



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 02:41 am BST

PDB ID : 2JA7  
Title : CPD lesion containing RNA Polymerase II elongation complex C  
Authors : Brueckner, F.; Hennecke, U.; Carell, T.; Cramer, P.  
Deposited on : 2006-11-23  
Resolution : 3.80 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
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

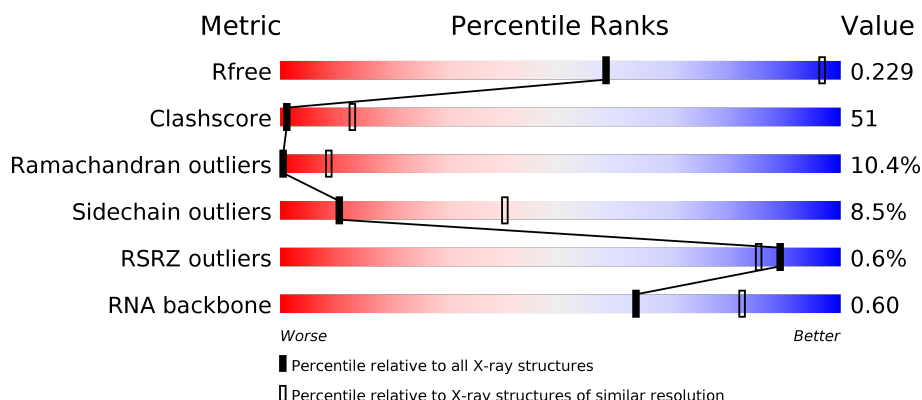
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.80 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1212 (4.00-3.60)
Clashscore	141614	1288 (4.00-3.60)
Ramachandran outliers	138981	1243 (4.00-3.60)
Sidechain outliers	138945	1237 (4.00-3.60)
RSRZ outliers	127900	1121 (4.00-3.60)
RNA backbone	3102	1036 (4.60-3.00)

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 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	1	14	<div> <div>29%</div> <div>21%</div> <div>50%</div> </div>
1	4	14	<div> <div>29%</div> <div>21%</div> <div>50%</div> </div>
2	2	25	<div> <div>52%</div> <div>16%</div> <div>28%</div> </div>
2	5	25	<div> <div>8%</div> <div>48%</div> <div>16%</div> <div>28%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	3	11	
3	6	11	
4	A	1733	
4	M	1733	
5	B	1224	
5	N	1224	
6	C	318	
6	O	318	
7	D	221	
7	P	221	
8	E	215	
8	Q	215	
9	F	155	
9	R	155	
10	G	171	
10	S	171	
11	H	146	
11	T	146	
12	I	122	
12	U	122	
13	J	70	
13	V	70	
14	K	120	
14	W	120	
15	L	70	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
15	X	70	<div> <div>20%</div> <div>36%</div> <div>9%</div> <div>34%</div> </div>

## 2 Entry composition

There are 17 unique types of molecules in this entry. The entry contains 63924 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(\*TP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP\*AP\*GP\*CP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	1	7	Total	C	N	O	P	0	0	0
			141	69	27	39	6			
1	4	7	Total	C	N	O	P	0	0	0
			141	69	27	39	6			

- Molecule 2 is a DNA chain called 5'-D(\*AP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*TTP\*CP\*CP\*BRUP\*GP\*GP\*TP\*CP\*AP\*TP\*T)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	2	18	Total	Br	C	N	O	P	0	0	0
			380	1	186	60	116	17			
2	5	18	Total	Br	C	N	O	P	0	0	0
			380	1	186	60	116	17			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	3	10	Total	C	N	O	P	0	0	0
			212	96	41	66	9			
3	6	10	Total	C	N	O	P	0	0	0
			212	96	41	66	9			

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II LARGEST SUB-UNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1421	Total	C	N	O	S	0	0	0
			11186	7048	1958	2118	62			
4	M	1421	Total	C	N	O	S	0	0	0
			11186	7048	1958	2118	62			

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE II 140 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1115	Total	C	N	O	S	0	0	0
			8866	5614	1553	1644	55			
5	N	1115	Total	C	N	O	S	0	0	0
			8866	5614	1553	1644	55			

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASE II 45KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	267	Total	C	N	O	S	0	0	0
			2101	1320	349	419	13			
6	O	267	Total	C	N	O	S	0	0	0
			2101	1320	349	419	13			

