



wwPDB X-ray Structure Validation Summary Report ⓘ

May 29, 2020 – 01:21 am BST

PDB ID : 2O5I
Title : Crystal structure of the T. thermophilus RNA polymerase elongation complex
Authors : Vassilyev, D.G.; Tahirov, T.H.; Vassilyeva, M.N.
Deposited on : 2006-12-06
Resolution : 2.50 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

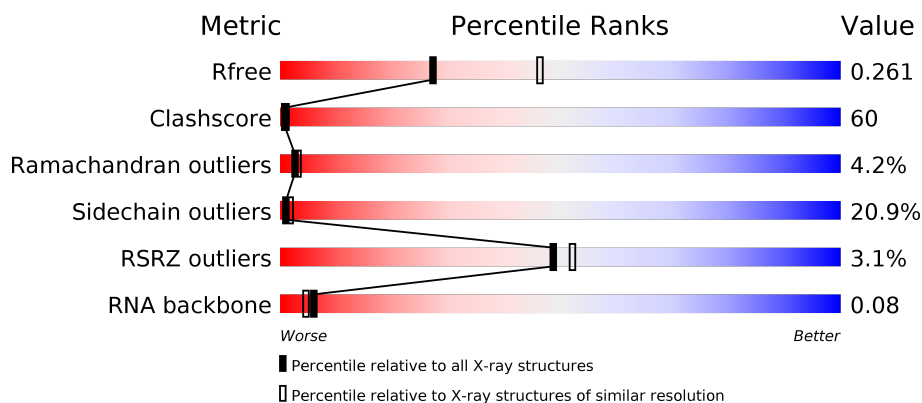
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)
RNA backbone	3102	1008 (2.84-2.16)

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	G	23	
1	X	23	
2	H	16	
2	Y	16	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	I	14	<div><div></div><div>36%50%7%7%</div></div>
3	Z	14	<div><div></div><div>29%57%7%7%</div></div>
4	A	315	<div><div></div><div>%19%45%9%27%</div></div>
4	B	315	<div><div></div><div>3%22%42%8%27%</div></div>
4	K	315	<div><div></div><div>2%22%42%9%27%</div></div>
4	L	315	<div><div></div><div>4%26%37%9%27%</div></div>
5	C	1119	<div><div></div><div>4%22%57%19%.</div></div>
5	M	1119	<div><div></div><div>3%25%59%15%.</div></div>
6	D	1524	<div><div></div><div>2%20%51%13%.15%</div></div>
6	N	1524	<div><div></div><div>2%24%48%12%.15%</div></div>
7	E	99	<div><div></div><div>3%24%54%15%..</div></div>
7	O	99	<div><div></div><div>5%25%49%19%..</div></div>

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 52719 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 DNA chain called 5'-D(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	23	Total	C	N	O	P	0	0	0
			467	220	80	144	23			
1	X	23	Total	C	N	O	P	0	0	0
			467	220	80	144	23			

- Molecule 2 is a RNA chain called 5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	16	Total	C	N	O	P	0	0	0
			347	153	64	114	16			
2	Y	16	Total	C	N	O	P	0	0	0
			347	153	64	114	16			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	13	Total	C	N	O	P	0	0	0
			270	126	57	74	13			
3	Z	13	Total	C	N	O	P	0	0	0
			270	126	57	74	13			

- Molecule 4 is a protein called DNA-directed RNA polymerase alpha chain.

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

Continued on next page...

Continued from previous page...

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

- Molecule 5 is a protein called DNA-directed RNA polymerase beta chain.

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

- Molecule 6 is a protein called DNA-directed RNA polymerase beta' chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	D	1303	Total	C	N	O	S	0	0	0
			10280	6508	1821	1919	32			
6	N	1303	Total	C	N	O	S	0	0	0
			10280	6508	1821	1919	32			

- Molecule 7 is a protein called DNA-directed RNA polymerase omega chain.

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

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

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

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	1	Total	Mg	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	N	1	Total 1	Mg 1	0	0

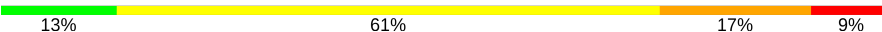
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	G	22	Total 22	O 22	0	0
10	H	18	Total 18	O 18	0	0
10	I	36	Total 36	O 36	0	0
10	X	25	Total 25	O 25	0	0
10	Y	16	Total 16	O 16	0	0
10	Z	16	Total 16	O 16	0	0
10	A	144	Total 144	O 144	0	0
10	B	159	Total 159	O 159	0	0
10	C	658	Total 658	O 658	0	0
10	D	760	Total 760	O 760	0	0
10	E	70	Total 70	O 70	0	0
10	K	132	Total 132	O 132	0	0
10	L	121	Total 121	O 121	0	0
10	M	575	Total 575	O 575	0	0
10	N	750	Total 750	O 750	0	0
10	O	61	Total 61	O 61	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: 5'-D(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*CP*G)-3'

Chain G: 



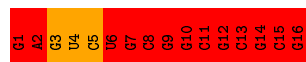
- Molecule 1: 5'-D(P*CP*CP*CP*TP*GP*TP*CP*TP*GP*GP*CP*GP*TP*TP*CP*GP*CP*GP*CP*GP*CP*CP*G)-3'

Chain X: 



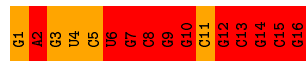
- Molecule 2: 5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3',

Chain H: 



- Molecule 2: 5'-R(P*GP*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP*G)-3',

Chain Y: 



- Molecule 3: 5'-D(*AP*AP*CP*GP*CP*CP*AP*GP*AP*CP*AP*GP*GP*G)-3'

Chain I: 




