



wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 09:51 pm BST

PDB ID : 2A6E
Title : Crystal structure of the T. Thermophilus RNA polymerase holoenzyme
Authors : Artsimovitch, I.; Vassilyeva, M.N.; Svetlov, D.; Svetlov, V.; Perederina, A.;
Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Tahirov, T.H.; Vassilyev, D.G.;
RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on : 2005-07-02
Resolution : 2.80 Å(reported)

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

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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

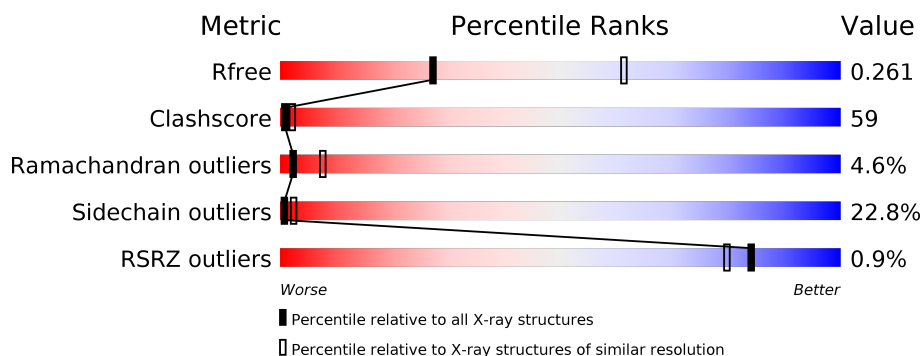
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	<div> <div>17%</div> <div>40%</div> <div>16%</div> <div>27%</div> </div>
1	B	315	<div> <div>3%</div> <div>17%</div> <div>45%</div> <div>11%</div> <div>27%</div> </div>
1	K	315	<div> <div>14%</div> <div>47%</div> <div>11%</div> <div>27%</div> </div>
1	L	315	<div> <div>%</div> <div>18%</div> <div>44%</div> <div>10%</div> <div>27%</div> </div>
2	C	1119	<div> <div>%</div> <div>22%</div> <div>60%</div> <div>17%</div> <div>•</div> </div>
2	M	1119	<div> <div>%</div> <div>25%</div> <div>57%</div> <div>17%</div> <div>•</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	1524	<div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>23%51%17%•9%</div></div>
3	N	1524	<div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>24%51%15%•9%</div></div>
4	E	99	<div><div></div><div><div></div><div></div><div></div><div></div></div><div>22%52%20%••</div></div>
4	O	99	<div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>28%48%19%•</div></div>
5	F	423	<div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>19%47%14%•18%</div></div>
5	P	423	<div><div>%</div><div><div></div><div></div><div></div><div></div></div><div>22%49%10%18%</div></div>

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 58679 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 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 ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	1	Total	Mg	0	0
			1	1		
7	N	1	Total	Mg	0	0
			1	1		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	191	Total	O	0	0
			191	191		
8	B	181	Total	O	0	0
			181	181		
8	C	767	Total	O	0	0
			767	767		
8	D	1100	Total	O	0	0
			1100	1100		
8	E	93	Total	O	0	0
			93	93		
8	F	333	Total	O	0	0
			333	333		
8	K	151	Total	O	0	0
			151	151		
8	L	179	Total	O	0	0
			179	179		
8	M	739	Total	O	0	0
			739	739		

Continued on next page...

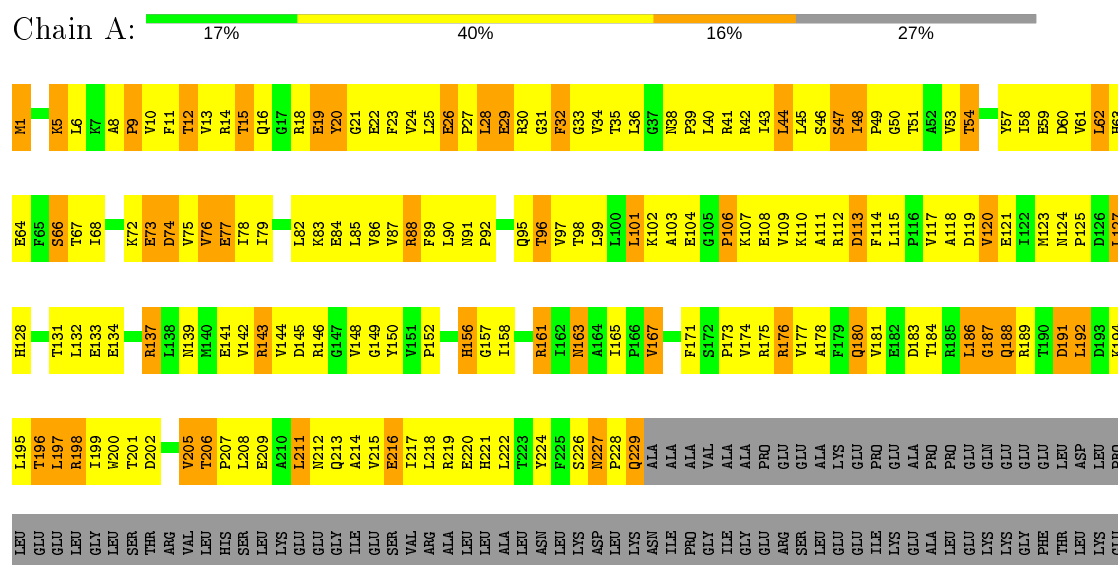
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	N	1038	Total 1038	O 1038	0	0
8	O	78	Total 78	O 78	0	0
8	P	267	Total 267	O 267	0	0

