



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:10 PM GMT

PDB ID : 4KMU  
Title : X-ray crystal structure of the Escherichia coli RNA polymerase in complex with Rifampin  
Authors : Murakami, K.S.  
Deposited on : 2013-05-08  
Resolution : 3.85 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

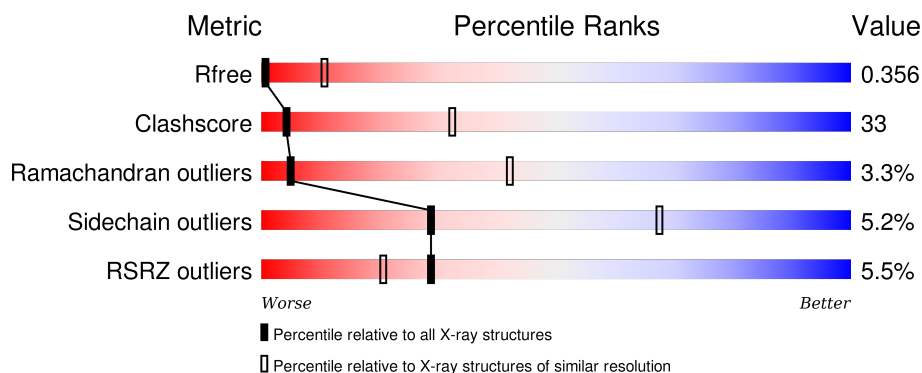
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.85 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	1000 (4.20-3.52)
Clashscore	102246	1090 (4.20-3.52)
Ramachandran outliers	100387	1046 (4.20-3.52)
Sidechain outliers	100360	1038 (4.20-3.52)
RSRZ outliers	91569	1004 (4.20-3.52)

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	329	<div> <div>3%</div> <div>51% 42% 5%</div> </div>
1	B	329	<div> <div>5%</div> <div>38% 27% 33%</div> </div>
1	F	329	<div> <div>6%</div> <div>43% 24% 30%</div> </div>
1	G	329	<div> <div>7%</div> <div>36% 27% 34%</div> </div>
2	C	1342	<div> <div>4%</div> <div>46% 48% 5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	1342	<div><div></div><div>7%</div><div>49%</div><div>45%</div><div>5%</div><div></div></div>
3	D	1407	<div><div></div><div>3%</div><div>32%</div><div>45%</div><div>5%</div><div>18%</div><div></div></div>
3	I	1407	<div><div></div><div>5%</div><div>35%</div><div>42%</div><div>5%</div><div>18%</div><div></div></div>
4	E	91	<div><div></div><div></div><div>54%</div><div>41%</div><div></div><div></div><div></div></div>
4	J	91	<div><div></div><div>4%</div><div>46%</div><div>33%</div><div></div><div></div><div>16%</div></div>
5	X	613	<div><div></div><div>6%</div><div>44%</div><div>37%</div><div></div><div></div><div>16%</div></div>
5	Y	613	<div><div></div><div>5%</div><div>37%</div><div>36%</div><div></div><div></div><div>25%</div></div>

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 56315 atoms, of which 116 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	323	Total	C	N	O	S	0	0	0
			2514	1571	443	492	8			
1	B	221	Total	C	N	O	S	0	0	0
			1706	1065	300	335	6			
1	F	229	Total	C	N	O	S	0	0	0
			1775	1106	313	350	6			
1	G	217	Total	C	N	O	S	0	0	0
			1671	1045	293	327	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			
2	H	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			
3	I	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			

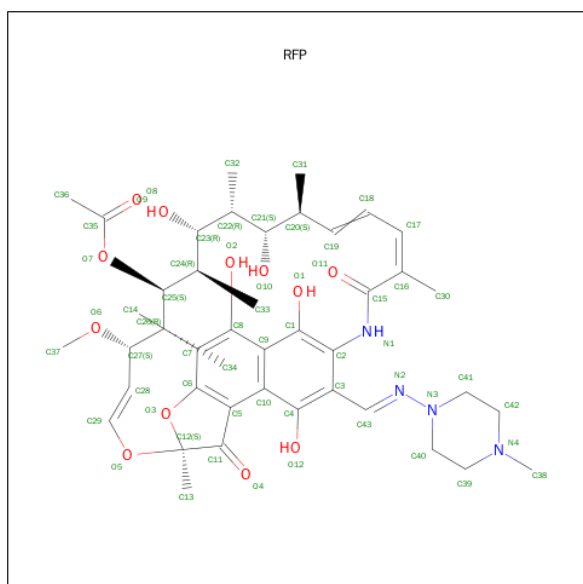
- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	J	76	Total	C	N	O	S	0	0	0
			605	368	115	121	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	X	517	Total	C	N	O	S	0	0	0
			4198	2621	745	806	26			
5	Y	458	Total	C	N	O	S	0	0	0
			3732	2335	671	703	23			

- Molecule 6 is RIFAMPICIN (three-letter code: RFP) (formula:  $C_{43}H_{58}N_4O_{12}$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	C	1	Total	C	H	N	O	0	0
			117	43	58	4	12		
6	H	1	Total	C	H	N	O	0	0
			117	43	58	4	12		

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

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

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	I	1	Total 1	Mg 1	0	0
8	D	1	Total 1	Mg 1	0	0