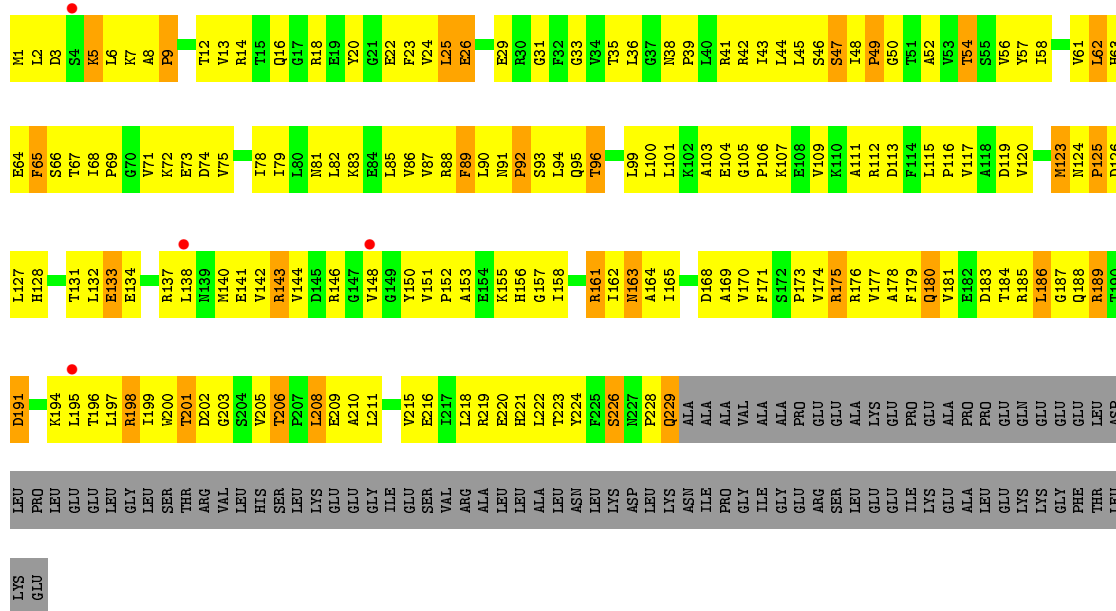
- Molecule 3: 5'-D(*AP*AP*CP*GP*CP*CP*AP*GP*AP*CP*AP*GP*GP*G)-3'

Chain Z: 




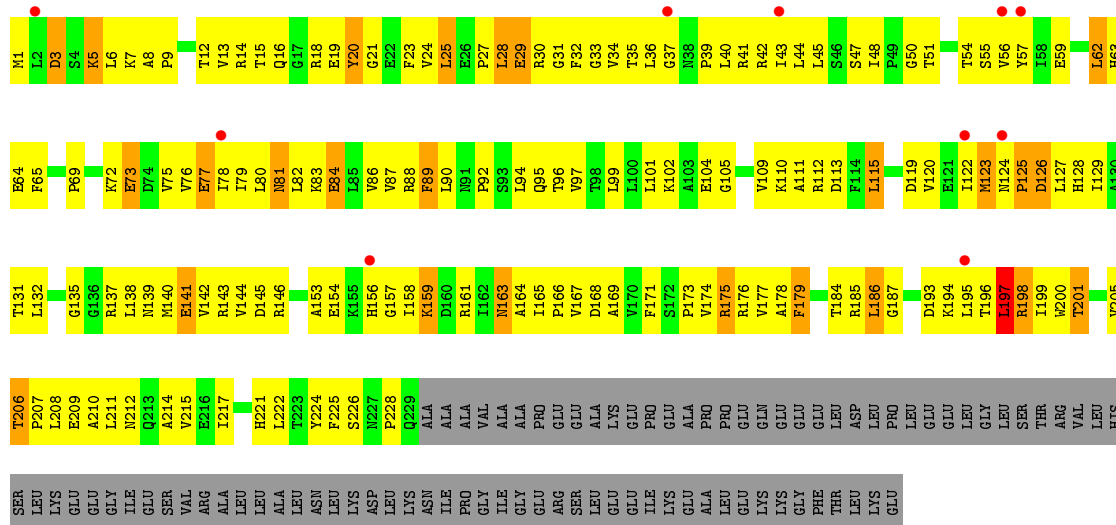
- Molecule 4: DNA-directed RNA polymerase alpha chain

Chain A: 

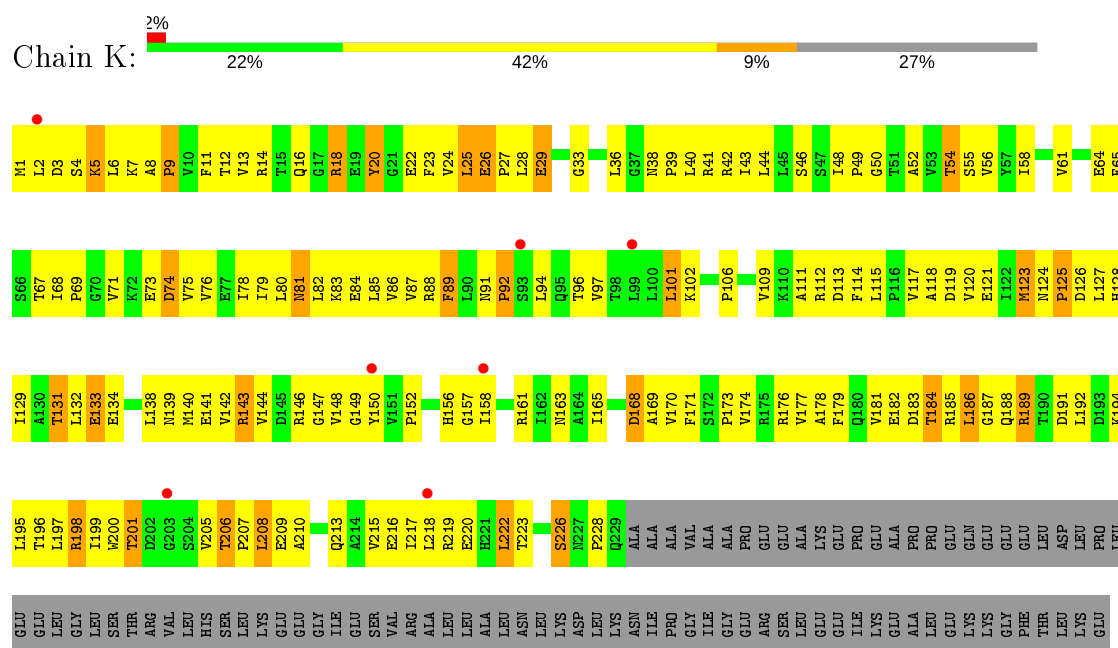


- Molecule 4: DNA-directed RNA polymerase alpha chain

Chain B: 

