	F48
	L963		Y826	L755	S691	Y629	Q567	E504	G436	E112	P247	E178		P178	G49
Q1026	K964	L892	V827	V756	E592	R630	A568	G505	R437	R376	L313	P248		W179	F52
F1027	E965	A893	A828	G757	B593	S631	V569	N506	L438	P377	T314	P248		G180	F53
G1028	L966	G894		R758	L694	N632	P570	R507	C439			D251		V181	P54
			K832	T759	L695	Q633	L571	L508	R378	E379	V317	K252		S182	I54
R1031	L968	L897	L833	S760	V696	G634		A509	E442		G319	A253		S183	E55
Q969	Q969	G898	Q834	F761	R697	T635		A510	T443		R318	R253		M184	E56
G1033	G970	B899		K762	D698	B636	Q575	E511	P444		E321	V257		K185	E57
E1034	K971	A900			F699	L637	A576	R512	E445			G259		V186	K59
M1035	V972					D638		V513	E445			D324		N187	
E1036	V973	L838			Y700	T701	P577	V513	I449			L260		K188	E123
V1037	L974	S903	A840	E766	S702	Q639	V578	A515	G450			D261		R189	G62
				T768	T703	R641	R640	A515	D450			S387		K190	L64
						P841	M580	B516	L451			R327		P191	G63
						R642		R513	D263			S388		S126	L65
							T581		L452			D389		E129	L66