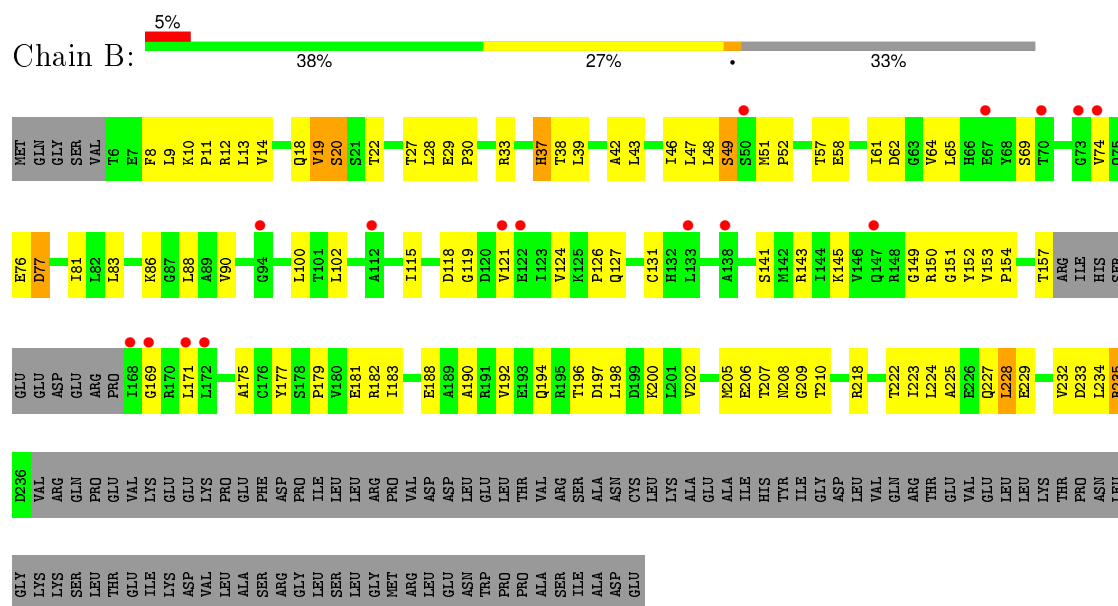
### 3 Residue-property plots

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 subunit alpha



- Molecule 1: DNA-directed RNA polymerase subunit alpha

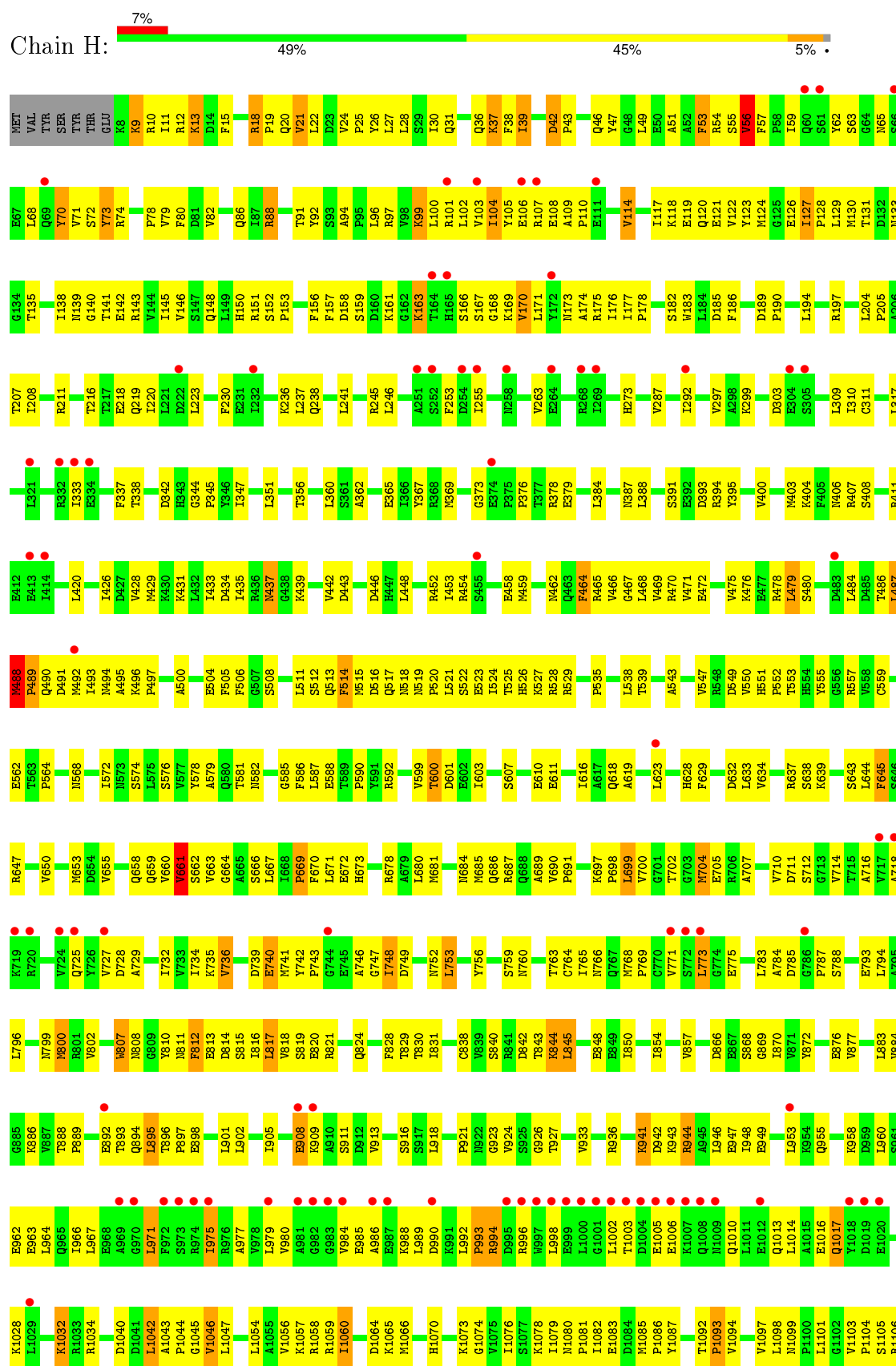