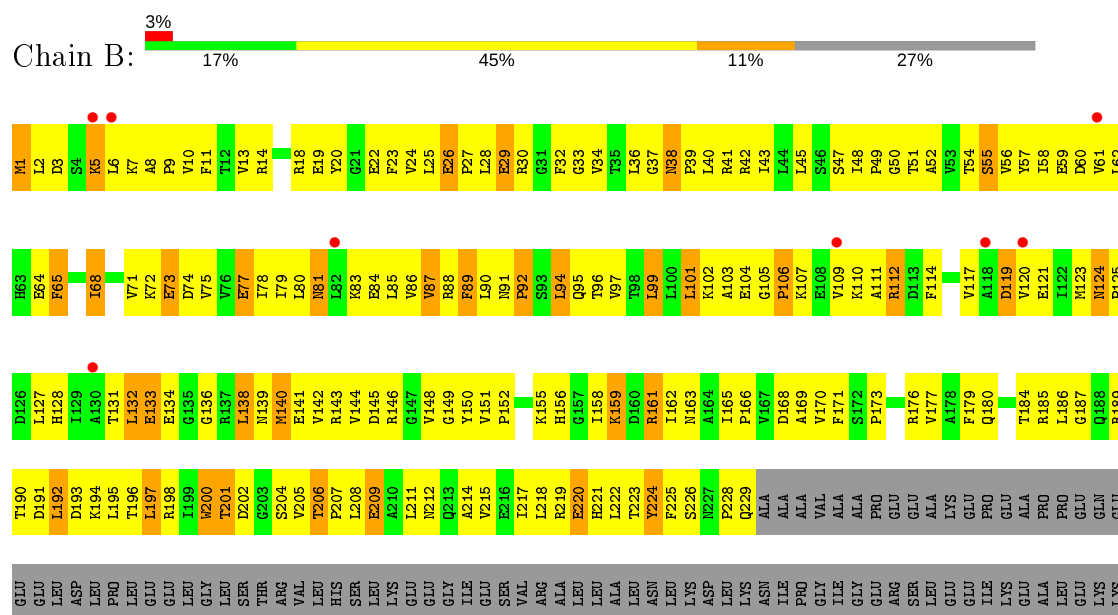
3 Residue-property plots

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

• Molecule 1: DNA-directed RNA polymerase alpha chain



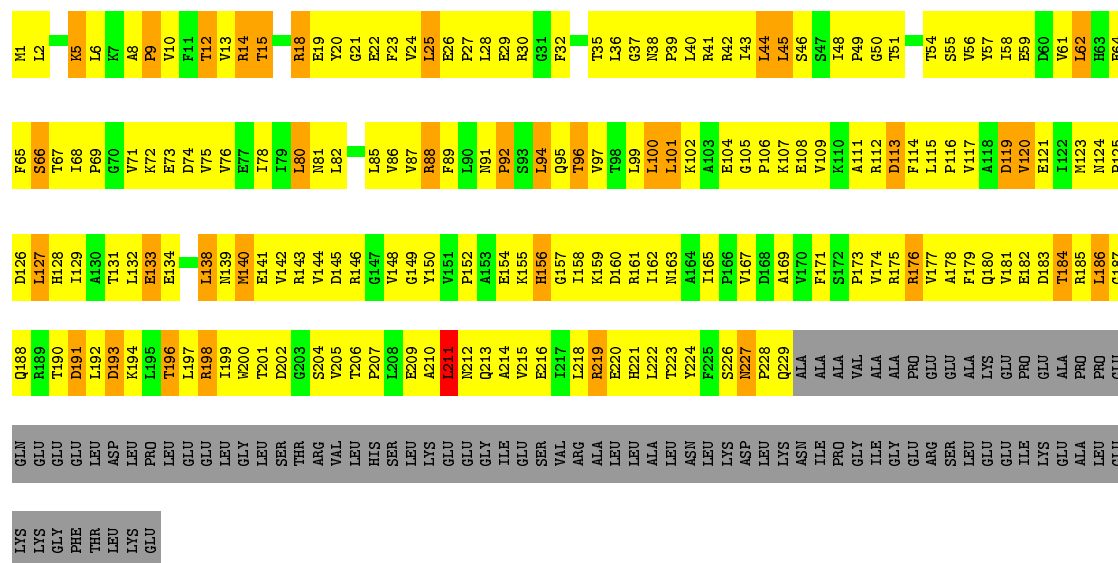
• Molecule 1: DNA-directed RNA polymerase alpha chain



