



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 09:27 PM GMT

PDB ID : 4ZH4  
Title : Crystal structure of Escherichia coli RNA polymerase in complex with CBRP18  
Authors : Feng, Y.; Ebright, R.H.  
Deposited on : 2015-04-24  
Resolution : 3.99 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

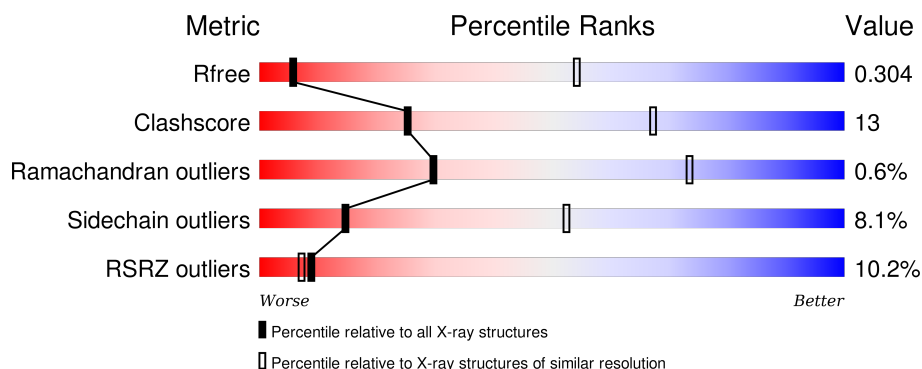
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.99 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	1010 (4.42-3.56)
Clashscore	102246	1052 (4.40-3.60)
Ramachandran outliers	100387	1005 (4.40-3.60)
Sidechain outliers	100360	1013 (4.42-3.58)
RSRZ outliers	91569	1013 (4.42-3.56)

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	335	<div> <div>7%</div> <div>60%</div> <div>26%</div> <div>10%</div> </div>
1	B	335	<div> <div>6%</div> <div>38%</div> <div>24%</div> <div>36%</div> </div>
1	G	335	<div> <div>5%</div> <div>41%</div> <div>24%</div> <div>33%</div> </div>
1	H	335	<div> <div>11%</div> <div>38%</div> <div>24%</div> <div>36%</div> </div>
2	C	1342	<div> <div>9%</div> <div>66%</div> <div>31%</div> <div>•</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	I	1342	
3	D	1407	
3	J	1407	
4	E	91	
4	K	91	
5	F	613	
5	L	613	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	4OE	C	2001	-	-	-	X
6	4OE	I	2001	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 57539 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	302	Total	C	N	O	S	0	0	0
			2328	1456	413	451	8			
1	B	216	Total	C	N	O	S	0	0	0
			1667	1041	294	326	6			
1	G	224	Total	C	N	O	S	0	0	0
			1730	1076	308	340	6			
1	H	216	Total	C	N	O	S	0	0	0
			1662	1038	292	326	6			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	MET	-	expression tag	UNP P0A7Z4
A	-4	HIS	-	expression tag	UNP P0A7Z4
A	-3	HIS	-	expression tag	UNP P0A7Z4
A	-2	HIS	-	expression tag	UNP P0A7Z4
A	-1	HIS	-	expression tag	UNP P0A7Z4
A	0	HIS	-	expression tag	UNP P0A7Z4
A	1	HIS	-	expression tag	UNP P0A7Z4
B	-5	MET	-	expression tag	UNP P0A7Z4
B	-4	HIS	-	expression tag	UNP P0A7Z4
B	-3	HIS	-	expression tag	UNP P0A7Z4
B	-2	HIS	-	expression tag	UNP P0A7Z4
B	-1	HIS	-	expression tag	UNP P0A7Z4
B	0	HIS	-	expression tag	UNP P0A7Z4
B	1	HIS	-	expression tag	UNP P0A7Z4
G	-5	MET	-	expression tag	UNP P0A7Z4
G	-4	HIS	-	expression tag	UNP P0A7Z4
G	-3	HIS	-	expression tag	UNP P0A7Z4
G	-2	HIS	-	expression tag	UNP P0A7Z4
G	-1	HIS	-	expression tag	UNP P0A7Z4
G	0	HIS	-	expression tag	UNP P0A7Z4
G	1	HIS	-	expression tag	UNP P0A7Z4

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
H	-5	MET	-	expression tag	UNP P0A7Z4
H	-4	HIS	-	expression tag	UNP P0A7Z4
H	-3	HIS	-	expression tag	UNP P0A7Z4
H	-2	HIS	-	expression tag	UNP P0A7Z4
H	-1	HIS	-	expression tag	UNP P0A7Z4
H	0	HIS	-	expression tag	UNP P0A7Z4
H	1	HIS	-	expression tag	UNP P0A7Z4

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1340	Total	C	N	O	S	0	0	0
			10570	6631	1841	2055	43			
2	I	1340	Total	C	N	O	S	0	0	0
			10566	6629	1840	2054	43			

- 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
			8992	5653	1608	1686	45			
3	J	1319	Total	C	N	O	S	0	0	0
			10254	6443	1824	1939	48			

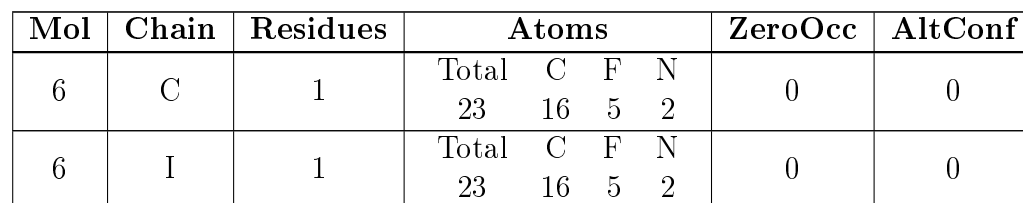
- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	89	Total	C	N	O	S	0	0	0
			691	421	129	140	1			
4	K	79	Total	C	N	O	S	0	0	0
			627	382	118	126	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	542	Total	C	N	O	S	0	0	0
			4204	2625	752	801	26			
5	L	539	Total	C	N	O	S	0	0	0
			4196	2619	749	802	26			

- Molecule 6 is 5-(4-fluorophenyl)-4-[4-fluoro-3-(trifluoromethyl)phenyl]-1H-pyrazole (three-letter code: 4OE) (formula: C<sub>16</sub>H<sub>9</sub>F<sub>5</sub>N<sub>2</sub>).