- Molecule 1: DNA-directed RNA polymerase subunit alpha



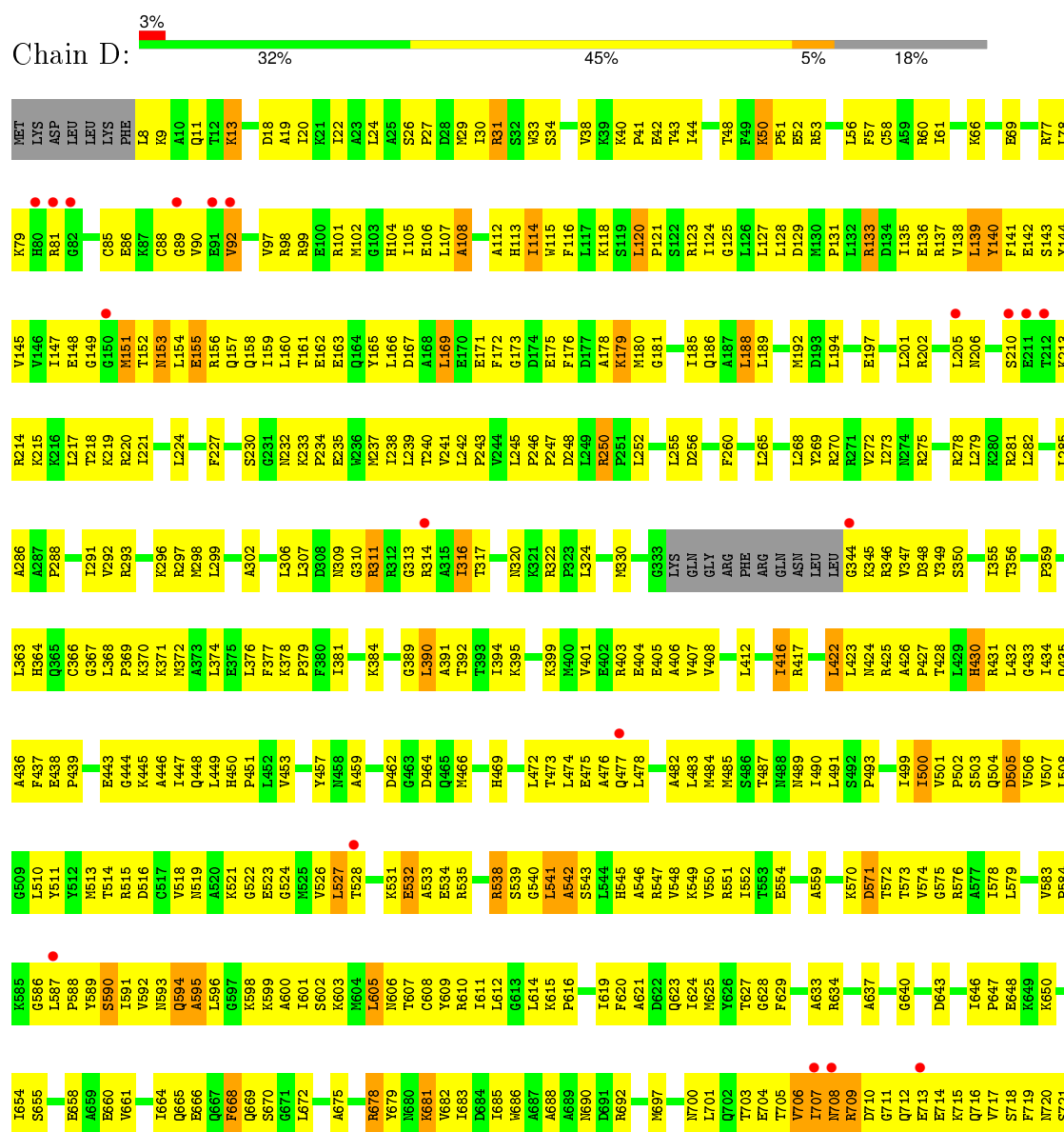


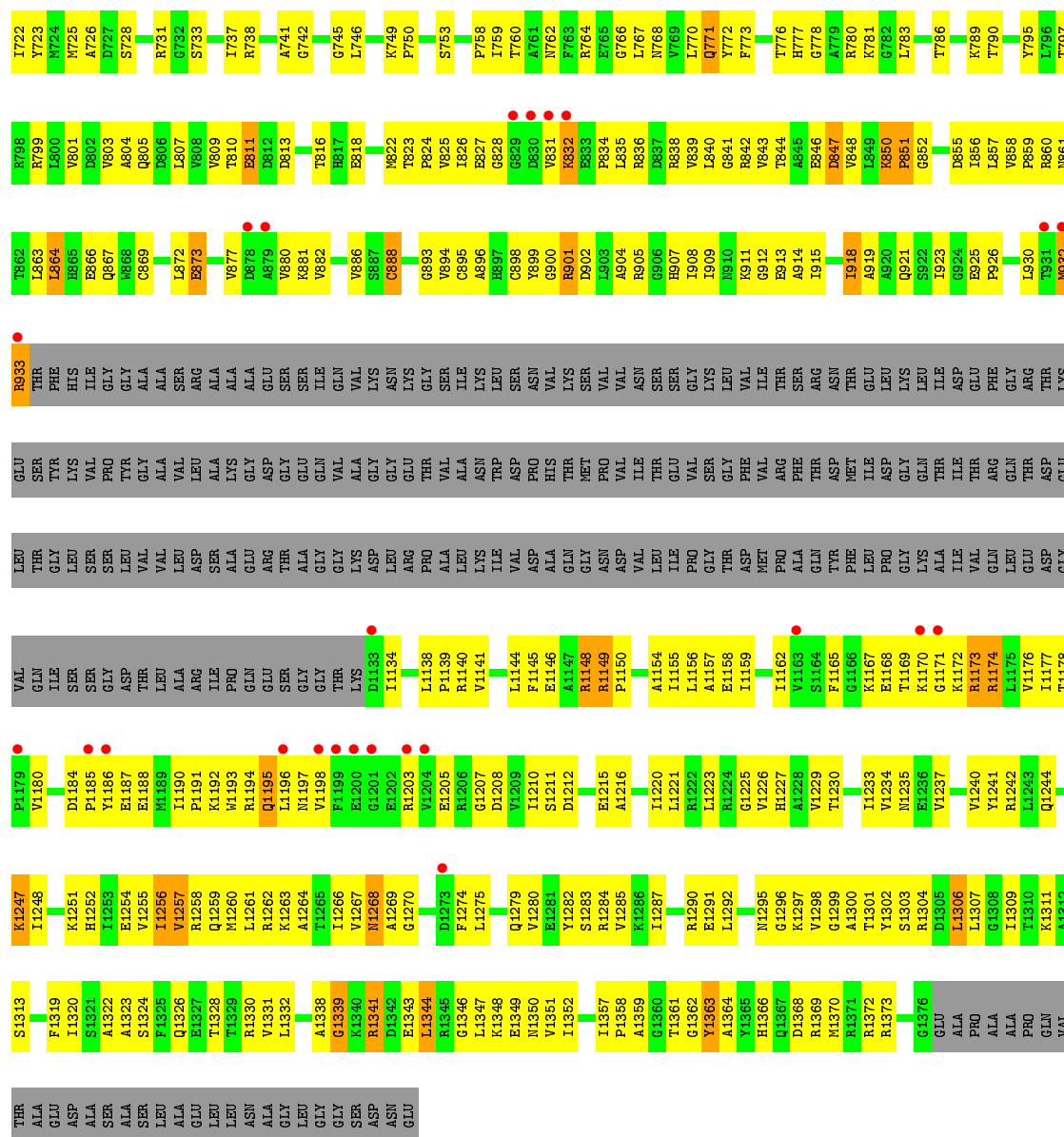
- Molecule 2: DNA-directed RNA polymerase subunit beta