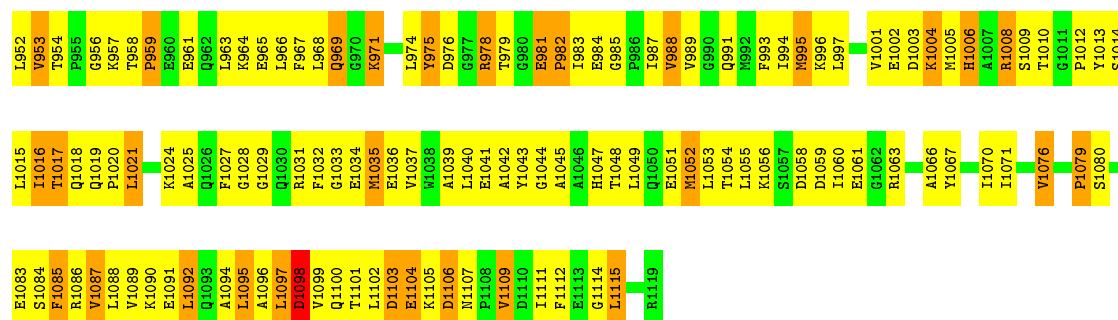
GLY
PHE
THR
LEU
LYS
GLU

• Molecule 1: DNA-directed RNA polymerase alpha chain

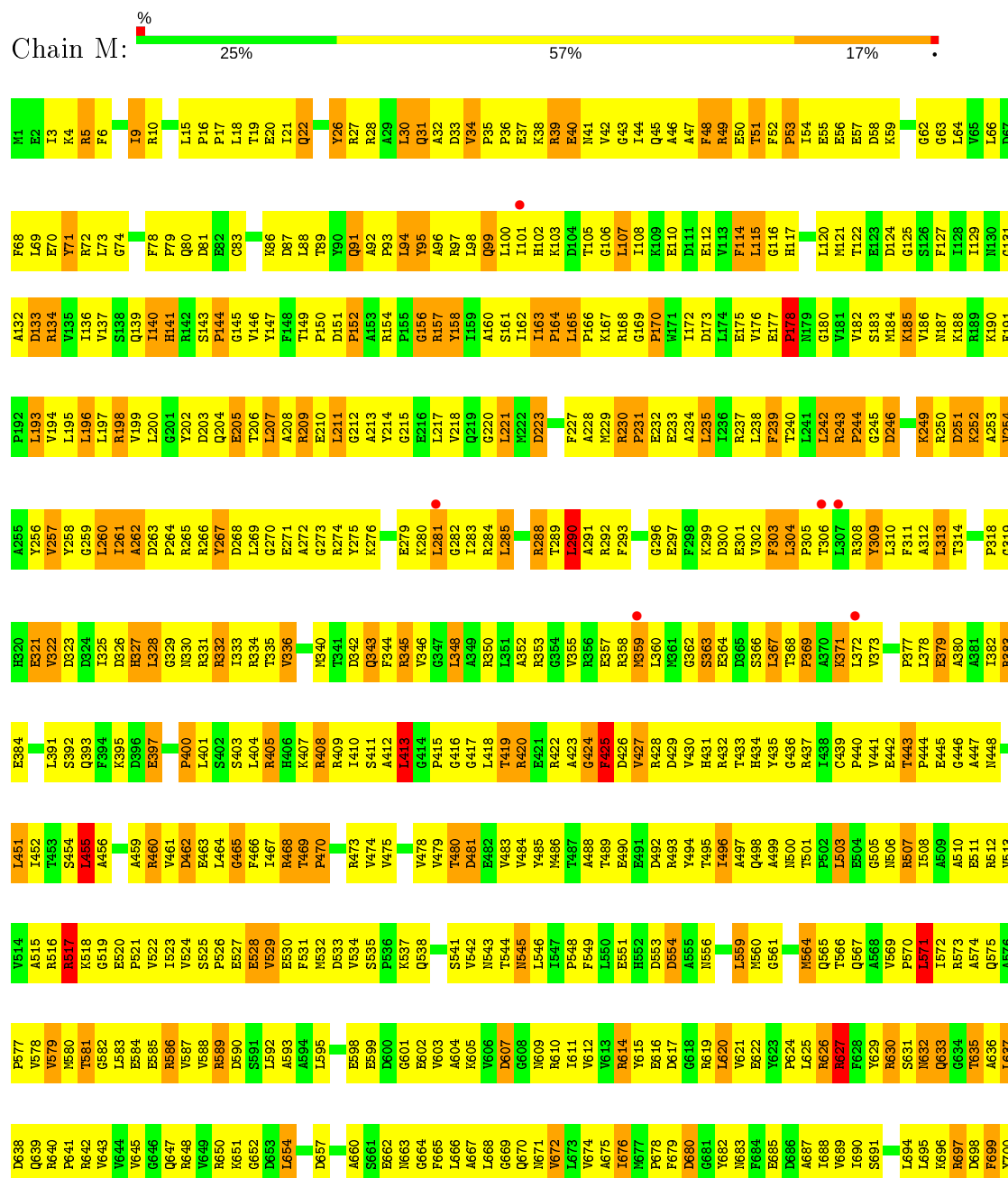
Chain K: 14% 47% 11% 27%

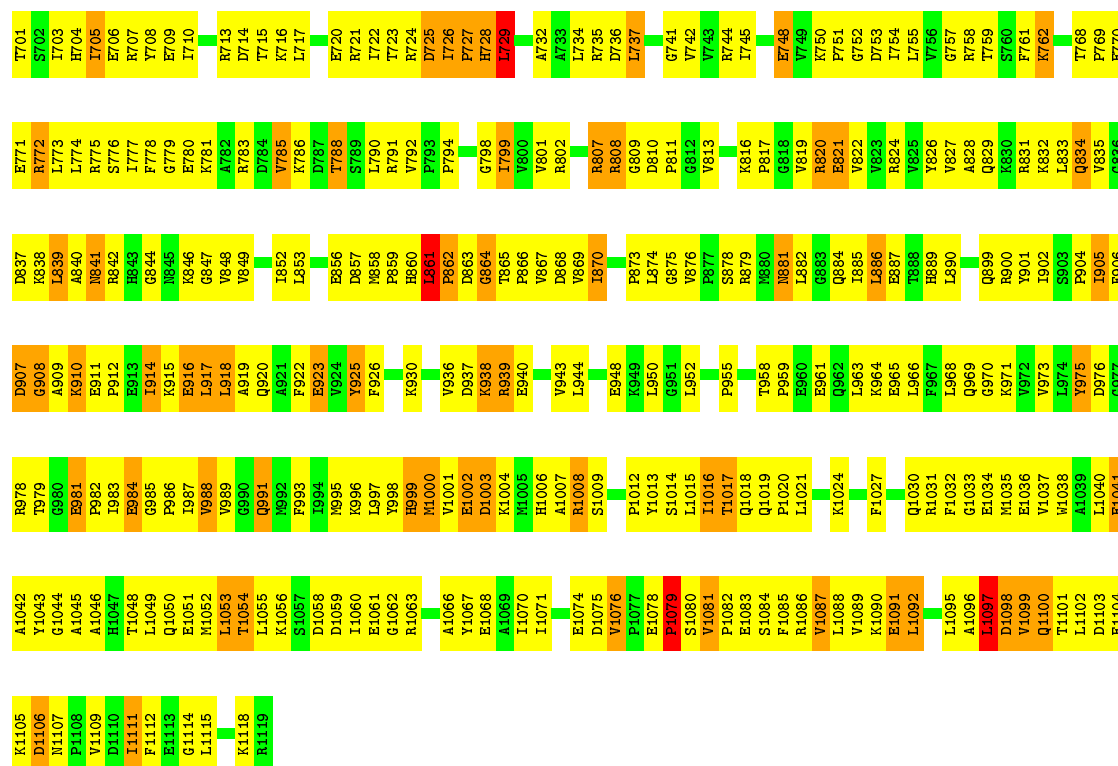




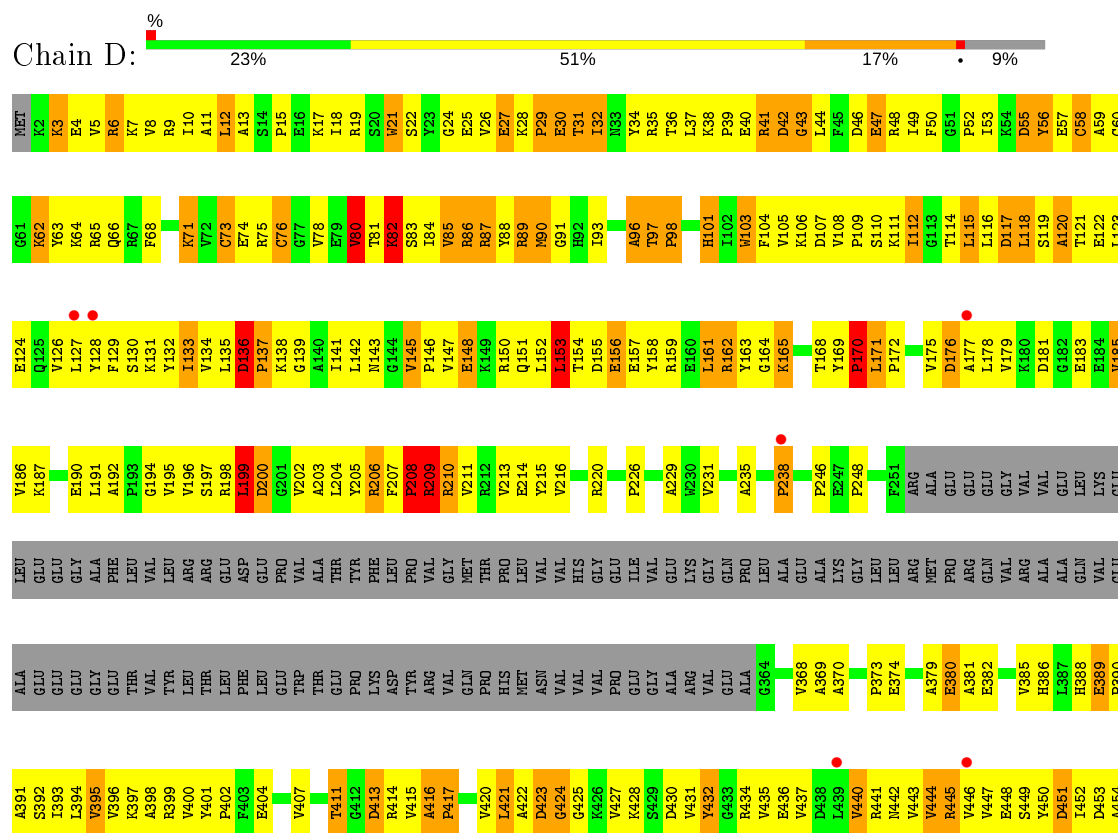


• Molecule 2: DNA-directed RNA polymerase beta chain



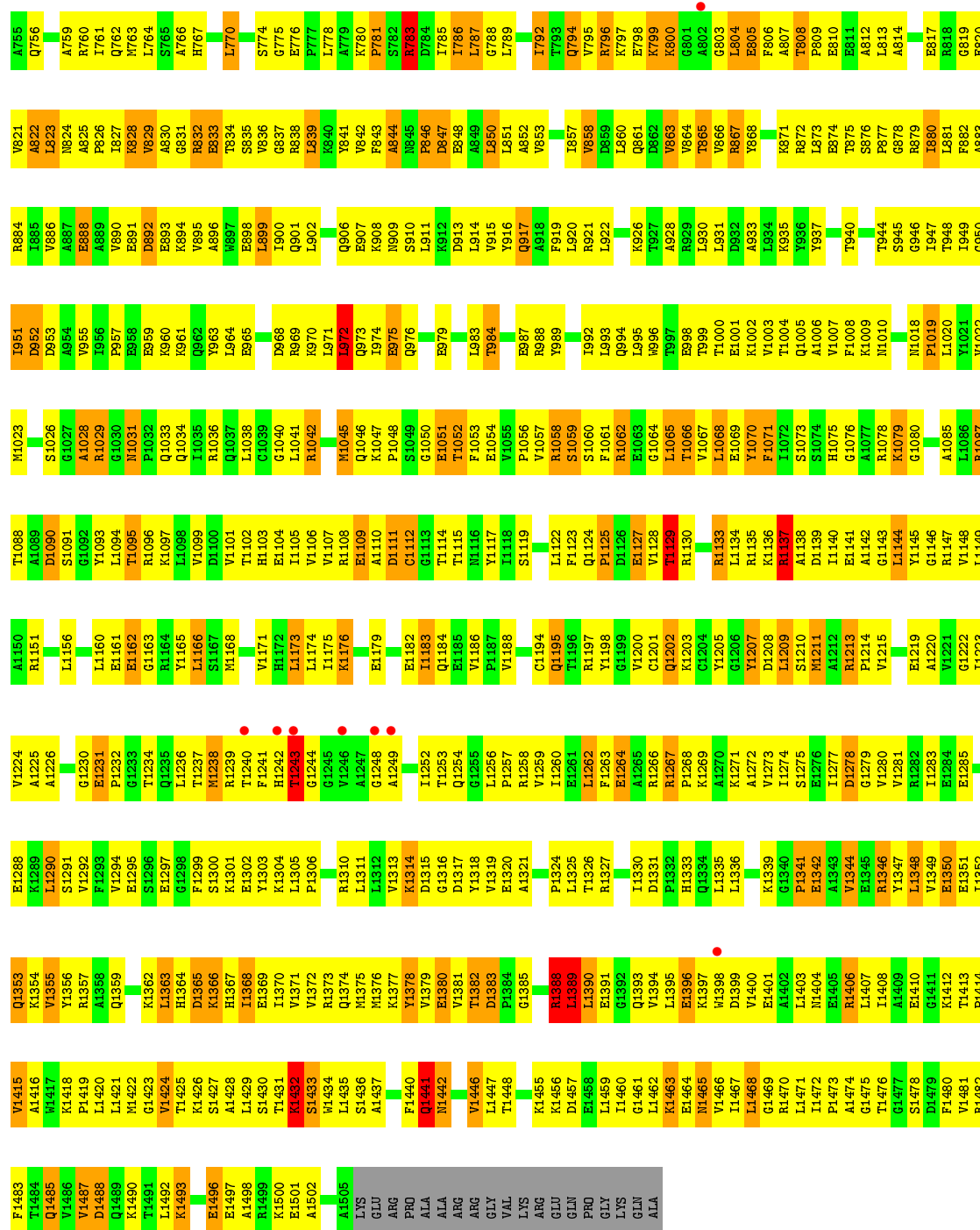


• Molecule 3: DNA-directed RNA polymerase beta' chain



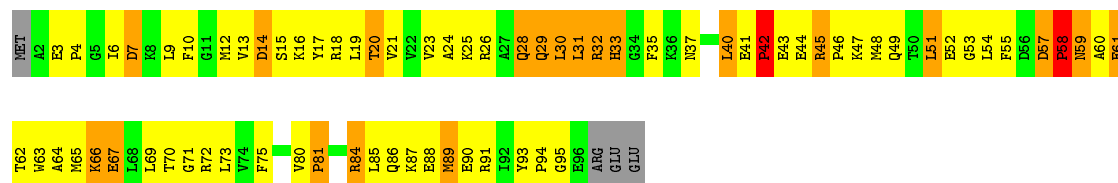
E1351	K1289	A1220	L1156	L1094	Q1033	L964	I900	G831	N768	N703	Q641	E576	A516	R455
L1352	G1289	V1221	G1157	T1095	Q1034	E965	Q901	R832	L769	R704	C642	E577	V517	R456
Q1353	L1290	G1222	V1158	R096	I1035	E966	L902	E833	L770	A705	G643	V578	P518	
K1354	S1291	I1223	R1159	K1097	R1036	A967	D903	T834	S771	P706	L644		V519	
	V092	V1224	L1160	L1098	Q1037	D968	V904	S935	T772	T707	P645	L581	L520	E459
R1357	F1293		E1161	V1099	L1038	R969	P905	V836	A773	L708	K646	L582	P521	A460
	G1230	G1230	E1162	D1100	C1039	K970	Q906	G837	H774	H709	R647	D583	P522	Q462
E1295	E1231	E1231	G1163	V1101	G1040	L971	R908	R838	G775	R710	M648	N584	D523	Q463
G1360	P1232	P1232	R1164	T1102	L1041	L972	K908	L839	E776	L711	A649	L524	L464	