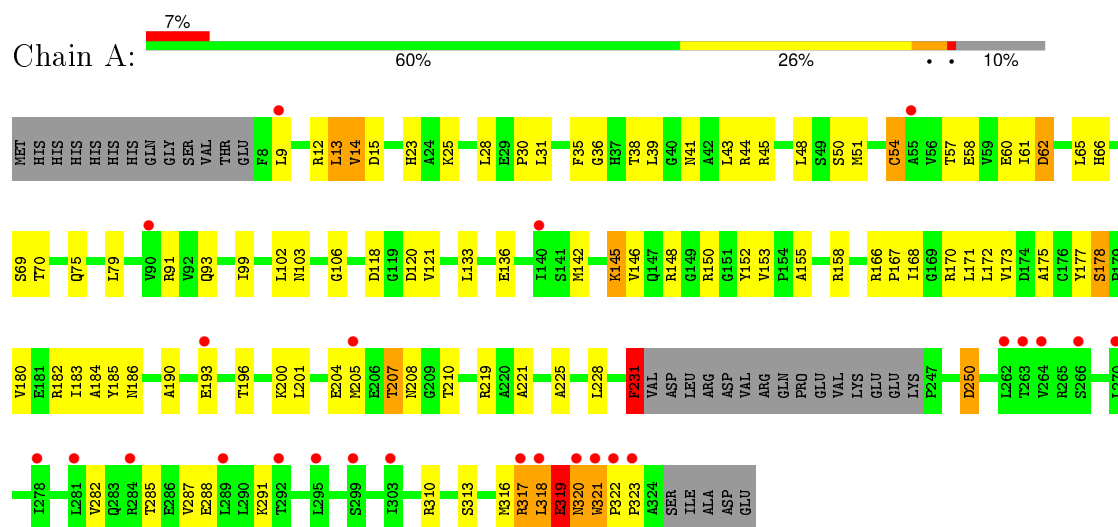
- | Mol | Chain | Residues | Atoms           | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 7   | J     | 1        | Total Mg<br>1 1 | 0       | 0       |
| 7   | D     | 1        | Total Mg<br>1 1 | 0       | 0       |

- | Mol | Chain | Residues | Atoms           | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 8   | J     | 2        | Total Zn<br>2 2 | 0       | 0       |
| 8   | D     | 2        | Total Zn<br>2 2 | 0       | 0       |

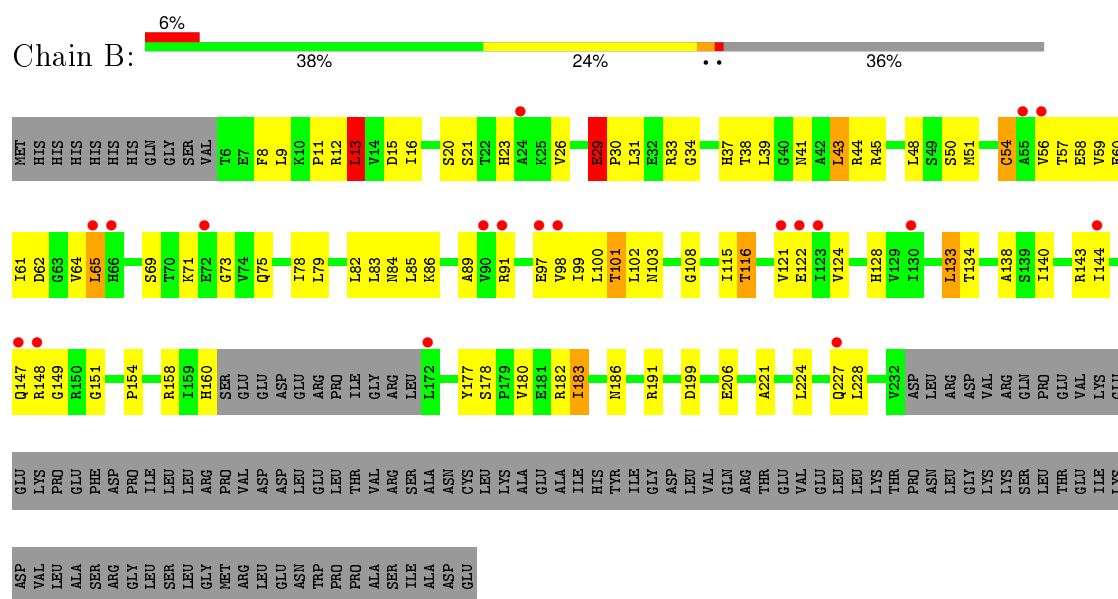
### 3 Residue-property plots [i](#)

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.

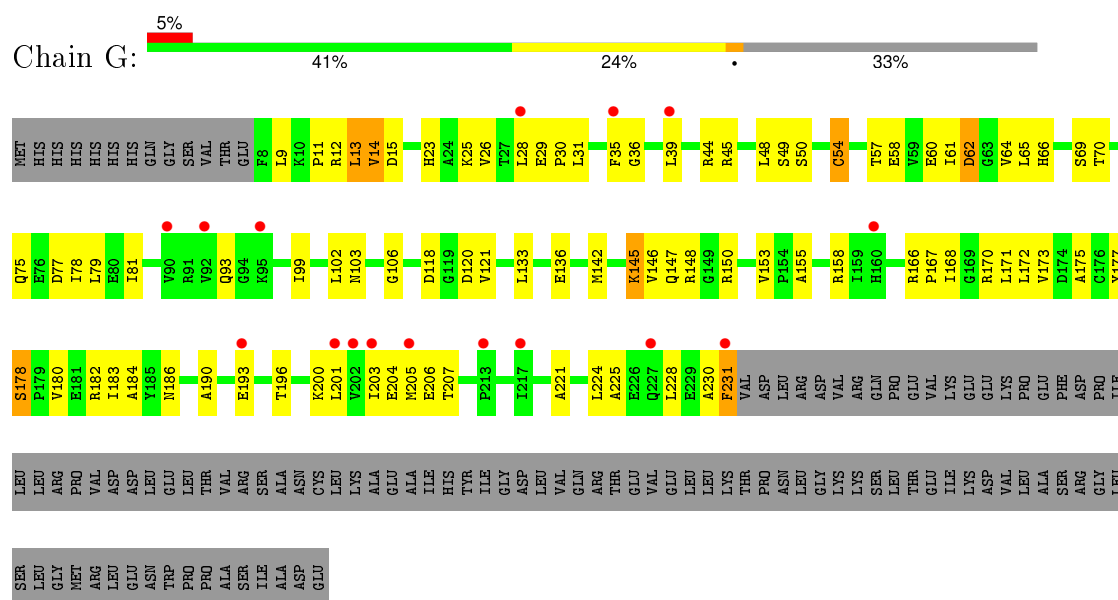
#### • Molecule 1: DNA-directed RNA polymerase subunit alpha