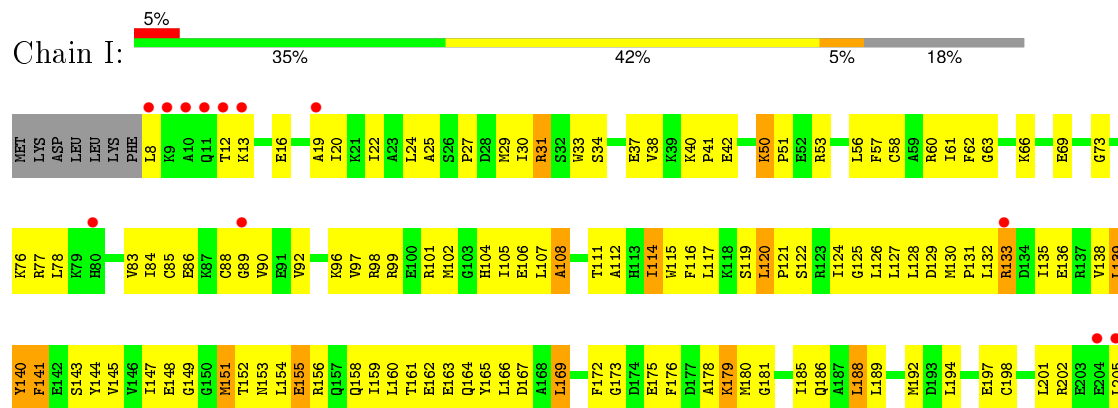


• Molecule 3: DNA-directed RNA polymerase subunit beta'

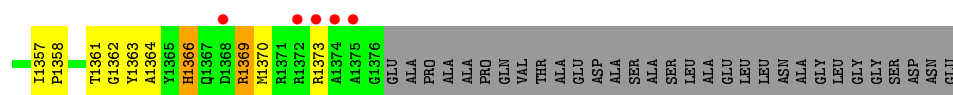




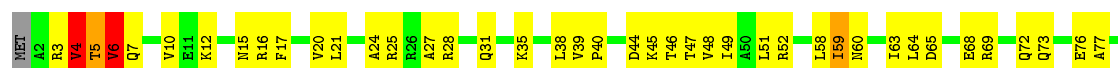
• Molecule 3: DNA-directed RNA polymerase subunit beta'