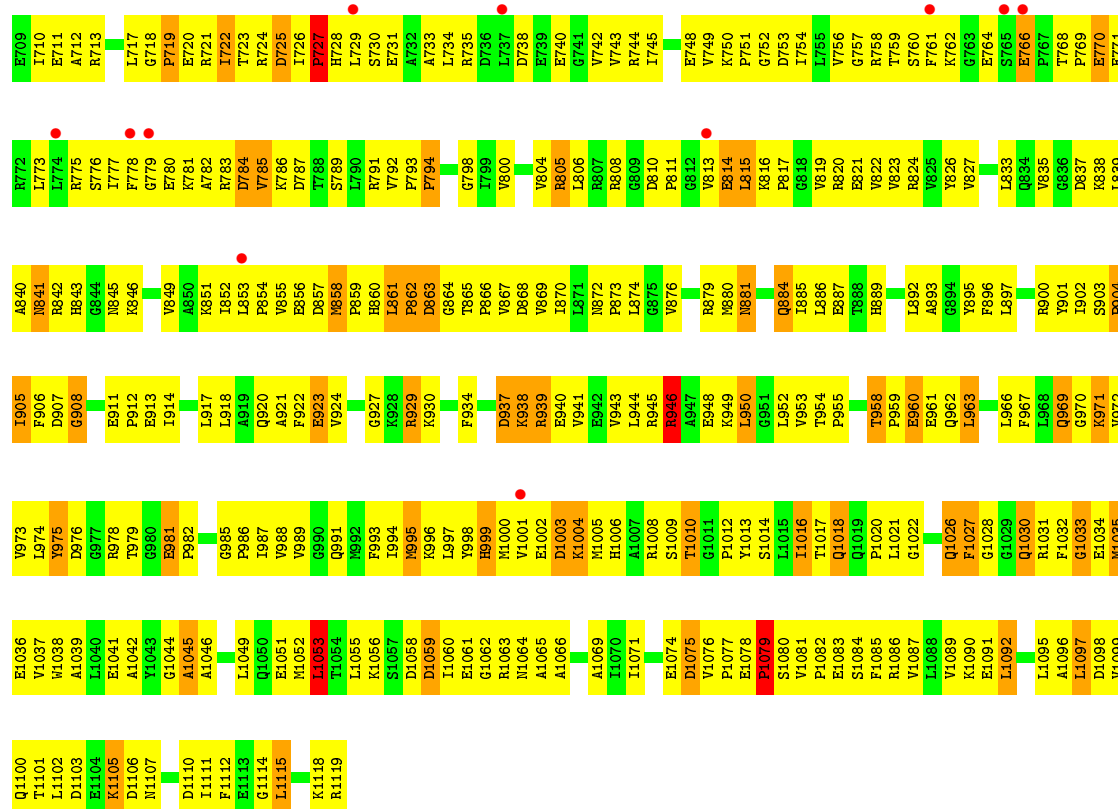


- Molecule 4: DNA-directed RNA polymerase alpha chain

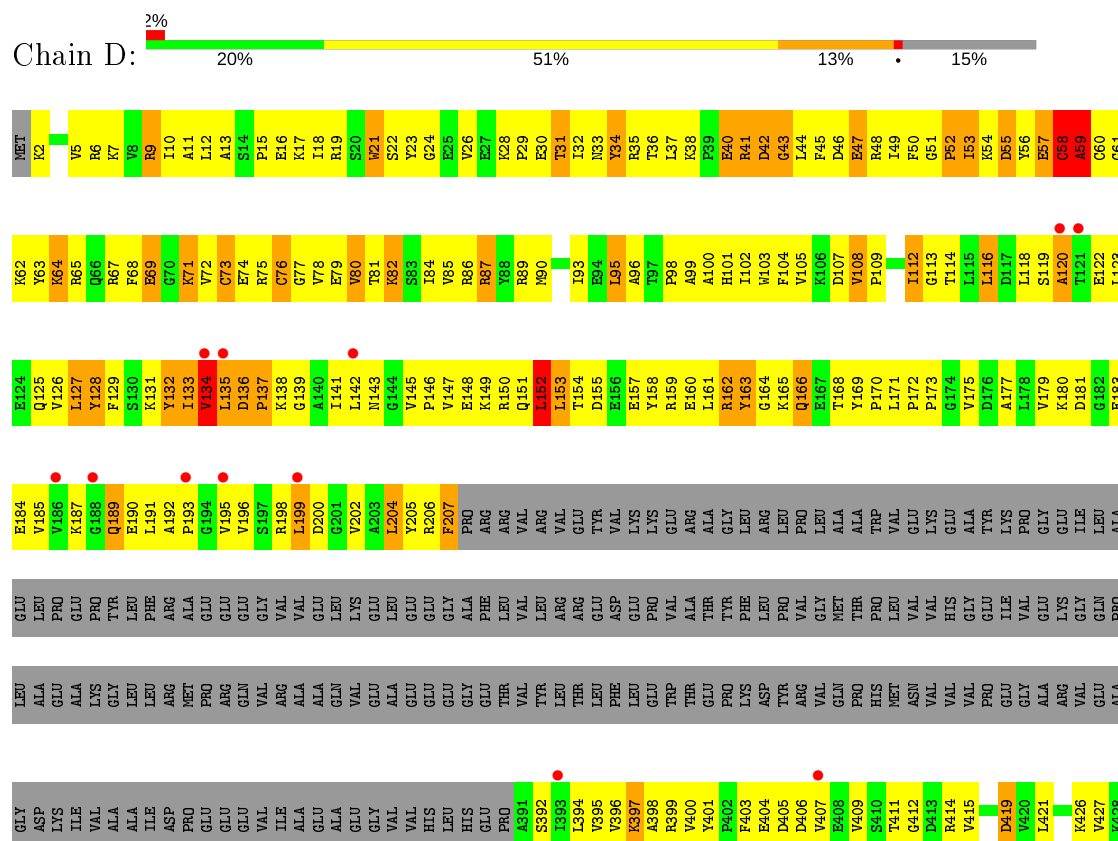


L944	R945	R879	P817	L755	E685	Y623	L559	A497	H434	K371	R308	P248	V186	D124
R946	R880	R880	G818	V756	D686	P624	S562	Q496	Y435	L372	Y309	K249	N187	G125
L882	R881	R881	V819	G757	A687	L625	S563	A499	G436	V373	F311	R250	K188	S126
R948	R882	R882	R758	R758	I688	R626	I563	I500	R437	N374	F311	K251	R189	F127
K949	R883	R883	E821	T759	V689	R627	I564	T501	I438	S375	A312	K252	K190	I128
L950	R884	R884	V822	S760	I690	G628	Q565	P502	C439	R376	L313	A253	F191	I129
G951	L885	L885	V823	F761	S691	Y629	T566	L503	P440	P377	T314	V254	P192	N130
L952	R824	R824	R824	K762	E692	R630	T567	L506	V441	L378	A315	A255	P192	G131
L953	R825	R825	R825	G763	E693	N631	A568	R507	T443	E379	P318	Y256	V194	A132
T954	V826	V826	V826	E764	S631	N632	I569	R507	A442	A380	G319	Y257	L195	D133
P955	V827	V827	V827	S765	L695	Q633	P570	R507	T443	V357	H320	Y258	L196	R134
G956	A828	A828	A828	E766	K696	G634	L571	A510	A447	I382	H320	G259	L197	V135
K957	Q829	Q829	Q829	T767	R697	T635	L572	A511	A448	R383	G321	L260	L197	I136
T958	R832	R832	R832	T768	D698	A636	R573	A512	I449	E384	V322	I261	V199	V137
P959	E770	E770	E770	F769	F699	L637	A574	A510	A447	I382	H320	L260	L197	I136
R960	L833	L833	L833	E770	Y700	D638	Q575	A511	A448	R383	G321	L261	V199	V137
E961	Q834	Q834	Q834	E771	T701	Q639	A576	A512	I449	E384	V322	I261	V199	V137
G961	V835	V835	V835	R772	S702	R640	P577	A515	G450	F386	D323	A262	L200	D133
Q962	G836	G836	G836	L773	I703	P641	V578	A516	L451	F386	D323	A262	L200	R134
L963	D837	D837	D837	L774	H704	R642	V579	A517	I449	E384	V322	I261	V199	V137
R963	R838	R838	R838	R775	H705	V643	I580	K518	A456	S392	H331	D267	V199	V137
E965	L839	L839	L839	S776	E706	V644	G582	E520	A457	Q393	R332	L269	L200	S138
L966	A840	A840	A840	I777	R707	V645	G582	P521	Y458	F394	I333	G270	A208	V146
F967	N841	N841	N841	F778	R707	G646	L583	V522	A459	K395	R334	E271	R209	Y147
L968	R842	R842	R842	G779	I710	Q647	E584	I523	R460	T335	T335	A272	R209	F148
Q969	E780	E780	E780	E780	R719	R648	E585	V524	Y461	N399	V336	G273	L211	T149
G970	K811	K811	K811	K811	R719	L655	L592	F531	A468	P400	G337	E279	L217	P150
K971	L852	L852	L852	L852	D714	A656	Y596	F532	R468	L401	E338	L281	V218	D151
V972	R788	R788	R788	T788	R720	D657	E597	M532	T469	K407	F344	G220	Q219	I158