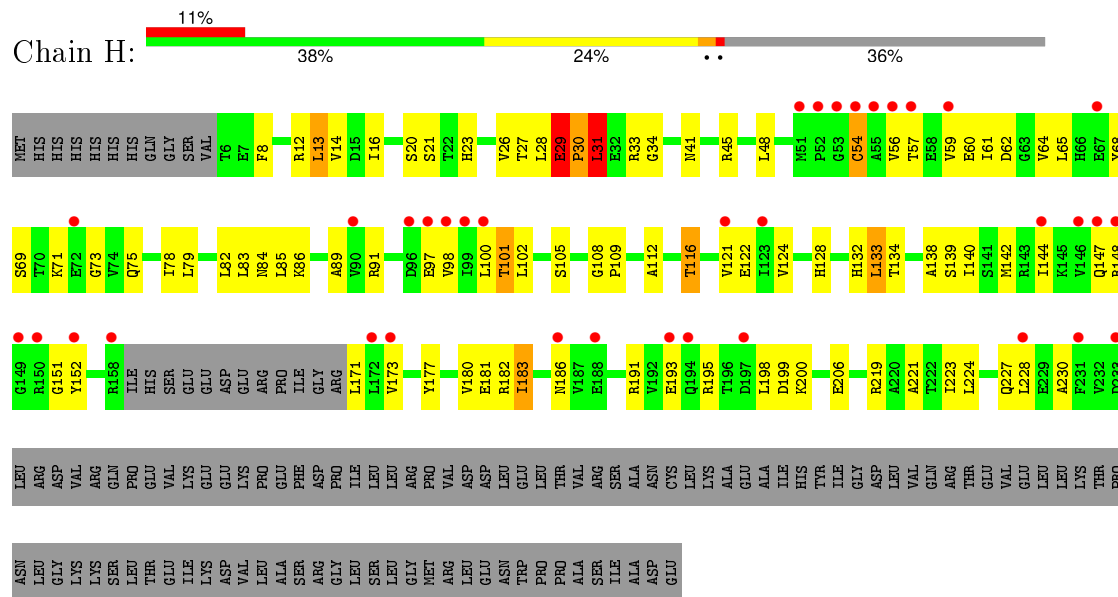
#### • Molecule 1: DNA-directed RNA polymerase subunit alpha



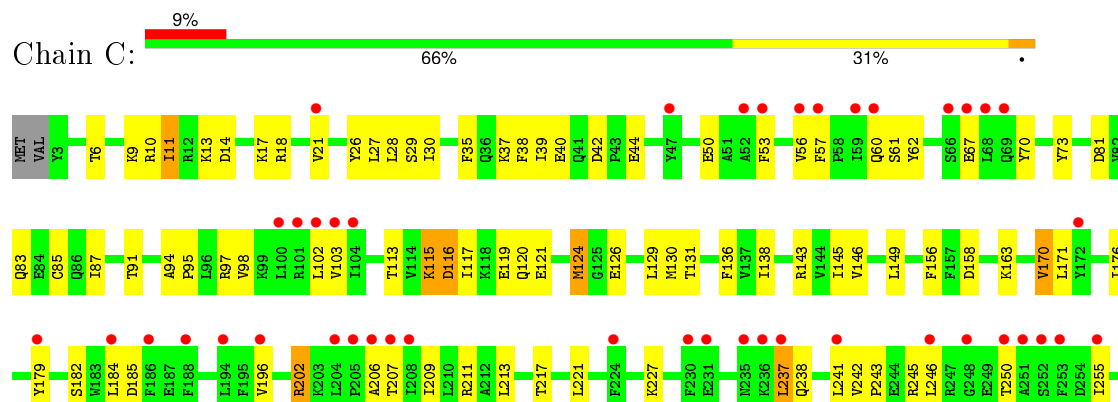
#### • Molecule 1: DNA-directed RNA polymerase subunit alpha



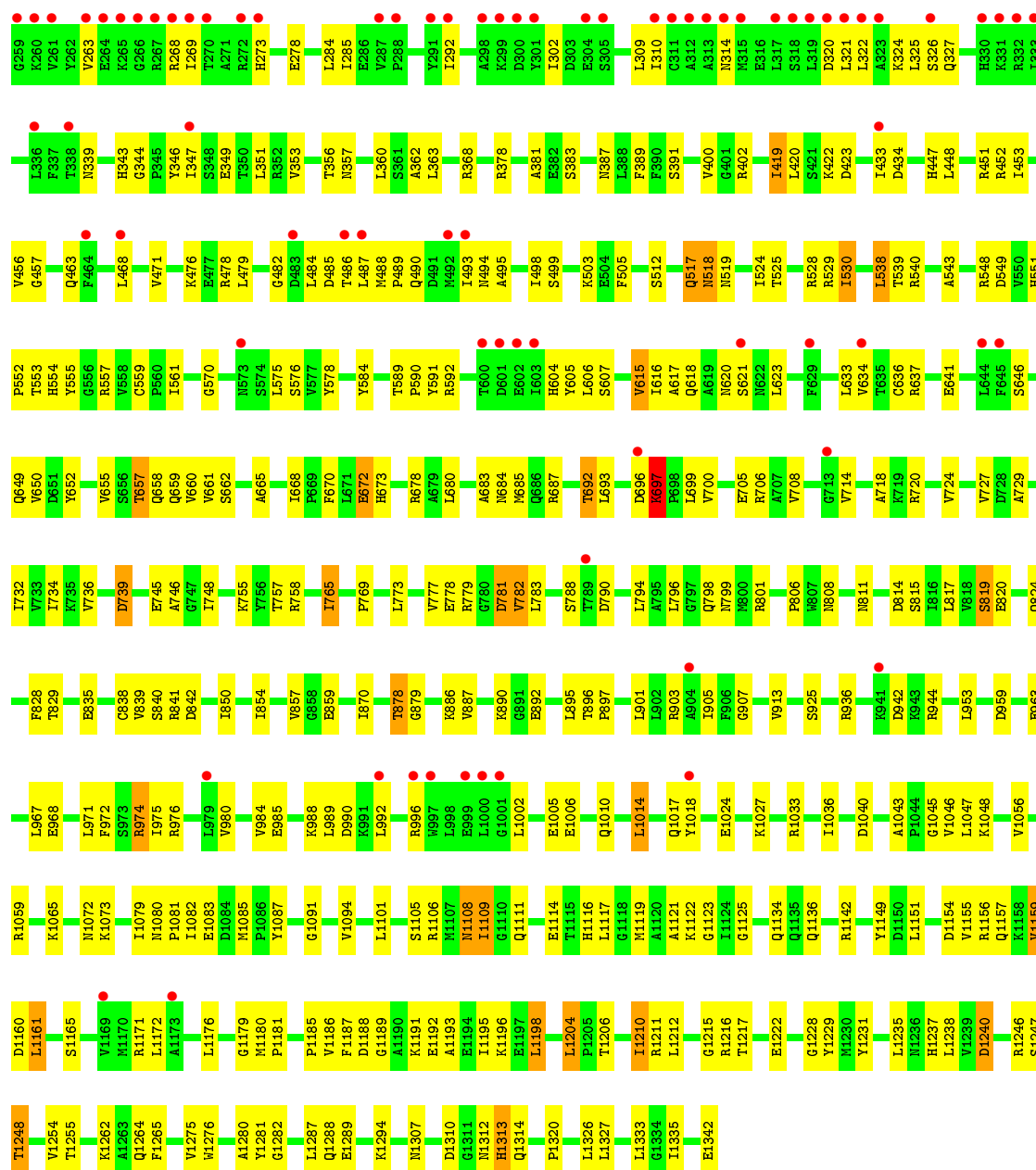
- Molecule 1: DNA-directed RNA polymerase subunit alpha