F1299	G1233		Y1165	H1103	R1042	Q973	S910	K940	P777	G712	L650	R586	R525	L465
S1300	L1166	Q1234	L1167	E1104	G1043	Y974	L911	V843	L778	I713	E651	R587	P526	K466
K1301	S1167	Q1235	S1167	L1105	L1044	E975	R912	V842	A779	Q714	L652		M527	E467
E1302	M1168	L1236	M1168	V1106	M1045	Q976	D913	F843	K780	A715	P653	P590	V528	L468
Y1303	D1169	T1237	D1169	V1107	Q1046	A977	D913	K1047		F716	Q529	T592	Q530	D469
K1366	K1304	M1238	D1170	R1108	K1047	Y978	L914		R783	Q717	P655		V530	L470
H1367	L1305	T1259	V1171	E1109	P1048	E979	V915		D784	F718	F656	N593	D531	E471
P1306	R1240			A1110	S1049	M980	X916		L785	V719	L657	P594	G532	A472
K1307	F1241	R1242	L1174	D1111	G1050	N980	Q917	E848	L786	L720	L658		G533	L473
			L1175	C1112	E1051	D985	A918	A849	L787	V721	K659	D597	R534	E474
		T1243	K1176	G1113	T1052	R986	F919	L851	G788	E722	K660	F535	F535	K475
		G1244	A1177	T1114	F1053	E987	L920	A852	L789	G723	M661	P599	A536	E476
		G1245	A1178	T1115	E1054	R988	R921	V853	Y790	Q724	E662	L600	T537	L477
				Y1116	V1055		L922		Y791	S725	E663	R601	S538	L478
		G1248	E1182	N1117	P1056	T992	G923	T857	L792	J726		S602	D539	E479
		A1249	I1183	I1118	V1057	L993	R924	N858	L793	Q727	T666	L540	L540	E480