Q973	L790	L790	L790	G658	I722	G657	E598	D533	R408	R345	G282	G282	Q219	I159
R974	R784	R784	R784	G659	T723	P659	E599	V534	R409	R409	G347	L283	L221	A160
E975	V785	V785	V785	A660	R724	A660	D600	S535	Y471	V346	G347	L283	L221	S161
D976	K861	K861	K861	L654	G718	L654	L592	E536	R472	I410	L348	L283	L221	S161
G977	R787	R787	R787	L655	R719	L655	L592	F531	A468	P400	G337	E279	L217	P150
R978	L853	L853	L853	A656	E720	A656	Y596	F532	R468	L401	E338	L281	V218	D151
T979	P854	P854	P854	G658	I722	G657	E598	D533	R408	R345	G282	G282	Q219	I158
G980	R791	R791	R791	P659	T723	P659	E599	V534	R409	R409	G347	L283	L221	A160
A921	V792	V792	V792	A660	R724	A660	D600	S535	Y471	V346	G347	L283	L221	S161
P982	D857	D857	D857	L654	G718	L654	L592	E536	R472	I410	L348	L283	L221	S161
E983	L858	L858	L858	A656	E720	A656	Y596	F532	R468	L401	E338	L281	V218	D151
V984	P859	P859	P859	G658	I722	G657	E598	D533	R408	R345	G282	G282	Q219	I158
G985	H860	H860	H860	P659	T723	P659	E599	V534	R409	R409	G347	L283	L221	A160
L986	P862	P862	P862	L654	G718	L654	L592	E536	R472	I410	L348	L283	L221	S161
R987	D863	D863	D863	A656	E720	A656	Y596	F532	R468	L401	E338	L281	V218	D151
V988	P864	P864	P864	G658	I722	G657	E598	D533	R408	R345	G282	G282	Q219	I158
G990	R791	R791	R791	P659	T723	P659	E599	V534	R409	R409	G347	L283	L221	A160
Q991	V792	V792	V792	A660	R724	A660	D600	S535	Y471	V346	G347	L283	L221	S161
L992	L861	L861	L861	L654	G718	L654	L592	E536	R472	I410	L348	L283	L221	S161
R993	D863	D863	D863	A656	E720	A656	Y596	F532	R468	L401	E338	L281	V218	D151
V994	P864	P864	P864	G658	I722	G657	E598	D533	R408	R345	G282	G282	Q219	I158
G995	H860	H860	H860	P659	T723	P659	E599	V534	R409	R409	G347	L283	L221	A160
L996	P862	P862	P862	L654	G718	L654	L592	E536	R472	I410	L348	L283	L221	S161
R997	D863	D863	D863	A656	E720	A656	Y596	F532	R468	L401	E338	L281	V218	D151
V998	P864	P864	P864	G658	I722	G657	E598	D533	R408	R345	G282	G282	Q219	I158
G999	R791	R791	R791	P659	T723	P659	E599	V534	R409	R409	G347	L283	L221	A160
M1000	E940	E940	E940	L874	R744	D680	G618	D564	T492	V430	S366	F303	L242	V181
V1001	R876	R876	R876	G875	I745	G881	G618	D564	T492	V430	S366	F303	L242	V181
E1002	P877	P877	P877	L815	D753	N682	V620	I555	Y494	H431	T368	P305	R243	V182
D1003	S878	S878	S878	F684	I754	F684	E622	A558	I496	T433	A370	T306	G245	S183