V1280	E1281	S1283	R1284	V1285	K1286	I1287	V1198	GLN	GLU	ARG	THR	GLU	GLY	ALA	Q867	V801	S721	G640	D571	L423	R346	L279	M206
E1282	S1283	R1284	V1285	K1286	I1287	V1208	R1203	GLU	ARG	THR	GLU	GLY	ASP	GLY	E873	D802	I722	D643	T572	R424	V347	K280	E207
E1203	S1283	R1284	V1285	K1286	I1287	V1209	R1204	GLY	THR	GLY	GLY	GLY	GLY	SER	A804	A803	Y723	M644	T573	A426	D348	L281	E208
E1206	R1206	V1207	D1208	V1209	I1133	I1134	R1207	GLN	THR	GLY	VAL	VAL	VAL	LYS	V806	D806	N725	V646	R576	P602	T428	L285	S210
E1207	K1206	V1207	D1208	V1209	I1133	I1134	R1208	GLY	THR	GLY	VAL	VAL	VAL	LYS	R807	D807	N726	D647	R577	S503	L429	A286	E211
E1209	V1209	S1211	A1216	D1219	L1138	P1139	R1210	ASN	THR	THR	VAL	VAL	VAL	LYS	R809	D810	S732	K649	L578	Q504	H430	A287	E212
E1211	S1211	A1216	D1219	L1138	P1139	R1140	R1211	GLY	THR	THR	VAL	VAL	VAL	LYS	R811	D812	S733	K650	L579	V507	G433	L361	K216
E1212	R1212	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R813	D814	S734	K651	L580	I434	L291	K215	K216
E1213	R1213	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R815	D816	S735	K652	L581	I435	L292	K217	K218
E1214	R1214	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R817	D818	S736	K653	L582	I436	L293	K219	K220
E1215	R1215	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R819	D820	S737	K654	L583	I437	L294	K221	K222
E1216	R1216	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R821	D822	S738	K655	L584	I438	L295	K223	K224
E1217	R1217	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R823	D824	S739	K656	L585	I439	L296	K225	K226
E1218	R1218	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R825	D826	S740	K657	L586	I440	L297	K227	K228
E1219	R1219	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R827	D828	S741	K658	L587	I441	L298	K229	K230
E1220	R1220	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R829	D830	S742	K659	L588	I442	L299	K231	K232
E1221	R1221	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R831	D832	S743	K660	L589	I443	L300	K233	K234
E1222	R1222	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R833	D834	S744	K661	L590	I444	L301	K235	K236
E1223	R1223	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R835	D836	S745	K662	L591	I445	L302	K237	K238
E1224	R1224	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R837	D838	S746	K663	L592	I446	L303	K239	K240
E1225	R1225	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R839	D840	S747	K664	L593	I447	L304	K241	K242
E1226	R1226	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R841	D842	S748	K665	L594	I448	L305	K243	K244
E1227	R1227	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R843	D844	S749	K666	L595	I449	L306	K245	K246
E1228	R1228	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R845	D846	S750	K667	L596	I450	L307	K247	K248
E1229	R1229	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R847	D848	S751	K668	L597	I451	L308	K249	K250
E1230	R1230	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R849	D850	S752	K669	L598	I452	L309	K251	K252
E1231	R1231	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R851	D852	S753	K670	L599	I453	L310	K253	K254
E1232	R1232	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R853	D854	S754	K671	L600	I454	L311	K255	K256
E1233	R1233	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R855	D856	S755	K672	L601	I455	L312	K257	K258
E1234	R1234	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R857	D858	S756	K673	L602	I456	L313	K259	K260
E1235	R1235	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R859	D860	S757	K674	L603	I457	L314	K261	K262
E1236	R1236	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R861	D862	S758	K675	L604	I458	L315	K263	K264
E1237	R1237	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R863	D864	S759	K676	L605	I459	L316	K265	K266
E1238	R1238	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R865	D866	S760	K677	L606	I460	L317	K267	K268
E1239	R1239	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R867	D868	S761	K678	L607	I461	L318	K269	K270
E1240	R1240	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R869	D870	S762	K679	L608	I462	L319	K271	K272
E1241	R1241	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R871	D872	S763	K680	L609	I463	L320	K273	K274
E1242	R1242	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R873	D874	S764	K681	L610	I464	L321	K275	K276
E1243	R1243	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R875	D876	S765	K682	L611	I465	L322	K277	K278
E1244	R1244	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R877	D878	S766	K683	L612	I466	L323	K279	K280
E1245	R1245	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R879	D880	S767	K684	L613	I467	L324	K281	K282
E1246	R1246	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R881	D882	S768	K685	L614	I468	L325	K283	K284
E1247	R1247	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R883	D884	S769	K686	L615	I469	L326	K285	K286
E1248	R1248	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R885	D886	S770	K687	L616	I470	L327	K287	K288
E1249	R1249	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R887	D888	S771	K688	L617	I471	L328	K289	K290
E1250	R1250	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R889	D890	S772	K689	L618	I472	L329	K291	K292
E1251	R1251	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R891	D892	S773	K690	L619	I473	L330	K293	K294
E1252	R1252	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R893	D894	S774	K691	L620	I474	L331	K295	K296
E1253	R1253	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R895	D896	S775	K692	L621	I475	L332	K297	K298
E1254	R1254	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R897	D898	S776	K693	L622	I476	L333	K299	K300
E1255	R1255	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R899	D899	S777	K694	L623	I477	L334	K301	K302
E1256	R1256	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R901	D900	S778	K695	L624	I478	L335	K303	K304
E1257	R1257	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R903	D902	S779	K696	L625	I479	L336	K305	K306
E1258	R1258	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R905	D904	S780	K697	L626	I480	L337	K307	K308
E1259	R1259	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R907	D906	S781	K698	L627	I481	L338	K309	K310
E1260	R1260	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R909	D908	S782	K699	L628	I482	L339	K311	K312
E1261	R1261	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R911	D910	S783	K700	L629	I483	L340	K313	K314
E1262	R1262	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R913	D912	S784	K701	L630	I484	L341	K315	K316
E1263	R1263	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R915	D914	S785	K702	L631	I485	L342	K317	K318
E1264	R1264	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R917	D916	S786	K703	L632	I486	L343	K319	K320
E1265	R1265	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R919	D918	S787	K704	L633	I487	L344	K321	K322
E1266	R1266	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R921	D920	S788	K705	L634	I488	L345	K323	K324
E1267	R1267	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R923	D922	S789	K706	L635	I489	L346	K325	K326
E1268	R1268	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R925	D924	S790	K707	L636	I490	L347	K327	K328
E1269	R1269	T218	K221	L222	L223	F227	S230	GLY	THR	THR	VAL	VAL	VAL	LYS	R927	D926	S791	K708	L637	I491	L348		