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II 32KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	D	177	Total	C	N	O	S	0	0	0
			1427	882	256	287	2			
7	P	177	Total	C	N	O	S	0	0	0
			1427	882	256	287	2			

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 27 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	E	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			
8	Q	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 23 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	F	87	Total	C	N	O	S	0	0	0
			705	451	119	132	3			
9	R	87	Total	C	N	O	S	0	0	0
			705	451	119	132	3			

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASE II 19KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	G	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			
10	S	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 14.5 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	H	135	Total	C	N	O	S	0	0	0
			1084	683	183	214	4			
11	T	135	Total	C	N	O	S	0	0	0
			1084	683	183	214	4			

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	I	116	Total	C	N	O	S	0	0	0
			944	581	172	181	10			
12	U	116	Total	C	N	O	S	0	0	0
			944	581	172	181	10			

- Molecule 13 is a protein called DNA-DIRECTED RNA POLYMERASES I/II/III SUBUNIT 10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			
13	V	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 14 is a protein called DNA-DIRECTED RNA POLYMERASE II 13.6 KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			
14	W	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 15 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III 7.7

## KDA POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	L	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			
15	X	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	1	Total	Mg	0	0
			1	1		
16	M	1	Total	Mg	0	0
			1	1		

- Molecule 17 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	J	1	Total	Zn	0	0
			1	1		
17	B	1	Total	Zn	0	0
			1	1		
17	I	2	Total	Zn	0	0
			2	2		
17	C	1	Total	Zn	0	0
			1	1		
17	V	1	Total	Zn	0	0
			1	1		
17	A	2	Total	Zn	0	0
			2	2		
17	N	1	Total	Zn	0	0
			1	1		
17	U	2	Total	Zn	0	0
			2	2		
17	X	1	Total	Zn	0	0
			1	1		
17	O	1	Total	Zn	0	0
			1	1		
17	L	1	Total	Zn	0	0
			1	1		
17	M	2	Total	Zn	0	0
			2	2		

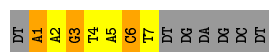


### 3 Residue-property plots [i](#)


These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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(\*TP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP \*AP\*GP\*CP\*T)-3'

Chain 1: 



- Molecule 1: 5'-D(\*TP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP \*AP\*GP\*CP\*T)-3'

Chain 4: 



- Molecule 2: 5'-D(\*AP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP \*CP\*TP\*TP\*TP\*TTP\*CP\*CP\*BRUP\*GP\*GP\*TP\*CP\*AP\*TP\*T)-3'

Chain 2: 



- Molecule 2: 5'-D(\*AP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP \*CP\*TP\*TP\*TP\*TTP\*CP\*CP\*BRUP\*GP\*GP\*TP\*CP\*AP\*TP\*T)-3'

Chain 5: 



- Molecule 3: 5'-R(\*UP\*UP\*CP\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*AP)-3'

Chain 3: 