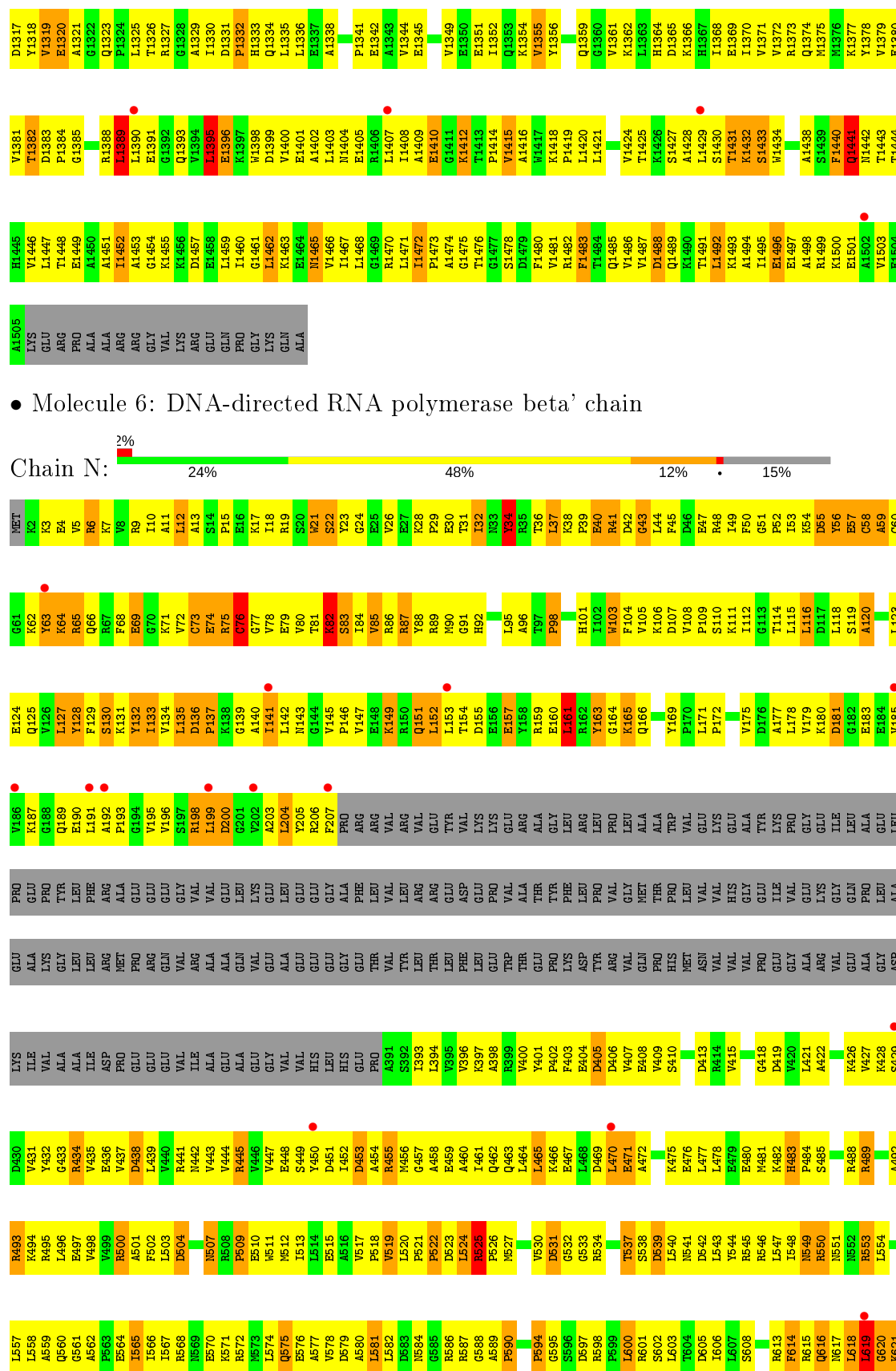


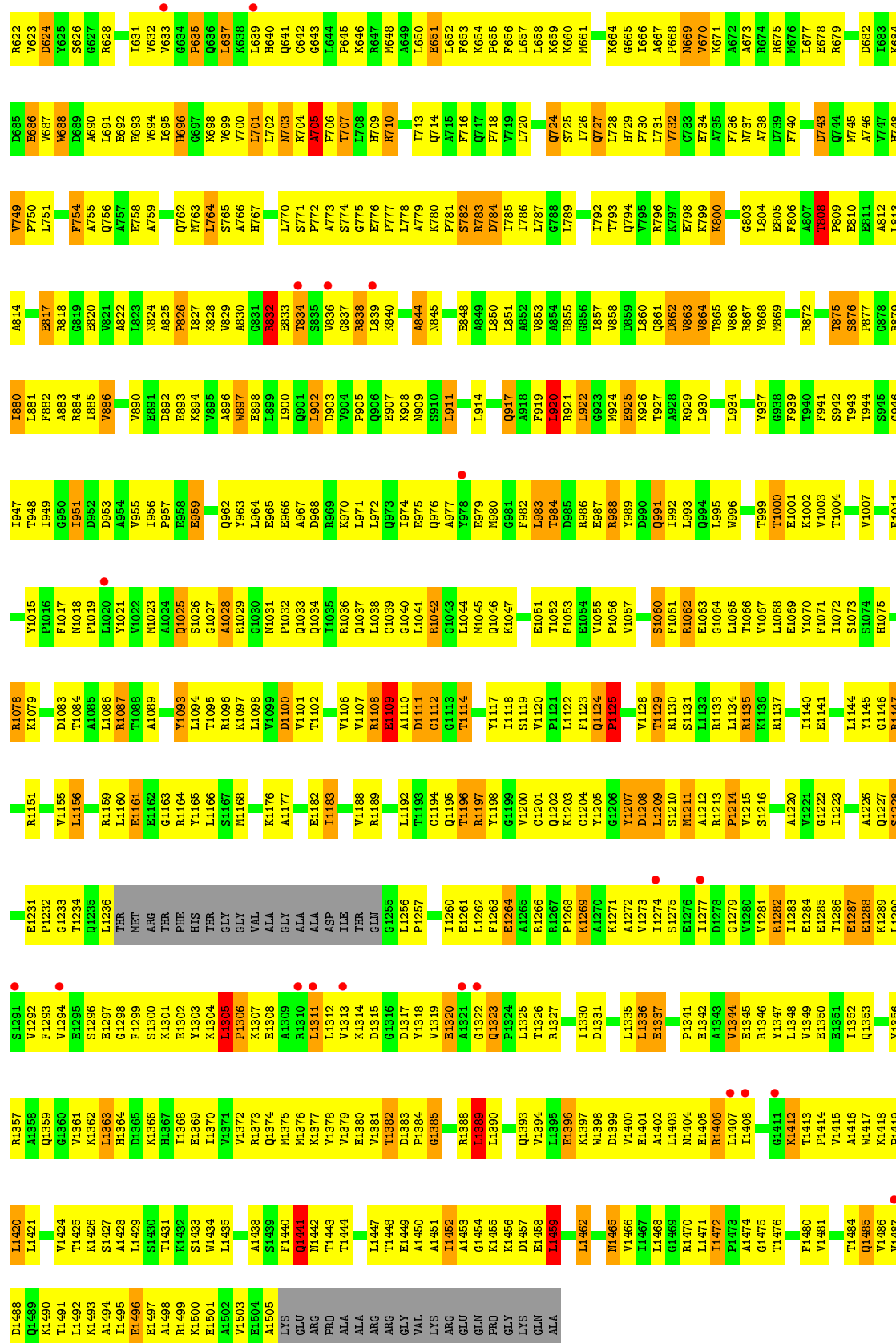


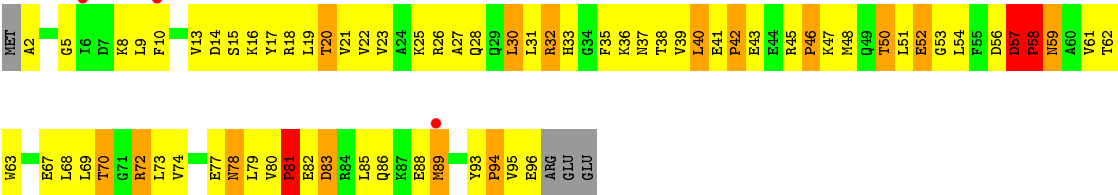
● Molecule 6: DNA-directed RNA polymerase beta' chain



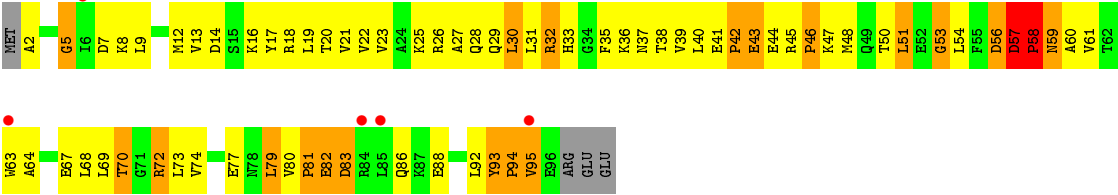
P1257	T1196	T1129	T1066	L934	F806	D743	D682	G620	H551	R489	S429
R1258	R1197	R1130	V1067	R935	A807	Q744	I683	R621	H552	A489	D430
V1259	Y1198	S1131	L1088	Y936	T808	M745	K684	R622	R553	A490	V431
L1260	G1199	E1001	E1069	Y937	P809	A746	R685	V623	L554	R491	Y432
E1261	V1200	K1002	K1003	Y938	E810	S876	E686	D624	K555	R492	K433
L1262	V1201	V1003	V1004	Y939	E811	H748	E687	G625	K556	R493	R434
F1263	Q1202	R1135	I1072	T940	A812	V749	H688	S626	L557	R494	V435
E1264	K1203	R1136	I1073	F941	L813	P750	D689	G627	L558	R495	V436
A1265	C1204	R1137	S1074	S942	A814	L751	A690	R628	A559	L496	V437
R1266	Y1205	A1138	H1075	T943	A815	S752	E692	S629	G561	E497	D438
R1267	G1206	D1139	H1075	T944	R816	S753	E693	V630	G561	V499	L439
P1268	V1207	T1140	F1011	S945	E817	F754	F694	I631	A562	R500	L440
K1269	D1208	E1141	Y1015	S946	R818	A755	V694	V632	I635	A501	R441
A1270	L1209	G1080	P1016	Y947	G819	Q756	I695	V633	I665	F502	R442
K1271	S1210	G1081	P1017	T948	E820	A757	H696	G634	I566	R503	V443
A1272	M1211	A1082	E888	T949	R821	E758	G697	P635	I567	L503	V444
I1273	A1212	D1083	G950	T950	A822	E759	K698	Q636	R568	D504	R445
I1274	R1213	R1147	Y890	Y951	L823	T761	K699	L637	R569	S505	R446
S1275	V1214	V1148	E891	D952	R824	Q762	V700	V638	E570	G506	V446
E1276	V1215	L1149	D953	D953	A825	M763	L701	L639	K571	R507	V447
I1277	V1216	A1150	A954	A954	P826	L764	L702	L640	R572	R508	E448
D1278	I1217	R1151	V955	V955	R828	S765	N703	Q641	R573	P509	S449
G1279	A1220	E1152	G1027	Y956	K828	A766	R704	C642	W511	E510	Y450
V1280	V1281	V1153	A1028	P957	W829	H767	A773	G643	W512	D451	D451
I1281	G1221	G1092	R1029	F958	A830	T768	P706	L644	A577	D453	I452
R1282	G1222	Y1093	G1030	E959	G831	L769	T707	P645	I513	A454	D453
I1283	I1223	T1095	M1031	Y963	R832	L770	L708	R646	I514	R455	A454
E1284	V1224	R1096	P1032	R833	E833	S771	H709	A580	E515	R456	R456