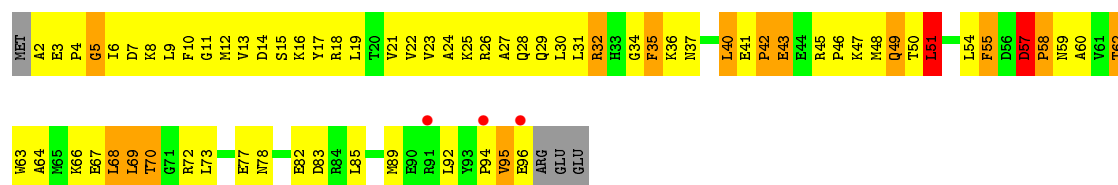




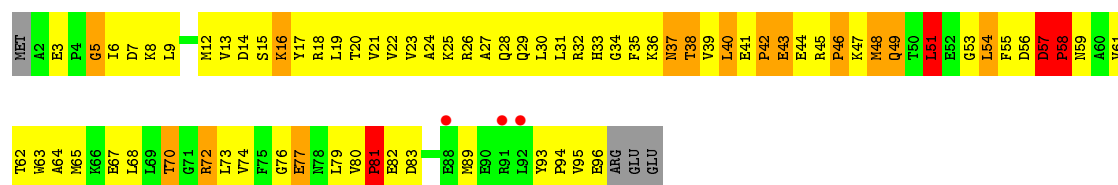
S1210	E1141	I1072	F1008	K935	K971	B810	H748	D685	G620	K555	R495	R434	ASP	ARG
M1211	G1146	H075	K1009	I936	R672	B811	V749	E686	K621	K556	L496	V435	PRO	ALA
A1212	R1147	H075	M1010	I937	L873	A812	P750	V887	R622	L557	E497	E436	GLU	GLU
R1213	L1148	A1082	F1011	G938	E874	L813	L751	M888	V623	L558	V498	V437	ARG	GLU
P1214	L1149	A1083	E1012	T939	T875	A814	S752	D689	D624	A569	V499	D438	GLN	GLY
S1215	A1150	T940	E1013	F941	S876	A815	S753	A690	V625	O560	R500	VAL	VAL	GLY
S1216	A1151	L1084	P1016	S942	P877	B816	F754	L891	S626	G561	A501	ILE	ARG	VAL
I1217	R1152	L1086	F1017	R943	G878	B817	A755	E892	G627	A562	F502	ALA	ALA	VAL
G1218	L1153	A1088	F1018	T944	R879	A818	Q756	E893	R628	P563	L503	GLU	ALA	VAL
E1219	L1154	L1088	P1019	S945	L880	G819	A757	V694	S629	E564	L504	ALA	GLN	LEU
A1220	V1155	A1089	G1020	G946	L881	B820	E758	I695	V630	S505	GLU	VAL	LYS	LYS
G1221	D1090	A1090	L1021	I947	P882	B821	A759	H696	V631	I566	GLY	VAL	GLU	LEU
I1222	Y1093	L1091	V1022	T948	A883	A822	R760	G697	V632	I567	VAL	VAL	ALA	LEU
I1223	R1158	L1098	M1023	T948	R884	L823	I761	K698	V633	R568	GLU	VAL	GLU	GLU
V1224	R1159	L1094	M1023	I949	L885	B824	Q762	V699	G634	P569	HIS	VAL	GLU	GLU
A1225	L1160	L1094	G950	T949	R886	A825	M763	V700	GLU	E510	LEU	GLU	GLY	GLY
E1226	E1161	T1095	G1027	I951	A887	P826	L764	R701	L637	K571	HIS	D451	ALA	ALA
Q1227	E1162	R1096	A1028	D952	E888	B827	S765	L702	K638	R572	VAL	I452	GLY	PHE
Q1228	G1163	K1097	R1029	D953	A889	K828	A766	N703	L639	M573	PRO	D463	LEU	LEU
I1229	R1164	L1098	G1030	P957	B890	B829	H767	A705	R640	L574	VAL	A454	THR	VAL
E1231	Y1165	V1099	M1031	E958	E891	A830	N768	A705	K641	E515	TYR	R455	VAL	VAL
P1232	L1166	D1100	P1032	E958	L892	B831	L769	P706	Q642	A516	TYR	R456	LEU	LEU
S1233	S1167	V1101	Q1033	E959	E893	B832	L770	T707	G643	A577	THR	I393	ARG	ARG
G1233	M1168	T1102	Q1034	K960	E894	B833	L779	L708	G643	P518	LEU	V395	GLU	GLU
T1234	H1103	H1035	I1085	K961	T834	B834	A779	H709	P645	D579	ASP	V396	ASP	ASP
Q1235	V1171	E1104	R1036	Q962	A896	S835	S774	R710	K646	A560	GLU	K397	GLU	GLU
L1236	H1172	I1105	L1037	T963	R897	B836	G775	R711	GLU	L581	PRO	A391	LEU	PRO
T1237	V1106	V1107	LEU	L964	E898	G837	E776	G712	A649	L582	TRP	A398	VAL	VAL
M1238	I1175	P1107	CYS	E965	L899	R838	P777	I743	L650	D583	THR	R399	ALA	ALA
R1239	K1176	L1108	GLY	E966	L900	L839	L778	Q714	B651	N584	THR	V400	THR	THR
T1240	A1177	E1109	LEU	A967	Q901	K840	A779	A715	L652	G585	PRO	R525	TYR	TYR
F1241	A1178	A1110	ARG	D968	L902	B841	K780	F716	P653	R586	LYS	L465	PHE	PHE
HIS	E1179	D1111	GLY	R969	Q907	B842	P781	Q717	K654	R587	ASP	D405	LEU	LEU
THR	A1180	G1112	MET	R970	E907	F843	S782	P718	P655	E588	PRO	D406	PRO	PRO
GLY	I1183	C1113	GLN	L972	K908	A844	D784	V719	F656	L468	ARG	V407	VAL	VAL
VAL	T1114	L1115	LYS	Q973	L913	P846	L785	L720	V530	Q529	GLN	V409	GLY	GLY
ALA	E1185	H1116	PRO	I974	K912	D847	L786	V721	E471	D631	PRO	S410	MET	MET
GLY	Y1117	Y1117	SER	E975	B854	D847	L786	E722	A472	G532	PRO	T411	THR	THR
ALA	I1118	I1118	GLY	Q976	B855	E848	L787	G723	L473	G533	HIS	G412	PRO	PRO
ALA	L1122	L1122	THR	A977	B856	A849	G788	G723	R660	P590	MET	D413	LEU	LEU
ASP	R1189	L1123	PHE	F982	B857	N845	D784	L720	L657	P590	ASN	R414	VAL	VAL
ILE	P1191	Q1124	THR	D985	B858	L851	L789	L728	B662	L600	VAL	V415	VAL	VAL
THR	L1192	F1125	GLU	D985	B859	A852	Y791	H729	K664	T837	VAL	A416	VAL	HIS
GLN	T1193	P1125	VAL	D985	B860	A853	L792	H730	L477	S838	VAL	A417	GLY	GLY
G1255	C1194	E1127	PRO	D985	B861	B862	Q794	L731	I666	T604	PRO	G418	PRO	GLY
L1256	Q1195	D1126	VAL	E986	B862	L860	V795	V732	A667	D605	GLU	D419	GLU	ILE
P1257	T1196	L1128	R1058	E987	B863	L851	L796	C733	P668	L606	ALA	L421	ALA	GLU
R1258	R1197	T1129	S1059	E988	B864	B865	R796	V733	P668	L607	ARG	L422	ARG	LYS
V1259	R1197	R1130	S1060	D990	B865	L860	R796	A735	K671	S608	VAL	A423	VAL	GLY
I1260	V1200	S1131	F1061	Q991	B866	L861	K799	F736	K671	G609	GLN	D424	GLN	GLN
E1261	C1201	L1132	R1062	I992	B867	Q861	K800	N737	R675	K610	PRO	G424	ALA	ALA
L1262	Q1202	R1133	L1065	I993	B868	Q862	G801	A738	R676	K611	GLY	G425	GLY	PRO
F1263	K1203	L1134	L1066	Q994	B869	B863	A802	D739	L677	G612	LEU	K426	LEU	LEU
E1264	C1204	R1135	L1067	L995	B870	B864	G803	F740	L678	R613	ASP	V427	ALA	ALA
A1265	Y1205	K1136	V1067	L996	B871	B865	L804	D741	B679	R614	LYS	K428	GLU	GLU
R1266	G1266	R1137	L1068	L999	B872	B866	E805	G742	Q680	R615	ILE	S429	ALA	ALA
R1267	Y1207	A1138	E1069	T999	B873	B867	F806	D743	R681	Q616	VAL	K430	LYS	LYS
D1268	L1208	A1139	E1070	T999	B874	B868	A807	Q744	D682	R617	GLY	V431	GLY	GLY
K1269	L1209	I1140	F1071	V1007	B875	B869	T808	M745	K684	L619	ALA	Y432	ALA	LEU