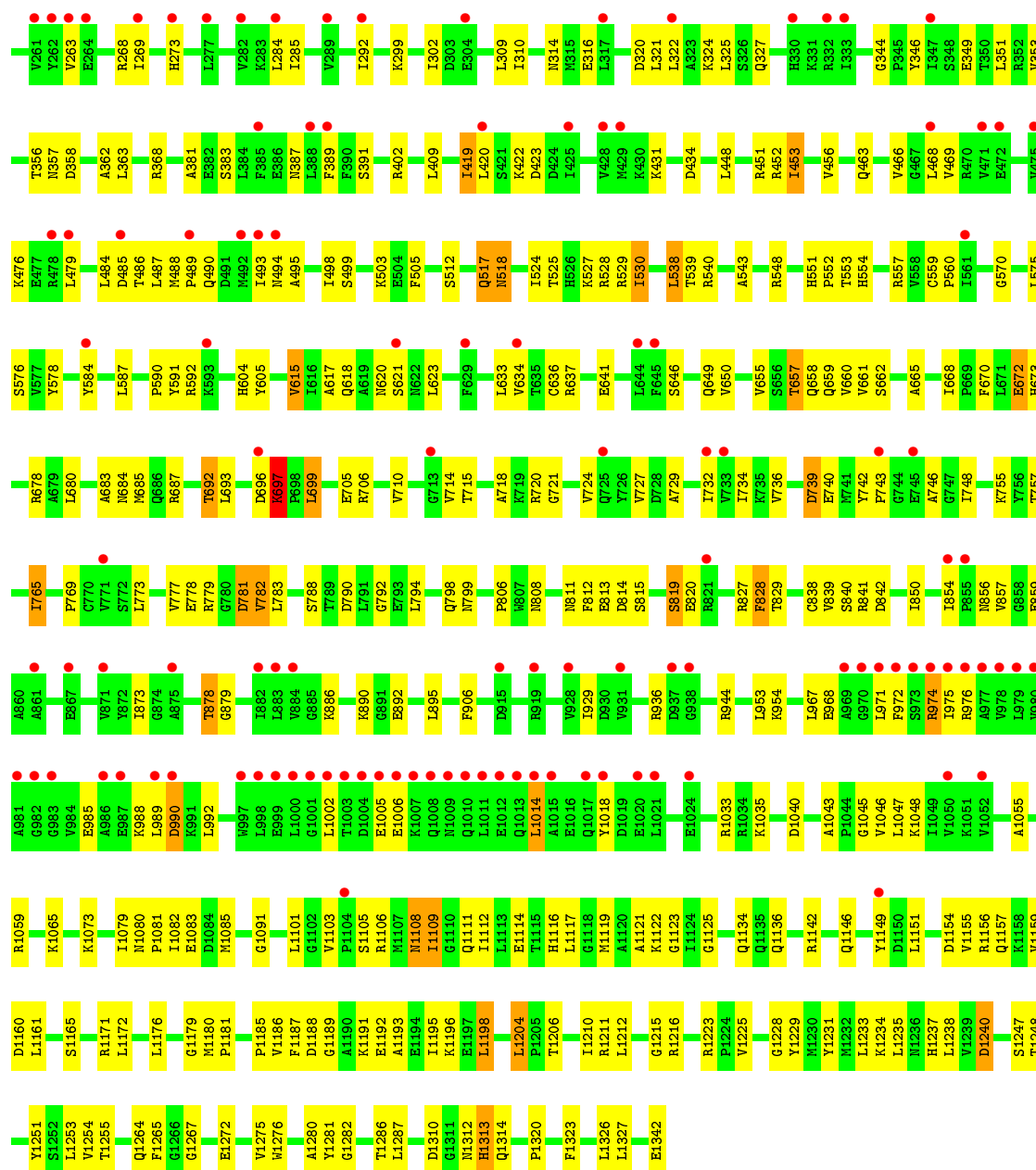


- Molecule 2: DNA-directed RNA polymerase subunit beta

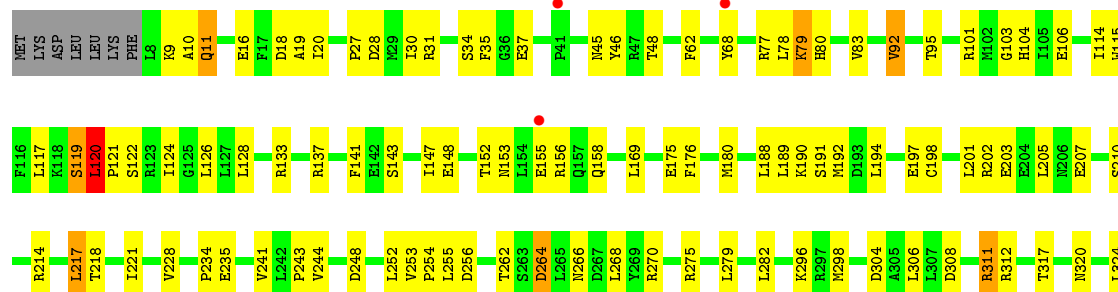


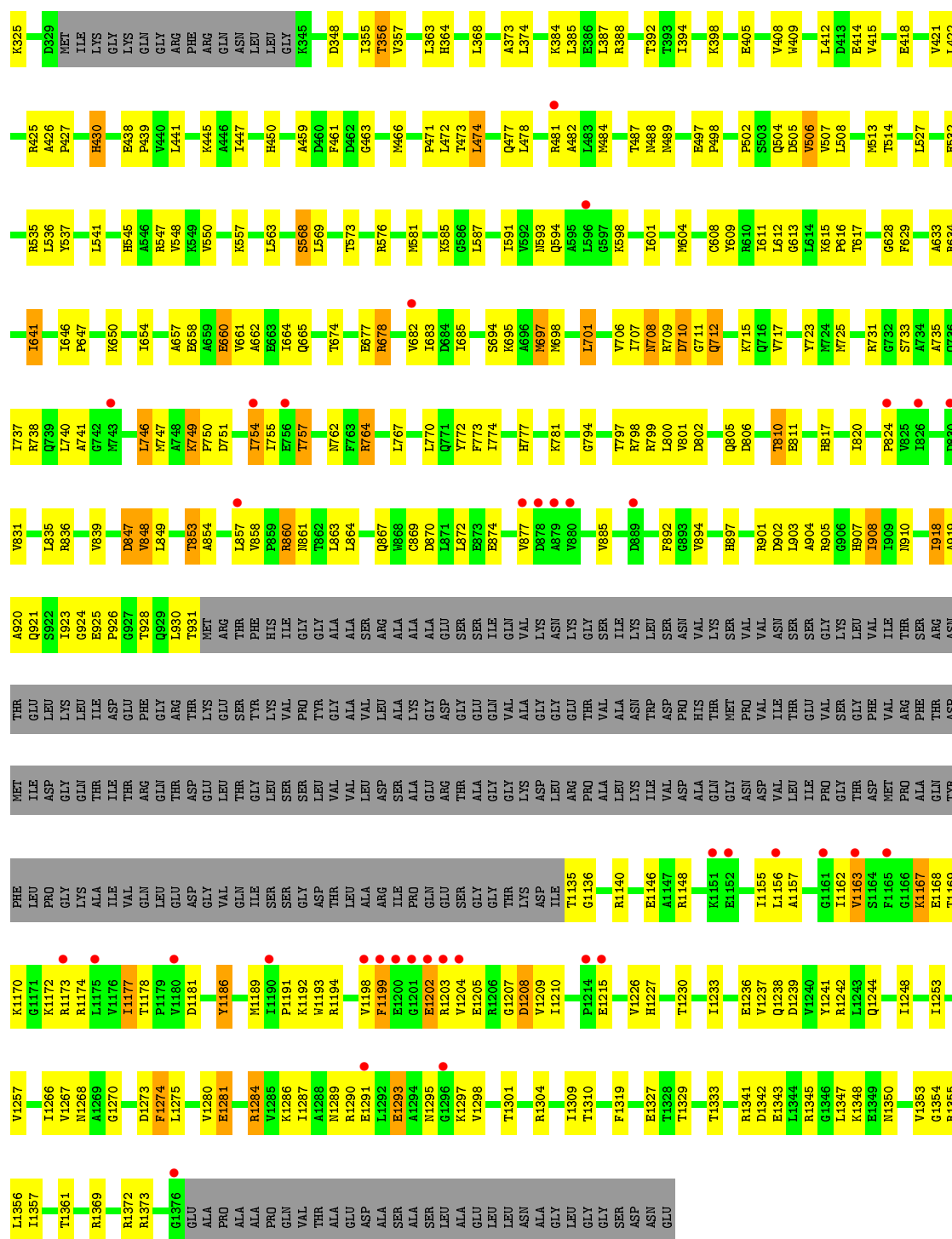






• Molecule 3: DNA-directed RNA polymerase subunit beta'



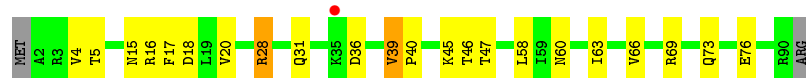
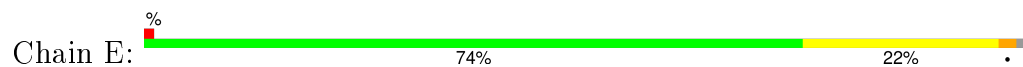


- Molecule 3: DNA-directed RNA polymerase subunit beta'