- Molecule 4: DNA-directed RNA polymerase subunit omega



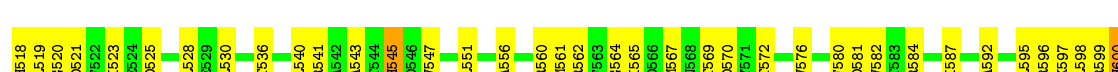
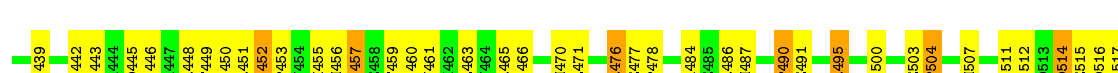
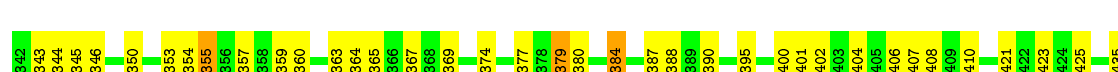
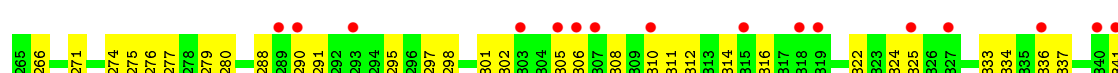
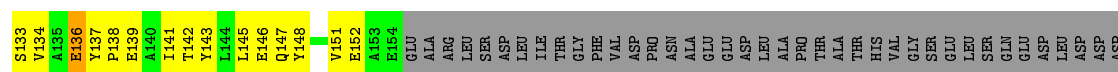
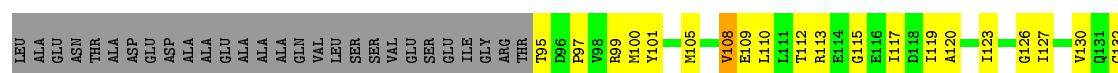
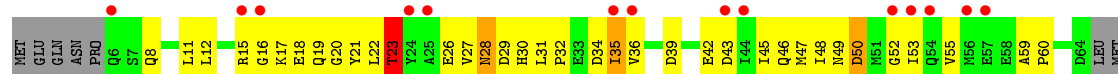
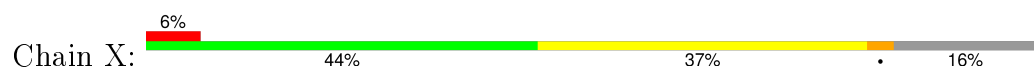
R91

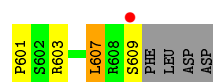
- Molecule 4: DNA-directed RNA polymerase subunit omega