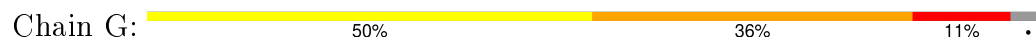
• Molecule 4: DNA-directed RNA polymerase subunit omega



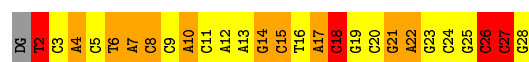
• Molecule 4: DNA-directed RNA polymerase subunit omega



• Molecule 5: DNA (28-MER)



• Molecule 5: DNA (28-MER)



- Molecule 6: RNA (5'-R(P\*CP\*CP\*AP\*GP\*CP\*CP\*GP\*GP\*CP\*GP\*CP\*UP\*CP\*GP\*CP\*A)-3')



- Molecule 6: RNA (5'-R(P\*CP\*CP\*AP\*GP\*CP\*CP\*GP\*GP\*CP\*GP\*CP\*UP\*CP\*GP\*CP\*A)-3')



- Molecule 7: DNA (5'-D(P\*GP\*TP\*AP\*GP\*CP\*TP\*TP\*GP\*TP\*GP\*GP\*TP\*AP\*GP\*TP\*GP\*AP\*CP\*GP\*AP\*G)-3')



- Molecule 7: DNA (5'-D(P\*GP\*TP\*AP\*GP\*CP\*TP\*TP\*GP\*TP\*GP\*GP\*TP\*AP\*GP\*TP\*GP\*AP\*CP\*GP\*AP\*G)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	155.98Å 155.98Å 495.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.15 – 4.31 44.15 – 4.31	Depositor EDS
% Data completeness (in resolution range)	99.9 (44.15-4.31) 99.4 (44.15-4.31)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.17	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.12 (at 4.28Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.281 , 0.311 0.281 , 0.311	Depositor DCC
$R_{free}$ test set	3812 reflections (4.83%)	DCC
Wilson B-factor (Å <sup>2</sup> )	102.9	Xtriage
Anisotropy	0.055	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.15 , 67.4	EDS
Estimated twinning fraction	0.499 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.27$ , $\langle L^2 \rangle = 0.12$	Xtriage
Outliers	0 of 79319 reflections	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	48166	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	250.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.72% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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