R1345	G1346	D1239	G1161	L1078	V1011	ILE	L849	V747	A562	L563	P471	Q365	L279	L189	V92
G1347	Y1241	Y1242	V1162	I1080	A1012	GLY	A850	A748	E563	S588	L474	C366	L282	K190	T95
K1348	R1242	L1243	V1163	I1081	G1013	ALA	T853	K750	I664	L569	L569	L368	A287	M192	K96
E1349	Q1244	L1245	K1167	A1082	V1017	ALA	A854	D751	Q668	K570	Q477	A373	I230	D193	V97
M1350		I1246	E1168	A1083	A1018	SER	L857	I754	Q669	I571	L478	E375	I230	R101	R101
V1353		I1248	T1169	A1084	W1020	ARG	P858	I755	S670	T572	P479	L374	I230	M102	M102
G1354		E1254	K1170	ALA	W1020	ALA	P859	E756	R678	T573	A460	E375	I230	G103	G103
I1357		E1254	K1171	ALA	W1020	ALA	R860	I757	R678	G575	R481	P379	V292	C198	H104
T1361		V1257	R1172	I1088	T1024	GLU	L863	N762	A681	R576	M484	K384	K296	L201	L201
D1368		I1266	K1173	I1089	P1026	S948	L864	N763	V682	L579	T487	K385	R202	R202	I114
R1371	A1269	V1267	R1174	I1090	P1027	S949	L864	N764	V683	L580	M488	L385	R203	E203	H115
R1372	G1270	N1268	L1175	P1091	I1028	I950	Q887	R767	D684	M581	M489	E386	E204	L117	F116
R1373		I1266	L1176	M1095	T1029	Q951	L872	L767	I685	I582	P602	R388	D304	L205	L117
A1374		V1267	T1177	Q1096	T1030	Q952	L873	L768	M686	I591	V507	E405	R311	R214	G125
GLY		F1274	E1186	Q1097	S1032	R955	C888	H777	A689	K585	P503	V408	R312	R214	G125
GLU		L1275	E1187	Q1098	G1033	Q956	C889	L770	M698	S586	V511	V409	R317	L217	L126
ALA		V1280	E1188	Q1099	F1034	S957	T880	L771	L701	L587	M513	M408	T218	T218	L127
PRO		E1281	I1189	Q1100	G1035	S958	T881	L772	V706	P588	M514	E414	N320	L128	L128
ALA		R1284	P1191	Q1101	F1036	S959	T882	L773	V707	L587	M515	V415	L324	R133	R133
ALA		V1285	K1192	Q1102	R1036	S960	T883	L774	I708	L588	M516	E418	K325	V228	V228
GLN		K1286	W1193	E1110	F1037	S961	T884	L775	M697	I591	L508	E418	P234	R137	R137
VAL			R1194			S962	T885	L776	M698	I592	L509	E418	P235	F141	F141
THR		N1289	E1195	Q1114	I1106	S963	C888	H777	M698	K585	V511	V421	D329	E142	E142
ALA		R1290	L1196	I1115	I1107	R964	T889	L778	L701	S586	M513	L422	ILE	S143	S143
GLU		E1293	R1197	S1116	Q1108	S965	T890	L779	V707	L587	M514	E418	GLY	V244	V244
ALA		A1294	F1198	S1117	Q1109	S966	T891	L780	I708	P588	M515	E418	LVS	I147	I147
ALA		N1295	F1199	Q1118	E1110	S967	T892	L781	I709	L588	M516	E418	LVS	E148	E148
SER						S968	T893	L782	M697	I591	L508	E418	GLN	E148	E148
ALA		E1297	E1202	D1119	D1051	R978	C895	L779	M697	K585	V511	V421	GLY	L252	L252
LEU		V1298	R1203	L1121	E1052	R979	C896	L780	M697	S586	M513	L422	ARG	N153	N153
ALA			V1204	A1122	L1053	T980	T897	L781	M697	L587	M514	E418	PHE	P254	P254
ALA		T1301	E1205	R1123	G1055	S984	T898	L782	M697	L587	M515	E418	GLN	L255	L255
GLU		Y1302	R1206	I1124	L1056	S985	T899	L783	M697	L587	M516	E418	ASN	E156	E156
LEU		S1303	G1207	P1125	S1057	D986	T900	L784	M697	L587	M517	E418	LEU	R156	R156
LEU		D1304	D1208	Q1126	S1058	E987	T901	L785	M697	L587	M518	E418	LEU	Q157	Q157
ASN			V1209	GLU	L1059	F988	T902	L786	M697	L587	M519	E418	LEU	Q158	Q158
ALA			I1210	SER	V1060	G989	T903	L787	M697	L587	M520	E418	GLY	G258	G258
GLY		I1309	E1215	GLY	V1061	R990	T904	L788	M697	L587	M521	E418	GLY	T161	T161
LEU			E1216	GLY	L1062	T991	T905	L789	M697	L587	M522	E418	ARG	T262	T262
GLY		S1318	F1217	THR	L1063	T992	T906	L790	M697	L587	M523	E418	PHE	S263	S263
GLY		F1319	P1217	LVS	S1064	R993	T907	L791	M697	L587	M524	E418	GLN	D264	D264
SER				ASP	A1065	S994	T908	L792	M697	L587	M525	E418	GLN	L285	L285