E1285	A1225	L1160	Q1033	Q901	T834	P772	R710	W648	A516	G457	G457
T1286	A1226	K1087	Q1034	L972	S835	A773	L711	A649	P517	E518	A458
G1227	G1227	L1098	I1035	Q973	W836	S774	G712	L650	V519	R459	R459
E1288	S1228	L1098	R1036	D968	G837	G775	I713	E651	L520	A460	A460
I1289	I1229	V1101	Q1037	R969	R838	E776	Q714	L652	P521	P461	P461
G1290	E1230	T1102	L1038	R970	L839	L778	A715	K654	P522	Q462	Q462
S1291	E1231	H1103	C1039	L971	R840	E779	P718	R655	D523	L464	L464
P1292	G1232	E1104	G1040	L972	Y841	A779	V719	P656	L524	L465	L465
F1293	P1233	I1105	L1041	Q973	W842	K780	L720	L657	R525	R466	R466
M1168	T1234	V1106	R1042	T974	F843	P781	L721	L658	P526	K467	K467
V1171	Q1235	V1107	G1043	E975	A844	S782	E722	L659	M527	L468	L468
H1172	L1236	R1108	L1044	Q976	D847	R783	E722	G595	V530	D469	D469
E1109	THR	E1108	M1045	A977	E848	D784	Q723	G596	D531	L470	L470
L1174	ARG	A1110	K1046	Y978	E849	I785	Q724	R661	G532	E471	E471
D1111	THR	C1112	K1047	E979	A849	I786	S725	E662	G533	A472	A472
C1112	PHE	D1113	P1048	N980	L850	L787	L726	R663	R601	L473	L473
G1113	HIS	T1114	S1049	Q981	L851	G788	Q727	R664	E474	E474	E474
T1115	THR	T1115	G1050	F982	W853	L789	L728	L603	T537	K475	K475
M1116	GLY	N1116	T1052	T984	E852	Y790	H729	T604	S538	E476	E476
Y1117	GLY	Y1117	F1053	D985	R857	I792	P730	L606	D539	L477	L477
I1118	VAL	I1118	E1054	R986	W858	T793	W732	L607	M541	L478	L478
S1119	ALA	S1119	V1055	E987	D859	Q794	C733	V670	D542	E479	E479
V1120	GLY	P1056	P1056	R988	E862	Y795	E734	R671	E480	E480	E480
P1121	ALA	V1057	V1057	Y989	R863	R796	A735	A672	L543	M481	M481
L1122	ALA	D990	K926	E925	W864	K799	F736	R613	L544	Y544	Y544
F1123	ASP	S1060	T927	K926	T865	W800	N737	R615	R545	R483	R483
Q1124	IIE	F1061	Q991	T927	T865	K800	A738	R616	R546	P484	P484
P1125	THR	R1062	L993	R929	T865	K800	D739	Q616	L547	S485	S485
D1126	THR	E1063	L994	R994	W868	G803	F740	R617	L548	R486	R486
E1127	GLN	G1064	L995	L995	G870	L804	D741	L618	M549	A487	A487
Q1195	L1256	L1065	L1065	W996	G870	E805	G742	L619	R550	R488	R488