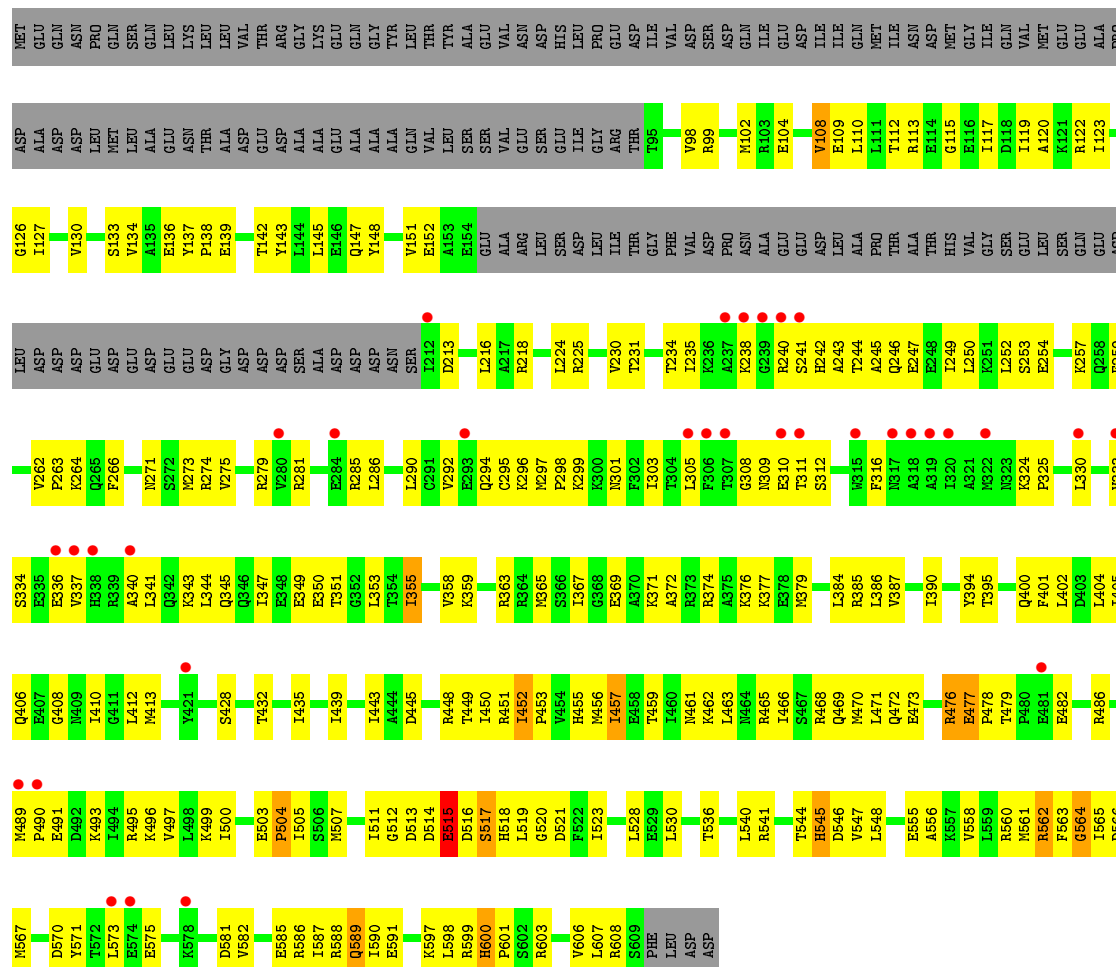
I1E  
ALA  
GLY  
GLY  
ARG  
ARG

- Molecule 5: RNA polymerase sigma factor RpoD





● Molecule 5: RNA polymerase sigma factor RpoD



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	184.52Å 203.87Å 307.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.97 – 3.85 30.75 – 3.81	Depositor EDS
% Data completeness (in resolution range)	92.4 (29.97-3.85) 85.5 (30.75-3.81)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.25 (at 3.75Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.264 , 0.321 0.310 , 0.356	Depositor DCC
$R_{free}$ test set	4808 reflections (4.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	112.9	Xtriage
Anisotropy	0.299	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 42.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.20$	Xtriage
Outliers	0 of 103271 reflections	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	56315	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.67% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, RFP, 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	0.19	0/2548	0.38	0/3454
1	B	0.20	0/1725	0.42	0/2337
1	F	0.20	0/1797	0.41	0/2436
1	G	0.20	0/1690	0.41	0/2290
2	C	0.21	0/10690	0.42	0/14423
2	H	0.22	0/10690	0.42	0/14423
3	D	0.20	0/9198	0.42	0/12413
3	I	0.20	0/9198	0.42	0/12413
4	E	0.19	0/710	0.38	0/956
4	J	0.19	0/607	0.39	0/817
5	X	0.20	0/4253	0.39	0/5719
5	Y	0.20	0/3783	0.39	0/5083
All	All	0.21	0/56889	0.41	0/76764

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2514	0	2566	170	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1706	0	1738	108	0
1	F	1775	0	1800	77	0
1	G	1671	0	1706	92	0
2	C	10523	0	10546	801	0
2	H	10523	0	10546	701	0
3	D	9060	0	9257	808	0
3	I	9060	0	9257	751	0
4	E	708	0	719	51	0
4	J	605	0	612	44	0
5	X	4198	0	4250	243	0
5	Y	3732	0	3809	211	0
6	C	59	58	56	9	0
6	H	59	58	56	14	0
7	D	2	0	0	0	0
7	I	2	0	0	0	0
8	D	1	0	0	0	0
8	I	1	0	0	0	0
All	All	56199	116	56918	3754	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 33.

The worst 5 of 3754 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:H:1119:MET:HG2	2:H:1228:GLY:HA2	1.20	1.17
3:D:1173:ARG:HA	3:D:1174:ARG:HB2	1.26	1.14
3:I:610:ARG:HG3	3:I:864:LEU:HD13	1.29	1.14
2:C:700:VAL:HG11	2:C:1114:GLU:HG3	1.30	1.12
3:I:1173:ARG:HA	3:I:1174:ARG:HB2	1.28	1.12

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	321/329 (98%)	254 (79%)	52 (16%)	15 (5%)	3	32
1	B	217/329 (66%)	188 (87%)	23 (11%)	6 (3%)	6	45
1	F	227/329 (69%)	194 (86%)	28 (12%)	5 (2%)	8	50
1	G	213/329 (65%)	188 (88%)	20 (9%)	5 (2%)	8	49
2	C	1333/1342 (99%)	1066 (80%)	225 (17%)	42 (3%)	5	42
2	H	1333/1342 (99%)	1065 (80%)	222 (17%)	46 (4%)	4	41
3	D	1154/1407 (82%)	919 (80%)	193 (17%)	42 (4%)	4	40
3	I	1154/1407 (82%)	925 (80%)	192 (17%)	37 (3%)	5	42
4	E	88/91 (97%)	76 (86%)	7 (8%)	5 (6%)	2	28
4	J	74/91 (81%)	64 (86%)	5 (7%)	5 (7%)	1	23
5	X	511/613 (83%)	444 (87%)	54 (11%)	13 (2%)	7	47
5	Y	454/613 (74%)	410 (90%)	33 (7%)	11 (2%)	7	48
All	All	7079/8222 (86%)	5793 (82%)	1054 (15%)	232 (3%)	5	42