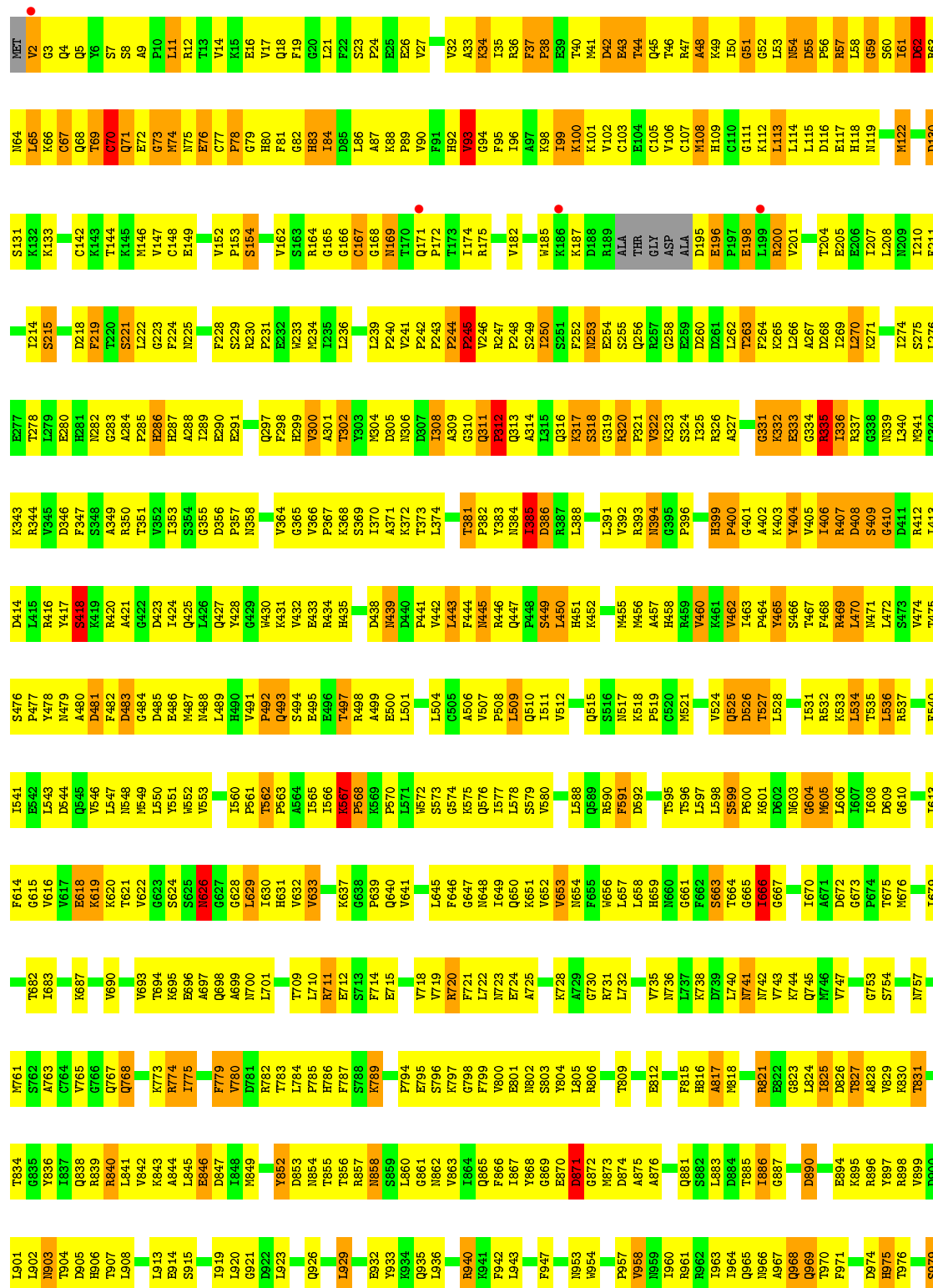
- Molecule 3: 5'-R(\*UP\*UP\*CP\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*AP)-3'