● Molecule 7: DNA-directed RNA polymerase omega chain



4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	156.21Å 156.21Å 499.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 19.99 – 2.50	Depositor EDS
% Data completeness (in resolution range)	90.2 (20.00-2.50) 83.7 (19.99-2.50)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.95 (at 2.50Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.238 , 0.267 0.238 , 0.261	Depositor DCC
R_{free} test set	19570 reflections (5.74%)	wwPDB-VP
Wilson B-factor (Å ²)	54.7	Xtriage
Anisotropy	0.093	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 124.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.149 for h,-k,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	52719	wwPDB-VP
Average B, all atoms (Å ²)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.95% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	G	1.12	2/520 (0.4%)	1.17	5/798 (0.6%)
1	X	1.05	1/520 (0.2%)	1.13	2/798 (0.3%)
2	H	1.84	1/387 (0.3%)	2.39	37/601 (6.2%)
2	Y	1.31	1/387 (0.3%)	2.56	40/601 (6.7%)
3	I	0.94	1/304 (0.3%)	0.89	0/467
3	Z	0.84	1/304 (0.3%)	0.90	0/467
4	A	0.74	0/1838	0.82	1/2498 (0.0%)
4	B	0.76	0/1838	0.79	3/2498 (0.1%)
4	K	0.72	0/1838	0.82	1/2498 (0.0%)
4	L	0.72	0/1838	0.80	4/2498 (0.2%)
5	C	0.78	0/8997	0.93	17/12164 (0.1%)
5	M	0.79	1/8997 (0.0%)	0.93	20/12164 (0.2%)
6	D	0.83	9/10452 (0.1%)	0.92	21/14116 (0.1%)
6	N	0.80	2/10452 (0.0%)	0.91	15/14116 (0.1%)
7	E	0.85	1/784 (0.1%)	1.18	6/1057 (0.6%)
7	O	0.82	1/784 (0.1%)	1.08	5/1057 (0.5%)
All	All	0.82	21/50240 (0.0%)	0.97	177/68398 (0.3%)

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
1	G	0	5
1	X	0	3
2	H	0	2
2	Y	0	3
3	I	0	1
3	Z	0	1
All	All	0	15

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	1	G	OP3-P	25.71	1.92	1.61
6	D	133	ILE	N-CA	9.15	1.64	1.46
6	D	132	TYR	CA-C	8.40	1.74	1.52
1	X	1	DC	OP3-P	-8.00	1.51	1.61
1	G	1	DC	OP3-P	-7.45	1.52	1.61

The worst 5 of 177 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	Y	2	A	N9-C1'-C2'	-18.13	90.43	114.00
7	E	94	PRO	CA-N-CD	-16.74	88.07	111.50
2	Y	3	G	O4'-C1'-N9	-13.31	97.55	108.20
2	H	7	G	N9-C1'-C2'	-11.21	99.43	114.00
2	Y	7	G	N9-C1'-C2'	-11.05	99.63	114.00

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	1	DC	Sidechain
1	G	13	DT	Sidechain
1	G	16	DG	Sidechain
1	G	17	DC	Sidechain
1	G	18	DG	Sidechain

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	467	0	259	45	0
1	X	467	0	259	34	0
2	H	347	0	174	68	0
2	Y	347	0	175	77	0
3	I	270	0	144	13	0
3	Z	270	0	144	15	0
4	A	1806	0	1861	227	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	1806	0	1861	179	0
4	K	1806	0	1861	192	0
4	L	1806	0	1861	172	0
5	C	8829	0	8933	1212	0
5	M	8829	0	8933	1123	0
6	D	10280	0	10510	1429	0
6	N	10280	0	10510	1343	0
7	E	770	0	784	104	0
7	O	770	0	784	108	0
8	D	2	0	0	0	0
8	N	2	0	0	0	0
9	D	1	0	0	0	0
9	N	1	0	0	0	0
10	A	144	0	0	51	0
10	B	159	0	0	38	0
10	C	658	0	0	189	0
10	D	760	0	0	210	0
10	E	70	0	0	15	0
10	G	22	0	0	4	0
10	H	18	0	0	1	0
10	I	36	0	0	4	0
10	K	132	0	0	39	0
10	L	121	0	0	23	0
10	M	575	0	0	168	0
10	N	750	0	0	226	0
10	O	61	0	0	23	0
10	X	25	0	0	5	0
10	Y	16	0	0	2	0
10	Z	16	0	0	2	0
All	All	52719	0	49053	5880	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 60.