ASP		E1327	V1226	ILE			T909	L793	M697	L587	M526	E418	GLN	N286	N286
ASN		T1328	H1227	THR	T1068	V997	T910	L794	M697	L587	M527	E418	GLN	D287	D287
GLU		T1329		GLY	A1069	P998	T911	L795	M697	L587	M528	E418	GLN	E175	E175
				THR	G1070	V999	T912	L796	M697	L587	M529	E418	GLN	F176	F176
				GLY	G1071	G1000	T913	L797	M697	L587	M530	E418	GLN	D177	D177
				LVS	K1072	A1001	T914	L798	M697	L587	M531	E418	GLN	A178	A178
				ASP			T915	L799	M697	L587	M532	E418	GLN	K179	K179
				THR	D1073		T916	L800	M697	L587	M533	E418	GLN	M180	M180
				GLY	L1074		T917	L801	M697	L587	M534	E418	GLN	R275	R275
				THR			T918	L802	M697	L587	M535	E418	GLN		
				GLY			T919	L803	M697	L587	M536	E418	GLN		
				ASP			T920	L804	M697	L587	M537	E418	GLN		
				THR			T921	L805	M697	L587	M538	E418	GLN		
				GLY			T922	L806	M697	L587	M539	E418	GLN		
				ASP			T923	L807	M697	L587	M540	E418	GLN		
				THR			T924	L808	M697	L587	M541	E418	GLN		
				GLY			T925	L809	M697	L587	M542	E418	GLN		
				ASP			T926	L810	M697	L587	M543	E418	GLN		
				THR			T927	L811	M697	L587	M544	E418	GLN		
				GLY			T928	L812	M697	L587	M545	E418	GLN		
				ASP			T929	L813	M697	L587	M546	E418	GLN		
				THR			T930	L814	M697	L587	M547	E418	GLN		
				GLY			T931	L815	M697	L587	M548	E418	GLN		
				ASP			T932	L816	M697	L587	M549	E418	GLN		
				THR			T933	L817	M697	L587	M550	E418	GLN		
				GLY			T934	L818	M697	L587	M551	E418	GLN		
				ASP			T935	L819	M697	L587	M552	E418	GLN		
				THR			T936	L820	M697	L587	M553	E418	GLN		
				GLY			T937	L821	M697	L587	M554	E418	GLN		
				ASP			T938	L822	M697	L587	M555	E418	GLN		
				THR			T939	L823	M697	L587	M556	E418	GLN		
				GLY			T940	L824	M697	L587	M557	E418	GLN		
				ASP			T941	L825	M697	L587	M558	E418	GLN		
				THR			T942	L826	M697	L587	M559	E418	GLN		
				GLY			T943	L827	M697	L587	M560	E418	GLN		
				ASP			T944	L828	M697	L587	M561	E418	GLN		
				THR			T945	L829	M697	L587	M562	E418	GLN		
				GLY			T946	L830	M697	L587	M563	E418	GLN		
				ASP			T947	L831	M697	L587	M564	E418	GLN		
				THR			T948	L832	M697	L587	M565	E418	GLN		
				GLY			T949	L833	M697	L587	M566	E418	GLN		
				ASP			T950	L834	M697	L587	M567	E418	GLN		
				THR			T951	L835	M697	L587	M568	E418	GLN		
				GLY			T952	L836	M697	L587	M569	E418	GLN		
				ASP			T953	L837	M697	L587	M570	E418	GLN		
				THR			T954	L838	M697	L587	M571	E418	GLN		
				GLY			T955	L839	M697	L587	M572	E418	GLN		
				ASP			T956	L840	M697	L587	M573	E418	GLN		
				THR			T957	L841	M697	L587	M574	E418	GLN		
				GLY			T958	L842	M697	L587	M575	E418	GLN		
				ASP			T959	L843	M697	L587	M576	E418	GLN		
				THR			T960	L844	M697	L587	M577	E418	GLN		
				GLY			T961	L845	M697	L587	M578	E418	GLN		
				ASP			T962	L846	M697	L587	M579	E418	GLN		
				THR			T963	L847	M697	L587	M580	E418	GLN		
				GLY			T964	L848	M697	L587	M581	E418	GLN		
				ASP			T965	L849	M697	L587	M582	E418	GLN		
				THR			T966	L850	M697	L587	M583	E418	GLN		
				GLY			T967	L851	M697	L587	M584	E418	GLN		
				ASP			T968	L852	M697	L587	M585	E418	GLN		
				THR			T969	L853	M697	L587	M586	E418	GLN		
				GLY			T970	L854	M697	L587	M587	E418	GLN		
				ASP			T971	L855	M697	L587	M588	E418	GLN		
				THR			T972	L856	M697	L587	M589	E418	GLN		
				GLY			T973	L857	M697	L587	M590	E418	GLN		
				ASP			T974	L858	M697	L587	M591	E418	GLN		