		A1250	Q1184	S1119	R1058	R994	K326	D859	Q794	L728	A667	T804	N541	M491
			E1185	V1120	S1059	L995	L860		V795	H729	P668	D805	D542	K482
		L1252	L1186	P1121	S1060	R996	Q861		R796	P730	N669	L606	L543	H483
			P1187	L1122	F1061	T997	D862		K797	L731	V670	I607	V544	P484
		Q1254	V1188	F1123	A1062	E998	R929	V863	E798		K671	S808	R545	S485
		G1255	E1189	Q1124	E1063	T999	L930	V864		E734	A672	G809	R546	R486
		L1256	S1190	P1125	G1064	T999	L831	T865	R800	A735	A673	K610	L547	L487
		P1257	P1191		L1065	E1001	Y936	R666		F736	R674	Q611	N549	R488
		R1258	L1192	V1128	V1067	K1002	X937	R867	G803	N737	M676	G612	L548	R489
		V1259	T1193	T1128	L1066	L1003	G838	X968	L804	A738	L677	R613	R550	A490
			R1130	R1130	L1068			M669	E805	D739	L677	F614	N551	K491
		L1260	Q1195		E1069	A1006	F939			F740	E678	R615	M552	K491
		E1261	T1196	S1131	L1068	I1007	R939	E874	R806	D741	R679	Q116	R553	R493
		L1262	L1197	L1132	Y1070	T940	F941	T875	A807		Q680	N617	L554	K494
		F1263	Y1198	R1133	F1071	F1008	R941	T875	T808		L681	L618	K555	L495
		G1392	A1329	L1134	S1072	P809	S942	S876	P809	Q744	R681	L619	K556	L496
		L1264	E1264	R1135	L1073	M1010	T943	P877	E810	M745	D682		L557	