The worst 5 of 5880 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)
6:D:132:TYR:C	6:D:132:TYR:CA	1.74	1.56
7:O:95:VAL:CG1	10:O:2132:HOH:O	1.89	1.21
2:H:2:A:OP2	6:D:671:LYS:HD2	1.47	1.14
6:D:165:LYS:HB2	6:D:397:LYS:HB2	1.31	1.13

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:N:619:LEU:HD12	6:N:621:LYS:HZ3	1.08	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	
4	A	227/315 (72%)	206 (91%)	16 (7%)	5 (2%)	6	10
4	B	227/315 (72%)	205 (90%)	18 (8%)	4 (2%)	8	14
4	K	227/315 (72%)	205 (90%)	17 (8%)	5 (2%)	6	10
4	L	227/315 (72%)	205 (90%)	18 (8%)	4 (2%)	8	14
5	C	1117/1119 (100%)	916 (82%)	142 (13%)	59 (5%)	2	2
5	M	1117/1119 (100%)	918 (82%)	145 (13%)	54 (5%)	2	2
6	D	1297/1524 (85%)	1081 (83%)	165 (13%)	51 (4%)	3	4
6	N	1297/1524 (85%)	1100 (85%)	147 (11%)	50 (4%)	3	4
7	E	93/99 (94%)	76 (82%)	8 (9%)	9 (10%)	0	0
7	O	93/99 (94%)	75 (81%)	9 (10%)	9 (10%)	0	0
All	All	5922/6744 (88%)	4987 (84%)	685 (12%)	250 (4%)	3	3

5 of 250 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	29	GLU
4	A	187	GLY
4	B	29	GLU
4	B	187	GLY
5	C	40	GLU

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	
4	A	202/273 (74%)	164 (81%)	38 (19%)	1	2
4	B	202/273 (74%)	171 (85%)	31 (15%)	2	5
4	K	202/273 (74%)	170 (84%)	32 (16%)	2	4
4	L	202/273 (74%)	166 (82%)	36 (18%)	2	3
5	C	941/941 (100%)	710 (76%)	231 (24%)	0	1
5	M	941/941 (100%)	740 (79%)	201 (21%)	1	2
6	D	1100/1279 (86%)	874 (80%)	226 (20%)	1	2
6	N	1100/1279 (86%)	879 (80%)	221 (20%)	1	2
7	E	84/88 (96%)	62 (74%)	22 (26%)	0	0
7	O	84/88 (96%)	66 (79%)	18 (21%)	1	2
All	All	5058/5708 (89%)	4002 (79%)	1056 (21%)	1	2

5 of 1056 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	D	1135	ARG
4	L	81	ASN
6	N	1100	ASP
6	D	1285	GLU
7	E	58	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 143 such sidechains are listed below:

Mol	Chain	Res	Type
6	D	1404	ASN
4	L	16	GLN
6	N	976	GLN
6	D	1445	HIS
4	K	81	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	H	16/16 (100%)	12 (75%)	8 (50%)
2	Y	15/16 (93%)	11 (73%)	7 (46%)
All	All	31/32 (96%)	23 (74%)	15 (48%)

5 of 23 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	H	2	A
2	H	3	G
2	H	6	U
2	H	7	G
2	H	8	C

5 of 15 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	H	13	C
2	H	15	C
2	Y	12	G
2	H	12	G
2	Y	9	G

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 ⓘ

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, 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	G	23/23 (100%)	-0.38	0	100	100	16, 38, 58, 69	0
1	X	23/23 (100%)	-0.45	0	100	100	18, 41, 74, 89	0
2	H	16/16 (100%)	-0.74	0	100	100	35, 64, 80, 84	0
2	Y	16/16 (100%)	-0.84	0	100	100	31, 54, 90, 95	0
3	I	13/14 (92%)	-0.56	0	100	100	32, 42, 61, 78	0
3	Z	13/14 (92%)	-0.54	0	100	100	37, 49, 76, 84	0
4	A	229/315 (72%)	0.09	4 (1%)	70	72	45, 75, 94, 101	0
4	B	229/315 (72%)	0.08	10 (4%)	34	37	57, 79, 94, 104	0
4	K	229/315 (72%)	0.09	7 (3%)	49	52	56, 76, 91, 97	0
4	L	229/315 (72%)	0.21	14 (6%)	21	22	48, 82, 94, 101	0
5	C	1119/1119 (100%)	0.08	46 (4%)	37	40	19, 68, 98, 119	0
5	M	1119/1119 (100%)	0.06	28 (2%)	57	61	32, 68, 96, 108	0
6	D	1303/1524 (85%)	0.07	38 (2%)	51	55	34, 68, 94, 110	0
6	N	1303/1524 (85%)	0.06	34 (2%)	56	59	35, 69, 95, 108	0
7	E	95/99 (95%)	-0.00	3 (3%)	47	51	50, 70, 90, 96	0
7	O	95/99 (95%)	0.22	5 (5%)	26	28	41, 72, 98, 102	0
All	All	6054/6850 (88%)	0.07	189 (3%)	49	52	16, 70, 95, 119	0

The worst 5 of 189 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	D	135	LEU	7.9
6	N	1408	ILE	7.5
5	M	779	GLY	7.3
5	C	221	LEU	6.9
6	D	452	ILE	6.3

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
8	ZN	N	5058	1/1	0.96	0.17	71,71,71,71	0
8	ZN	D	4058	1/1	0.96	0.10	82,82,82,82	0
9	MG	N	8002	1/1	0.98	0.08	25,25,25,25	0
9	MG	D	8001	1/1	0.99	0.11	27,27,27,27	0
8	ZN	D	6112	1/1	0.99	0.14	65,65,65,65	0
8	ZN	N	7112	1/1	0.99	0.14	73,73,73,73	0

6.5 Other polymers [i](#)

There are no such residues in this entry.