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	0.38	0/1838	0.69	0/2498
1	B	0.43	0/1838	0.68	0/2498
1	K	0.45	0/1838	0.70	0/2498
1	L	0.42	0/1838	0.67	0/2498
2	C	0.49	0/8997	0.79	3/12164 (0.0%)
2	M	0.48	0/8997	0.78	4/12164 (0.0%)
3	D	0.51	1/9249 (0.0%)	0.83	10/12482 (0.1%)
3	N	0.51	0/10344	0.81	8/13968 (0.1%)
4	E	0.50	0/784	0.87	2/1057 (0.2%)
4	O	0.46	0/784	0.84	2/1057 (0.2%)
5	G	0.99	1/614 (0.2%)	1.41	9/943 (1.0%)
5	X	0.93	0/614	1.43	11/943 (1.2%)
6	H	1.14	3/378 (0.8%)	1.56	6/585 (1.0%)
6	Y	1.15	3/353 (0.8%)	1.45	5/546 (0.9%)
7	I	0.94	1/400 (0.2%)	1.46	7/616 (1.1%)
7	Z	0.90	1/400 (0.2%)	1.28	3/616 (0.5%)
All	All	0.53	10/49266 (0.0%)	0.84	70/67133 (0.1%)

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
5	G	0	10
5	X	0	11
6	H	0	4
6	Y	0	5
7	I	0	6
7	Z	0	4
All	All	0	40



The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	Z	1	DG	OP3-P	-7.99	1.51	1.61
7	I	1	DG	OP3-P	-7.84	1.51	1.61
6	Y	11	C	N1-C2	-7.40	1.32	1.40
6	H	11	C	N1-C2	-7.38	1.32	1.40
6	H	1	C	OP3-P	-7.06	1.52	1.61

The worst 5 of 70 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	I	1	DG	O5'-P-OP1	-18.71	88.24	110.70
7	Z	1	DG	O5'-P-OP1	-16.85	90.48	110.70
6	H	5	C	N1-C1'-C2'	-15.07	94.41	114.00
6	Y	5	C	N1-C1'-C2'	-9.98	101.02	112.00
7	I	1	DG	O5'-P-OP2	9.33	121.90	110.70

There are no chirality outliers.

5 of 40 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	G	10	DA	Sidechain
5	G	2	DT	Sidechain
5	G	3	DC	Sidechain
5	G	4	DA	Sidechain
5	G	9	DC	Sidechain

## 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	1806	0	1861	257	0
1	B	1806	0	1861	180	0
1	K	1806	0	1861	189	0
1	L	1806	0	1861	203	0
2	C	8829	0	8933	1430	0
2	M	8829	0	8933	1425	0
3	D	9097	0	9316	1626	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	N	10175	0	10401	1763	0
4	E	770	0	784	124	0
4	O	770	0	784	151	0
5	G	548	0	301	113	0
5	X	548	0	301	93	0
6	H	340	0	176	76	0
6	Y	318	0	165	56	0
7	I	357	0	194	68	0
7	Z	357	0	194	71	0
8	D	2	0	0	0	0
8	N	2	0	0	0	0
All	All	48166	0	47926	7162	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 75.

The worst 5 of 7162 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:C:52:PHE:CZ	2:C:68:PHE:HB2	1.22	1.67
3:D:897:TRP:HA	3:D:900:ILE:CG1	1.33	1.56
3:D:1041:LEU:HD11	3:D:1045:MET:SD	1.55	1.44
3:D:705:ALA:HB1	6:H:14:G:N2	1.24	1.40
3:D:705:ALA:CB	6:H:14:G:H21	1.40	1.32

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	227/315 (72%)	194 (86%)	24 (11%)	9 (4%)	<b>4</b> 37

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	227/315 (72%)	195 (86%)	23 (10%)	9 (4%)	4	37
1	K	227/315 (72%)	193 (85%)	26 (12%)	8 (4%)	4	41
1	L	227/315 (72%)	201 (88%)	17 (8%)	9 (4%)	4	37
2	C	1117/1119 (100%)	893 (80%)	151 (14%)	73 (6%)	1	26
2	M	1117/1119 (100%)	889 (80%)	153 (14%)	75 (7%)	1	25
3	D	1143/1524 (75%)	903 (79%)	167 (15%)	73 (6%)	2	26
3	N	1280/1524 (84%)	1011 (79%)	190 (15%)	79 (6%)	2	27
4	E	93/99 (94%)	67 (72%)	15 (16%)	11 (12%)	0	9
4	O	93/99 (94%)	69 (74%)	13 (14%)	11 (12%)	0	9
All	All	5751/6744 (85%)	4615 (80%)	779 (14%)	357 (6%)	2	27