		A1265	L1265	L1135	S1073	F1011	T944	G878	E811		L683		L557	E497
		R1266	C1201	K1136	A1074	F1011	T944	G878	E811		L683		L557	E497
		P1267	Q1202	R1137	H1075	E1012	S945	R879	A812	Y749	K684	R622	L558	L498
		L1333	K1203	A1138	G1076	E1013	G946	R879	L813		D685	V623	A559	V499
		Q1334	L1203	D1139	A1077	N1014	I947	R880	A814	S752	E686	D624	Q560	L500
		L1335	L1335	L1140	R1078	Y1015	T948	A883	S753		V687	Y625	G561	A501
		A1270	Y1205	E1140	K1078	Y1015	T948	A883	S753		V687	Y625	G561	A501
		K1271	G1206	E1141	K1079		I949	R884	F754		K688		A562	F502
		L1272	Y1207	A1142		P1019	G850		F817	A755	D689	R628	P563	F502
		L1273	D1208	G1143	A1082	L1020	R951	E888	G819	Q756	A690	R628	P563	L503
		L1274	L1209	L1144	D1083	Y1021	D952		E819	Q756	A690	R628	P563	L503
				Y1145	T1084	Y1022	R821		E821	A757	L691	S829	L565	L503
				G1146	A1085	V1022	R955		V821	E758	E692	L631	L566	G506
				L1147	L1086	A1023	V955		A822	A759	E693	V632	L567	N507
				V1148	R1087	Q1025	R957		L823	R760	V694	V633	R568	R508
					T1088	Q1025	R957		R824	A761	I695	G634	N569	F509
					A1089	S1026	E958		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509
					A1089	G1027	R959		R824	Q762	I695	G634	N569	F509