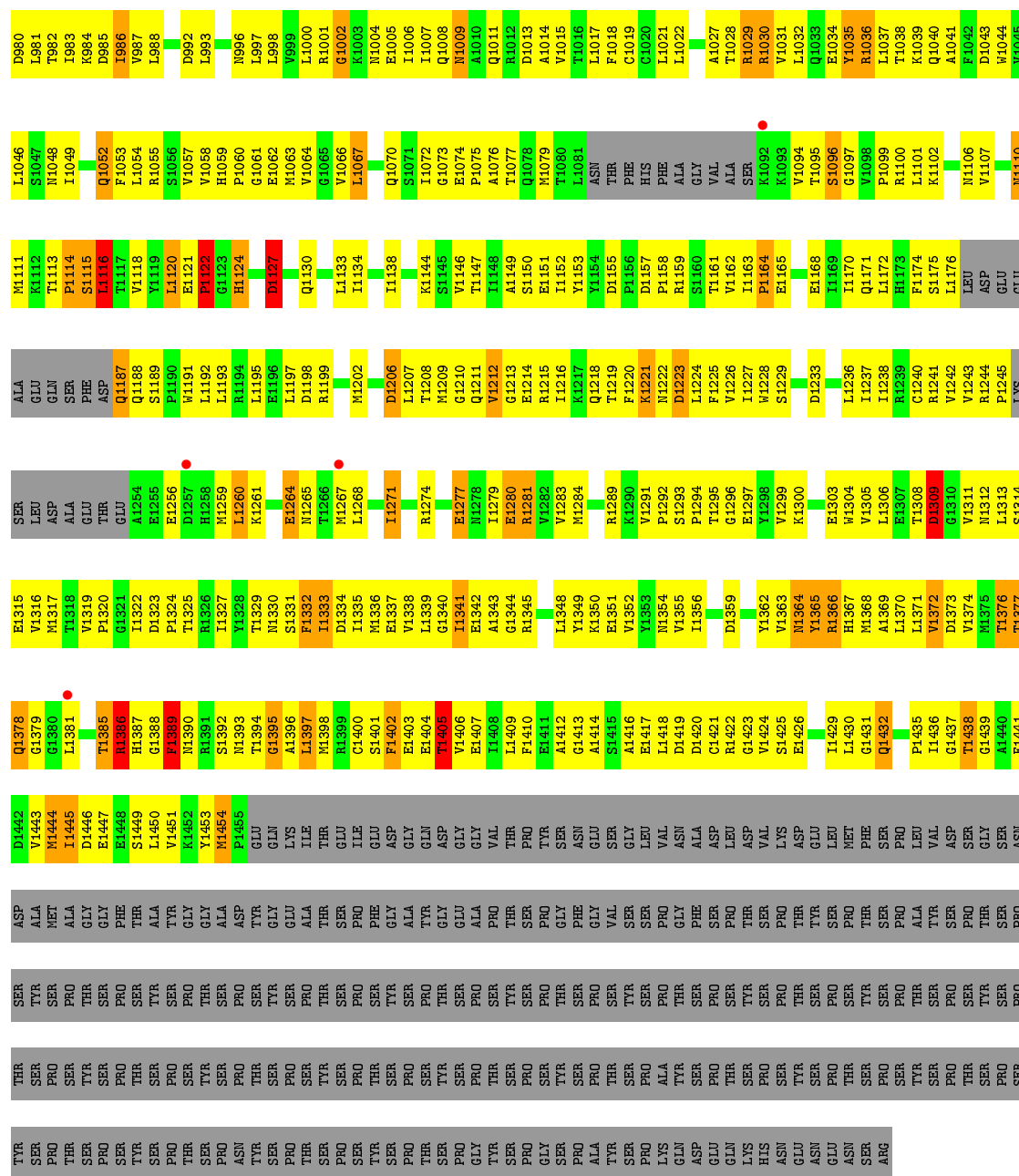
Chain 6: 