5 of 357 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	3	ASP
1	A	29	GLU
1	A	118	ALA
1	A	133	GLU
1	A	187	GLY

### 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%)	178 (88%)	24 (12%)	6	34
1	B	202/273 (74%)	182 (90%)	20 (10%)	10	42
1	K	202/273 (74%)	182 (90%)	20 (10%)	10	42
1	L	202/273 (74%)	182 (90%)	20 (10%)	10	42
2	C	941/941 (100%)	790 (84%)	151 (16%)	3	22
2	M	941/941 (100%)	794 (84%)	147 (16%)	3	24
3	D	968/1279 (76%)	806 (83%)	162 (17%)	3	20

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	N	1088/1279 (85%)	919 (84%)	169 (16%)	3	24
4	E	84/88 (96%)	70 (83%)	14 (17%)	3	20
4	O	84/88 (96%)	70 (83%)	14 (17%)	3	20
All	All	4914/5708 (86%)	4173 (85%)	741 (15%)	3	25

5 of 741 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	1269	LYS
2	M	34	VAL
3	N	1210	SER
3	D	1441	GLN
1	K	54	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	M	406	HIS
2	M	567	GLN
3	N	1031	ASN
2	M	393	GLN
3	N	714	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	H	15/16 (93%)	5 (33%)	1 (6%)
6	Y	14/16 (87%)	3 (21%)	1 (7%)
All	All	29/32 (90%)	8 (27%)	2 (6%)

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	H	4	G
6	H	5	C
6	H	6	C
6	H	15	C
6	H	16	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
6	H	5	C
6	Y	5	C

## 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 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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%)	-0.75	1 (0%) 93 90	154, 243, 347, 433	0
1	B	229/315 (72%)	-0.46	9 (3%) 43 34	202, 343, 460, 525	0
1	K	229/315 (72%)	-0.57	4 (1%) 73 63	157, 255, 375, 500	0
1	L	229/315 (72%)	-0.46	4 (1%) 73 63	193, 363, 452, 538	0
2	C	1119/1119 (100%)	-0.74	8 (0%) 89 84	12, 213, 336, 454	0
2	M	1119/1119 (100%)	-0.71	8 (0%) 89 84	60, 238, 426, 549	0
3	D	1151/1524 (75%)	-0.72	15 (1%) 79 71	8, 205, 382, 548	0
3	N	1288/1524 (84%)	-0.57	40 (3%) 52 41	81, 249, 452, 545	0
4	E	95/99 (95%)	-0.51	3 (3%) 51 40	156, 217, 337, 395	0
4	O	95/99 (95%)	-0.60	3 (3%) 51 40	205, 297, 426, 451	0
5	G	27/28 (96%)	-0.52	0 100 100	152, 210, 372, 381	0
5	X	27/28 (96%)	-0.44	0 100 100	238, 279, 353, 374	0
6	H	16/16 (100%)	0.32	1 (6%) 23 16	97, 189, 363, 385	0
6	Y	15/16 (93%)	-0.18	0 100 100	208, 249, 351, 378	0
7	I	17/21 (80%)	-0.37	0 100 100	215, 276, 418, 428	0
7	Z	17/21 (80%)	-0.45	0 100 100	264, 305, 348, 352	0
All	All	5902/6874 (85%)	-0.65	96 (1%) 74 65	8, 240, 416, 549	0

The worst 5 of 96 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	140	ALA	13.5
3	N	175	VAL	13.0
3	D	141	ILE	10.1
1	B	3	ASP	9.1
1	L	68	ILE	8.0

## 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, 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
8	ZN	N	1602	1/1	0.98	0.11	-0.53	72,72,72,72	0
8	ZN	N	1601	1/1	0.97	0.13	-1.18	72,72,72,72	0
8	ZN	D	1602	1/1	0.99	0.14	-1.22	72,72,72,72	0
8	ZN	D	1601	1/1	0.98	0.05	-1.32	72,72,72,72	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.