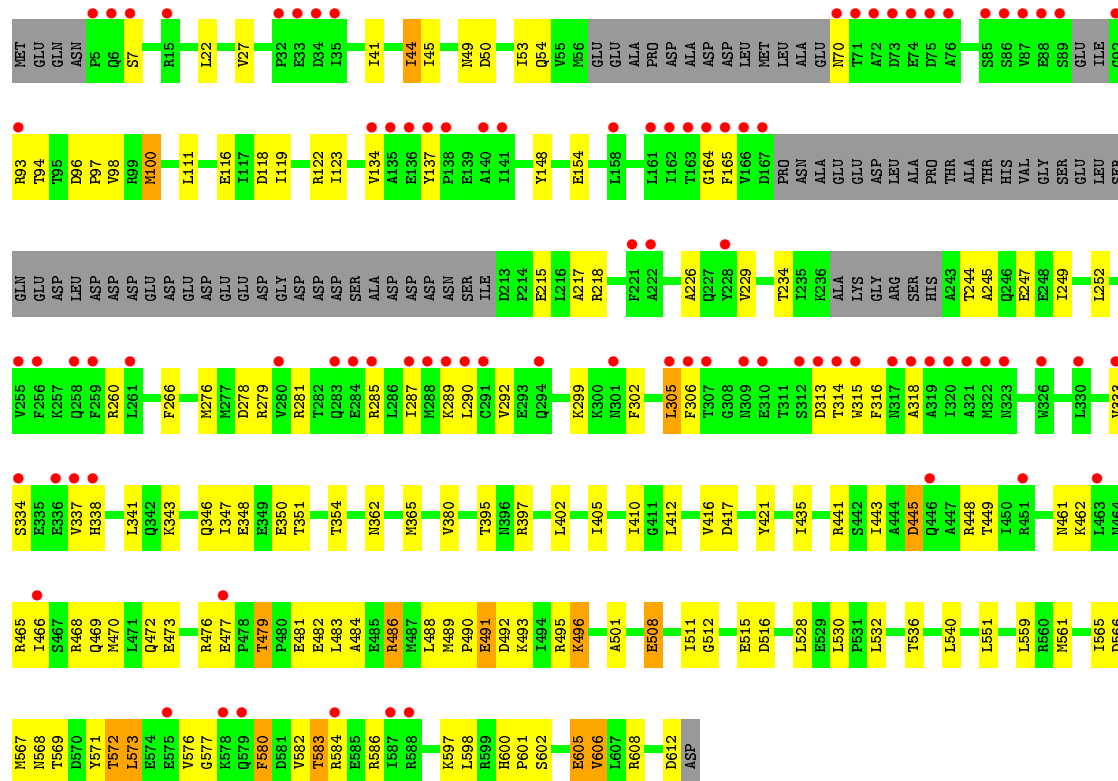
- Molecule 4: DNA-directed RNA polymerase subunit omega



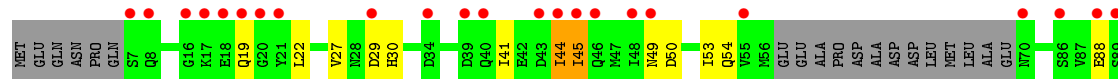
- Molecule 4: DNA-directed RNA polymerase subunit omega



- 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, $\alpha$ , $\beta$ , $\gamma$	185.58Å 204.10Å 308.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.20 – 3.99 49.79 – 3.99	Depositor EDS
% Data completeness (in resolution range)	93.1 (49.20-3.99) 92.3 (49.79-3.99)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.00 (at 4.00Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.250 , 0.283 0.287 , 0.304	Depositor DCC
$R_{free}$ test set	1978 reflections (2.19%)	DCC
Wilson B-factor (Å <sup>2</sup> )	156.4	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 88.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 98970 reflections	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	57539	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.57% 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, MG, 4OE

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	0.39	1/2358 (0.0%)	0.67	2/3197 (0.1%)
1	B	0.39	0/1687	0.73	1/2286 (0.0%)
1	G	0.33	0/1751	0.66	0/2373
1	H	0.37	0/1681	0.70	2/2278 (0.1%)
2	C	0.28	0/10739	0.49	0/14489
2	I	0.28	1/10735 (0.0%)	0.47	0/14484
3	D	0.29	0/9130	0.50	1/12325 (0.0%)
3	J	0.27	0/10409	0.48	1/14059 (0.0%)
4	E	0.29	0/693	0.49	0/935
4	K	0.26	0/629	0.48	0/847
5	F	0.31	2/4254 (0.0%)	0.51	1/5731 (0.0%)
5	L	0.28	0/4246	0.49	0/5720
All	All	0.30	4/58312 (0.0%)	0.52	8/78724 (0.0%)

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	D	0	1
3	J	0	1
All	All	0	2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	954	LYS	CD-CE	-6.99	1.33	1.51
5	F	605	GLU	CD-OE2	5.91	1.32	1.25
1	A	231	PHE	CD2-CE2	-5.76	1.27	1.39
5	F	605	GLU	CD-OE1	5.01	1.31	1.25



The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	250	ASP	CB-CG-OD1	-9.38	109.86	118.30
5	F	605	GLU	OE1-CD-OE2	9.01	134.11	123.30
1	B	29	GLU	C-N-CD	8.36	145.96	128.40
1	H	13	LEU	CA-CB-CG	6.42	130.07	115.30
3	D	120	LEU	N-CA-C	5.85	126.80	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	120	LEU	Peptide
3	J	120	LEU	Peptide

## 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	2328	0	2380	72	0
1	B	1667	0	1692	63	0
1	G	1730	0	1756	69	0
1	H	1662	0	1687	67	0
2	C	10570	0	10582	277	0
2	I	10566	0	10576	262	0
3	D	8992	0	9180	271	0
3	J	10254	0	10461	313	0
4	E	691	0	695	14	0
4	K	627	0	634	19	0
5	F	4204	0	4106	95	0
5	L	4196	0	4103	103	0
6	C	23	0	9	5	0
6	I	23	0	9	4	0
7	D	1	0	0	0	0
7	J	1	0	0	0	0
8	D	2	0	0	0	0
8	J	2	0	0	0	0
All	All	57539	0	57870	1457	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 13.