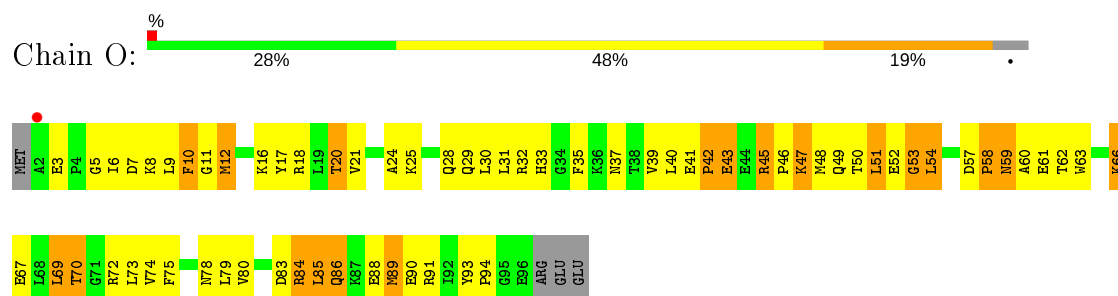


• Molecule 4: RNA polymerase omega chain

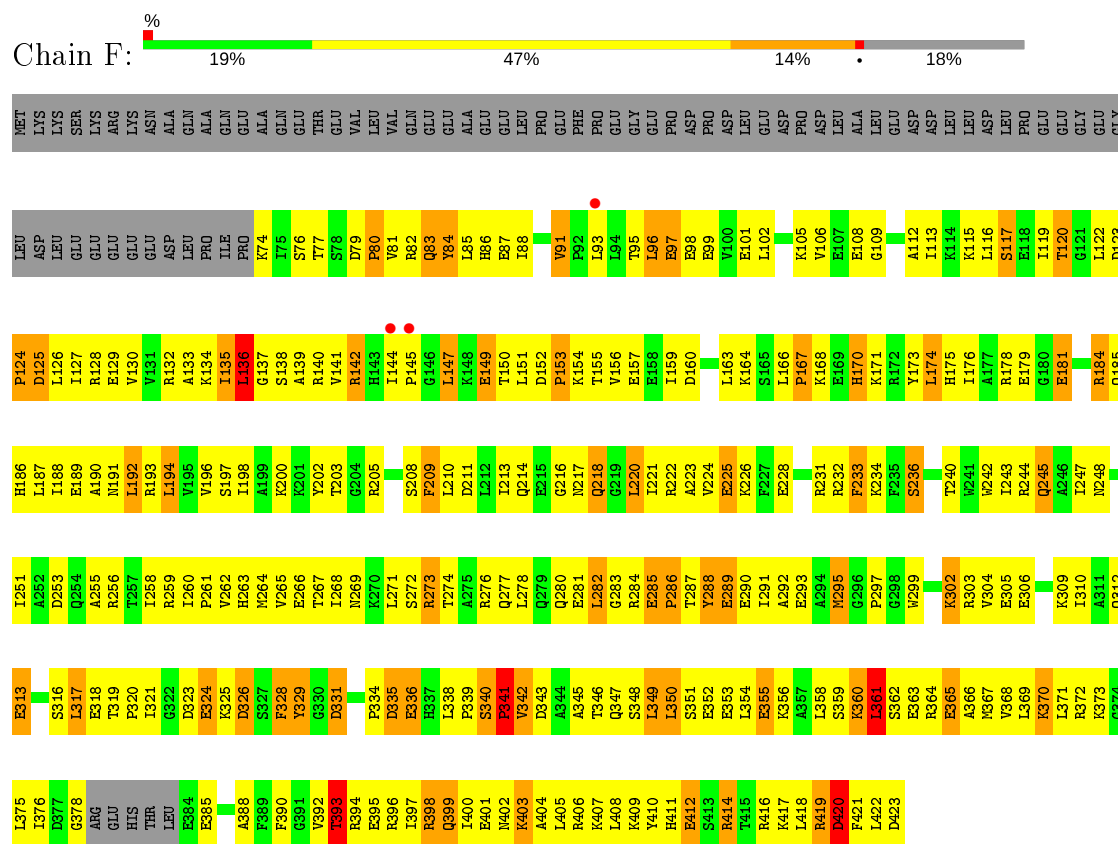
Chain E: 22% 52% 20%



- Molecule 4: RNA polymerase omega chain



- Molecule 5: RNA polymerase sigma factor rpoD



ARG	T319	D253	Q185
GLU	F320		H186
HIS	I321	R256	L187
THR	G322	T257	I188
LEU	D323	I258	E189
E384	E324	R259	A190
E385	K325	I260	N191
	D326	P261	L192
	S327	V262	R193
F390	F328	H263	L194
G391	Y329	M264	V195
V392	G330	V265	V196
T393	D331	E266	S197
R394	F332	T267	
E395	I333	I268	
R396	F334	N269	K200
	P335	K270	K201
Q398	D335	L271	Y202
Q399	E336	S272	T203
I400	H337	R273	G204
E401	L338	I274	R205
R402	F339	T275	G206
K403	S340	A276	L207
A404	P341	R277	S208
L405	V342	Q277	F209
R406	D343	L278	L210
K407	A344	Q279	D211
L408	A345	Q280	L212
K409	T346		L213
K409	Q347	G283	Q214
Y410	S348		E215
H411	I349	P286	G216
E412	L350	T287	N217
	S351	Y288	Q218
R416	E352	E289	G219
K417	E353	E290	L220
L418	R419	I291	I221
R420	E355	A292	R222
F421	K356		A223
I422	A357	M295	V224
D423	L358	G296	E225
	S359	P297	K226
	K360	G298	
	L361	W299	R232
	S362		P233
	E363	K302	K234
	R364	R303	
	E365	V304	Y238
	A366		
	K367		
	Y368	T307	W241
	L369	L308	W242
	K370	I310	I243
	L371	K309	R244
	R372	A311	Q245
	K373	Q312	A246
	G374	E313	I247
	L375	P314	N248
	I376	V315	R249
	D377	S316	A250
	G378	L317	I251
		E318	A252

4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	239.50Å 239.50Å 253.10Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.00 – 2.80 24.96 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.80) 92.0 (24.96-2.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.16 (at 2.80Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.231 , 0.268 0.227 , 0.261	Depositor DCC
R_{free} test set	21166 reflections (5.78%)	wwPDB-VP
Wilson B-factor (Å ²)	58.9	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 79.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	0.499 for -h,-k,l 0.068 for h,-h-k,-l 0.068 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	58679	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.92% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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	0.80	1/1838 (0.1%)	0.88	1/2498 (0.0%)
1	B	0.75	0/1838	0.82	2/2498 (0.1%)
1	K	0.75	0/1838	0.86	2/2498 (0.1%)
1	L	0.72	0/1838	0.78	1/2498 (0.0%)
2	C	0.84	0/8997	0.90	8/12164 (0.1%)
2	M	0.82	0/8997	0.89	7/12164 (0.1%)
3	D	0.84	0/10975	0.94	20/14836 (0.1%)
3	N	0.83	0/10975	0.93	18/14836 (0.1%)
4	E	0.84	0/783	0.97	0/1054
4	O	0.88	0/783	1.00	1/1054 (0.1%)
5	F	0.75	0/2812	0.82	3/3781 (0.1%)
5	P	0.75	0/2812	0.80	1/3781 (0.0%)
All	All	0.82	1/54486 (0.0%)	0.90	64/73662 (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
3	N	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	48	ILE	C-N	5.57	1.44	1.34

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	199	LEU	CA-CB-CG	-8.69	95.30	115.30
1	B	138	LEU	CA-CB-CG	8.01	133.72	115.30
3	N	1389	LEU	CA-CB-CG	7.77	133.18	115.30
5	P	136	LEU	CA-CB-CG	7.49	132.51	115.30
3	N	76	CYS	CA-CB-SG	6.73	126.11	114.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	N	132	TYR	Sidechain

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1806	0	1861	216	0
1	B	1806	0	1861	199	0
1	K	1806	0	1861	208	0
1	L	1806	0	1861	206	0
2	C	8829	0	8933	1184	0
2	M	8829	0	8933	1106	0
3	D	10797	0	10873	1450	0
3	N	10797	0	10873	1345	0
4	E	769	0	775	97	0
4	O	769	0	775	108	0
5	F	2771	0	2844	336	0
5	P	2771	0	2844	342	0
6	D	2	0	0	0	0
6	N	2	0	0	0	0
7	D	1	0	0	0	0
7	N	1	0	0	0	0
8	A	191	0	0	37	0
8	B	181	0	0	34	0
8	C	767	0	0	174	0
8	D	1100	0	0	234	0
8	E	93	0	0	14	0
8	F	333	0	0	58	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	K	151	0	0	30	0
8	L	179	0	0	49	0
8	M	739	0	0	195	0
8	N	1038	0	0	225	0
8	O	78	0	0	24	0
8	P	267	0	0	61	0
All	All	58679	0	54294	6401	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 59.