5 of 232 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	PRO
1	B	20	SER
1	B	52	PRO
2	C	21	VAL
2	C	39	ILE

### 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	281/286 (98%)	270 (96%)	11 (4%)	39	73
1	B	189/286 (66%)	184 (97%)	5 (3%)	54	80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	F	197/286 (69%)	191 (97%)	6 (3%)	48	78
1	G	185/286 (65%)	180 (97%)	5 (3%)	52	79
2	C	1150/1157 (99%)	1084 (94%)	66 (6%)	25	65
2	H	1150/1157 (99%)	1084 (94%)	66 (6%)	25	65
3	D	971/1168 (83%)	911 (94%)	60 (6%)	23	62
3	I	971/1168 (83%)	913 (94%)	58 (6%)	24	63
4	E	74/75 (99%)	72 (97%)	2 (3%)	52	79
4	J	65/75 (87%)	63 (97%)	2 (3%)	47	78
5	X	460/540 (85%)	442 (96%)	18 (4%)	39	73
5	Y	407/540 (75%)	388 (95%)	19 (5%)	32	70
All	All	6100/7024 (87%)	5782 (95%)	318 (5%)	29	67

5 of 318 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	6	VAL
2	H	46	GLN
3	I	1306	LEU
5	X	136	GLU
5	X	607	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 97 such sidechains are listed below:

Mol	Chain	Res	Type
4	E	31	GLN
5	X	437	GLN
5	Y	242	HIS
5	X	28	ASN
5	X	242	HIS

### 5.3.3 RNA ⓘ

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 8 ligands modelled in this entry, 6 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	$\# Z  > 2$
6	RFP	C	1401	-	63,63,63	2.00	10 (15%)	82,94,94	2.33	30 (36%)
6	RFP	H	1401	-	63,63,63	2.06	9 (14%)	82,94,94	2.07	26 (31%)

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
6	RFP	C	1401	-	-	0/60/85/85	0/1/5/5
6	RFP	H	1401	-	-	0/60/85/85	0/1/5/5

The worst 5 of 19 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	1401	RFP	C12-C11	-3.63	1.36	1.54
6	H	1401	RFP	C12-C11	-3.59	1.36	1.54
6	H	1401	RFP	O7-C25	-3.55	1.39	1.44
6	C	1401	RFP	O7-C25	-2.88	1.40	1.44
6	C	1401	RFP	C18-C19	2.04	1.42	1.33

The worst 5 of 56 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	1401	RFP	C2-C3-C43	-7.45	115.01	123.56
6	C	1401	RFP	O4-C11-C5	-4.80	122.02	132.02
6	H	1401	RFP	O4-C11-C5	-4.62	122.40	132.02
6	H	1401	RFP	C17-C18-C19	-4.57	112.60	124.18
6	H	1401	RFP	C12-O3-C6	-4.31	102.36	107.82

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	1401	RFP	9	0
6	H	1401	RFP	14	0

## 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 [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, 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	323/329 (98%)	0.12	11 (3%)	49	37	0, 73, 165, 263	0
1	B	221/329 (67%)	0.41	16 (7%)	18	12	3, 97, 189, 266	0
1	F	229/329 (69%)	0.52	19 (8%)	14	9	16, 121, 201, 266	0
1	G	217/329 (65%)	0.50	22 (10%)	9	7	39, 111, 186, 215	0
2	C	1335/1342 (99%)	-0.01	47 (3%)	48	36	0, 48, 166, 284	0
2	H	1335/1342 (99%)	0.22	90 (6%)	21	13	1, 86, 201, 341	0
3	D	1160/1407 (82%)	0.00	43 (3%)	45	34	0, 40, 157, 284	0
3	I	1160/1407 (82%)	0.16	70 (6%)	25	17	1, 52, 180, 322	0
4	E	90/91 (98%)	-0.31	0	100	100	0, 40, 109, 159	0
4	J	76/91 (83%)	0.26	4 (5%)	30	22	5, 76, 155, 167	0
5	X	517/613 (84%)	0.26	36 (6%)	19	13	3, 99, 228, 365	0
5	Y	458/613 (74%)	0.19	33 (7%)	18	12	2, 102, 219, 328	0
All	All	7121/8222 (86%)	0.15	391 (5%)	29	21	0, 70, 190, 365	0

The worst 5 of 391 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	1002	LEU	16.2
2	H	1001	GLY	12.4
3	I	10	ALA	10.7
5	X	319	ALA	10.3
2	C	231	GLU	9.6

### 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
6	RFP	H	1401	59/59	0.83	0.36	1.33	20,20,20,20	0
6	RFP	C	1401	59/59	0.90	0.29	0.81	20,20,20,20	0
7	ZN	I	1502	1/1	0.96	0.17	-0.54	49,49,49,49	0
7	ZN	D	1502	1/1	0.97	0.16	-0.71	8,8,8,8	0
7	ZN	I	1501	1/1	0.97	0.05	-1.46	60,60,60,60	0
7	ZN	D	1501	1/1	0.99	0.07	-1.48	54,54,54,54	0
8	MG	D	1503	1/1	0.89	0.16	-	24,24,24,24	0
8	MG	I	1503	1/1	0.97	0.70	-	20,20,20,20	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.