The worst 5 of 1457 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)
3:D:660:GLU:HB3	3:D:685:ILE:HD12	1.41	1.01
3:J:660:GLU:HB3	3:J:685:ILE:HD12	1.46	0.97
1:H:29:GLU:HB3	1:H:30:PRO:HD3	1.48	0.95
3:D:1280:VAL:HG21	3:D:1304:ARG:HE	1.39	0.88
2:C:1105:SER:HB2	3:D:731:ARG:HG2	1.56	0.87

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	298/335 (89%)	269 (90%)	20 (7%)	9 (3%)	5	45
1	B	212/335 (63%)	191 (90%)	18 (8%)	3 (1%)	14	59
1	G	222/335 (66%)	200 (90%)	16 (7%)	6 (3%)	6	47
1	H	212/335 (63%)	192 (91%)	17 (8%)	3 (1%)	14	59
2	C	1338/1342 (100%)	1236 (92%)	95 (7%)	7 (0%)	34	76
2	I	1338/1342 (100%)	1234 (92%)	98 (7%)	6 (0%)	39	79
3	D	1145/1407 (81%)	1050 (92%)	90 (8%)	5 (0%)	39	79
3	J	1311/1407 (93%)	1196 (91%)	112 (8%)	3 (0%)	52	86
4	E	87/91 (96%)	79 (91%)	8 (9%)	0	100	100
4	K	77/91 (85%)	73 (95%)	4 (5%)	0	100	100
5	F	532/613 (87%)	481 (90%)	50 (9%)	1 (0%)	52	86
5	L	529/613 (86%)	481 (91%)	48 (9%)	0	100	100
All	All	7301/8246 (88%)	6682 (92%)	576 (8%)	43 (1%)	30	73

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	193	GLU
1	A	319	GLU
1	A	320	ASN
1	B	29	GLU
2	C	237	LEU

### 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	257/292 (88%)	243 (95%)	14 (5%)	27	67
1	B	184/292 (63%)	166 (90%)	18 (10%)	10	43
1	G	191/292 (65%)	181 (95%)	10 (5%)	29	68
1	H	183/292 (63%)	168 (92%)	15 (8%)	14	51
2	C	1155/1157 (100%)	1064 (92%)	91 (8%)	15	54
2	I	1154/1157 (100%)	1062 (92%)	92 (8%)	15	53
3	D	964/1168 (82%)	877 (91%)	87 (9%)	12	47
3	J	1106/1168 (95%)	1012 (92%)	94 (8%)	13	51
4	E	72/75 (96%)	66 (92%)	6 (8%)	14	51
4	K	67/75 (89%)	60 (90%)	7 (10%)	9	40
5	F	426/540 (79%)	392 (92%)	34 (8%)	15	53
5	L	428/540 (79%)	393 (92%)	35 (8%)	14	51
All	All	6187/7048 (88%)	5684 (92%)	503 (8%)	15	52

5 of 503 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	F	395	THR
2	I	91	THR
4	K	36	ASP
5	F	488	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	133	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	560	ASN
3	D	1367	GLN
3	J	777	HIS
3	D	897	HIS
5	F	345	GLN

### 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 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	4OE	C	2001	-	24,25,25	4.18	6 (25%)	30,37,37	1.39	5 (16%)
6	4OE	I	2001	-	24,25,25	4.16	6 (25%)	30,37,37	1.39	5 (16%)

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	4OE	C	2001	-	-	0/14/14/14	0/3/3/3
6	4OE	I	2001	-	-	0/14/14/14	0/3/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	2001	4OE	CAR-CAV	-14.35	1.33	1.49
6	I	2001	4OE	CAR-CAV	-14.25	1.33	1.49
6	C	2001	4OE	CAT-CAS	-9.61	1.33	1.49
6	I	2001	4OE	CAT-CAS	-9.55	1.33	1.49
6	C	2001	4OE	NAN-NAO	-6.73	1.23	1.37

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	C	2001	4OE	CAW-CAU-CAQ	-2.39	117.29	121.98
6	I	2001	4OE	CAW-CAU-CAQ	-2.38	117.30	121.98
6	I	2001	4OE	FAD-CAW-CAU	-2.36	108.65	112.67
6	C	2001	4OE	FAD-CAW-CAU	-2.35	108.67	112.67
6	C	2001	4OE	CAH-CAQ-CAU	-2.22	120.75	123.80

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	2001	4OE	5	0
6	I	2001	4OE	4	0

## 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	302/335 (90%)	0.38	25 (8%) 14 10	29, 105, 230, 394	0
1	B	216/335 (64%)	0.49	19 (8%) 12 9	42, 129, 228, 299	0
1	G	224/335 (66%)	0.26	16 (7%) 19 13	57, 116, 198, 279	0
1	H	216/335 (64%)	0.72	36 (16%) 2 3	57, 136, 228, 269	0
2	C	1340/1342 (99%)	0.42	120 (8%) 12 8	13, 89, 219, 379	0
2	I	1340/1342 (99%)	0.54	164 (12%) 5 5	14, 112, 225, 339	0
3	D	1151/1407 (81%)	0.15	40 (3%) 48 37	15, 70, 173, 283	0
3	J	1319/1407 (93%)	0.60	149 (11%) 7 6	19, 94, 230, 319	0
4	E	89/91 (97%)	0.10	1 (1%) 82 75	27, 78, 123, 186	0
4	K	79/91 (86%)	0.44	3 (3%) 44 34	60, 124, 203, 243	0
5	F	542/613 (88%)	0.78	90 (16%) 2 3	27, 158, 265, 368	0
5	L	539/613 (87%)	0.69	86 (15%) 3 3	41, 153, 254, 312	0
All	All	7357/8246 (89%)	0.48	749 (10%) 9 7	13, 105, 230, 394	0

The worst 5 of 749 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	J	1054	THR	15.9
5	F	70	ASN	13.7
5	F	167	ASP	13.1
5	F	259	PHE	13.0
5	F	89	SER	12.1

### 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 [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	4OE	I	2001	23/23	0.73	0.96	4.70	100,113,125,126	0
6	4OE	C	2001	23/23	0.91	0.67	2.94	48,74,93,102	0
8	ZN	J	1503	1/1	0.98	0.25	-0.01	54,54,54,54	0
8	ZN	D	1503	1/1	0.99	0.23	-0.61	33,33,33,33	0
8	ZN	D	1502	1/1	0.96	0.12	-0.76	84,84,84,84	0
8	ZN	J	1502	1/1	0.91	0.16	-0.82	94,94,94,94	0
7	MG	J	1501	1/1	0.83	0.71	-	69,69,69,69	0
7	MG	D	1501	1/1	0.80	0.53	-	28,28,28,28	0

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

There are no such residues in this entry.