The worst 5 of 6401 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:M:1054:THR:HG21	2:M:1079:PRO:HB3	1.28	1.11
1:A:95:GLN:HA	1:A:146:ARG:HH12	1.12	1.08
2:C:135:VAL:HG11	2:C:407:LYS:HA	1.36	1.04
1:A:42:ARG:HH12	2:C:857:ASP:HB3	1.14	1.04
3:N:1036:ARG:HH21	3:N:1042:ARG:HA	1.20	1.04

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%)	202 (89%)	21 (9%)	4 (2%)	8	28
1	B	227/315 (72%)	202 (89%)	21 (9%)	4 (2%)	8	28
1	K	227/315 (72%)	200 (88%)	23 (10%)	4 (2%)	8	28
1	L	227/315 (72%)	204 (90%)	19 (8%)	4 (2%)	8	28
2	C	1117/1119 (100%)	917 (82%)	150 (13%)	50 (4%)	2	8

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	M	1117/1119 (100%)	907 (81%)	159 (14%)	51 (5%)	2	7
3	D	1388/1524 (91%)	1123 (81%)	191 (14%)	74 (5%)	2	6
3	N	1388/1524 (91%)	1110 (80%)	195 (14%)	83 (6%)	1	4
4	E	93/99 (94%)	74 (80%)	16 (17%)	3 (3%)	4	13
4	O	93/99 (94%)	75 (81%)	15 (16%)	3 (3%)	4	13
5	F	341/423 (81%)	283 (83%)	42 (12%)	16 (5%)	2	7
5	P	341/423 (81%)	285 (84%)	40 (12%)	16 (5%)	2	7
All	All	6786/7590 (89%)	5582 (82%)	892 (13%)	312 (5%)	2	7

5 of 312 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	178	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%)	147 (73%)	55 (27%)	0	1
1	B	202/273 (74%)	163 (81%)	39 (19%)	1	4
1	K	202/273 (74%)	152 (75%)	50 (25%)	0	2
1	L	202/273 (74%)	158 (78%)	44 (22%)	1	3
2	C	941/941 (100%)	734 (78%)	207 (22%)	1	2
2	M	941/941 (100%)	730 (78%)	211 (22%)	1	2
3	D	1123/1279 (88%)	846 (75%)	277 (25%)	0	2
3	N	1123/1279 (88%)	866 (77%)	257 (23%)	1	2
4	E	83/87 (95%)	58 (70%)	25 (30%)	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	O	83/87 (95%)	64 (77%)	19 (23%)	1	2
5	F	295/370 (80%)	228 (77%)	67 (23%)	1	2
5	P	295/370 (80%)	249 (84%)	46 (16%)	2	8
All	All	5692/6446 (88%)	4395 (77%)	1297 (23%)	1	2

5 of 1297 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	30	LEU
1	L	25	LEU
3	N	1353	GLN
5	F	80	PRO
5	F	361	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 179 such sidechains are listed below:

Mol	Chain	Res	Type
4	E	86	GLN
1	L	212	ASN
3	N	1374	GLN
5	F	337	HIS
1	K	156	HIS

5.3.3 RNA ⓘ

There are no RNA molecules 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 6 ligands modelled in this entry, 6 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

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	229/315 (72%)	-0.38	0 100 100	38, 65, 92, 117	0
1	B	229/315 (72%)	-0.27	8 (3%) 44 34	53, 95, 115, 121	0
1	K	229/315 (72%)	-0.37	0 100 100	41, 66, 93, 122	0
1	L	229/315 (72%)	-0.39	3 (1%) 77 72	52, 94, 114, 127	0
2	C	1119/1119 (100%)	-0.35	8 (0%) 87 84	23, 81, 110, 119	0
2	M	1119/1119 (100%)	-0.39	6 (0%) 91 88	27, 79, 109, 122	0
3	D	1392/1524 (91%)	-0.33	12 (0%) 84 80	17, 68, 113, 130	0
3	N	1392/1524 (91%)	-0.35	16 (1%) 80 75	27, 69, 110, 138	0
4	E	95/99 (95%)	-0.35	0 100 100	47, 85, 115, 133	0
4	O	95/99 (95%)	-0.49	1 (1%) 80 75	40, 76, 97, 111	0
5	F	345/423 (81%)	-0.40	3 (0%) 84 80	46, 84, 112, 131	0
5	P	345/423 (81%)	-0.39	5 (1%) 75 70	54, 87, 112, 125	0
All	All	6818/7590 (89%)	-0.36	62 (0%) 84 80	17, 77, 111, 138	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	130	ALA	5.1
3	D	1240	THR	4.7
3	N	1243	THR	4.7
3	D	1243	THR	4.5
5	F	145	PRO	4.4

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. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
7	MG	N	9002	1/1	0.94	0.13	64,64,64,64	0
7	MG	D	9001	1/1	0.97	0.15	67,67,67,67	0
6	ZN	N	7113	1/1	0.98	0.10	87,87,87,87	0
6	ZN	D	7058	1/1	0.98	0.07	109,109,109,109	0
6	ZN	D	7112	1/1	0.99	0.14	75,75,75,75	0
6	ZN	N	7059	1/1	0.99	0.11	100,100,100,100	0

6.5 Other polymers [i](#)

There are no such residues in this entry.