• Molecule 4: DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT

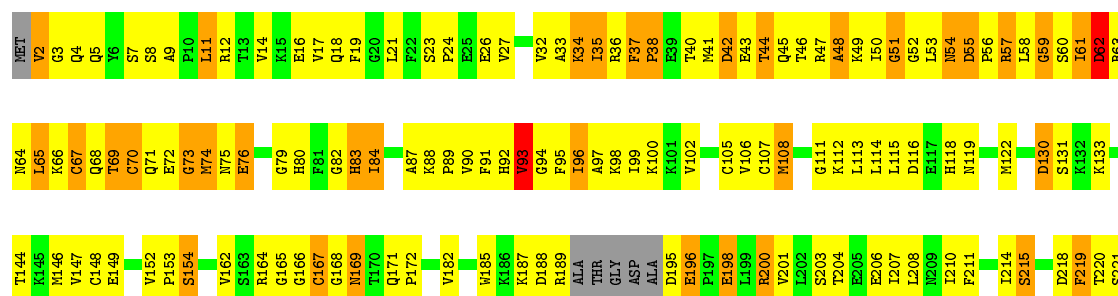
Chain A: 26% 45% 10% 18%



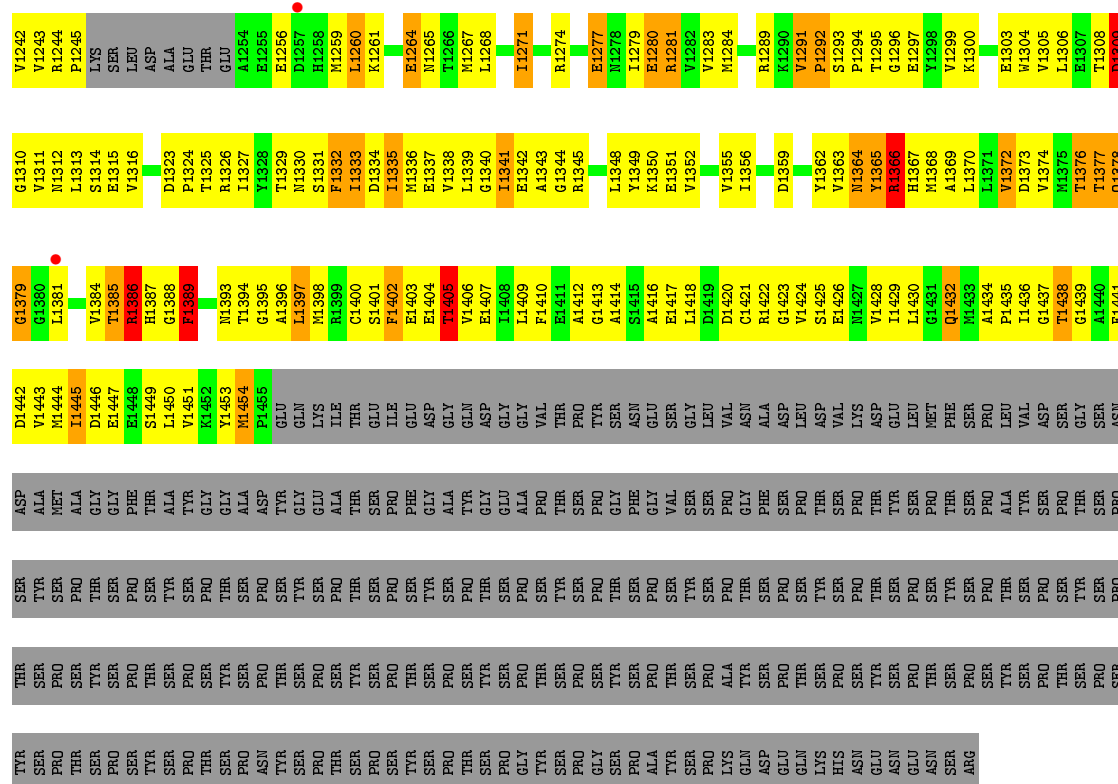


# • Molecule 4: DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT

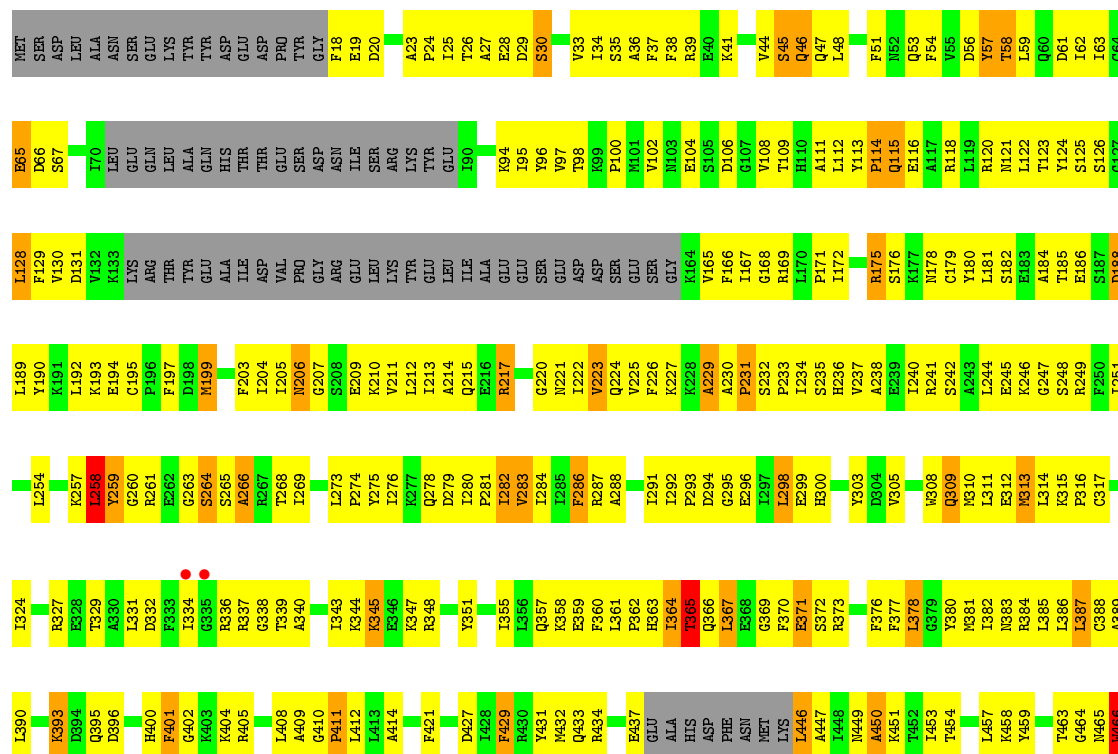
Chain M: 26% 45% 10% 18%

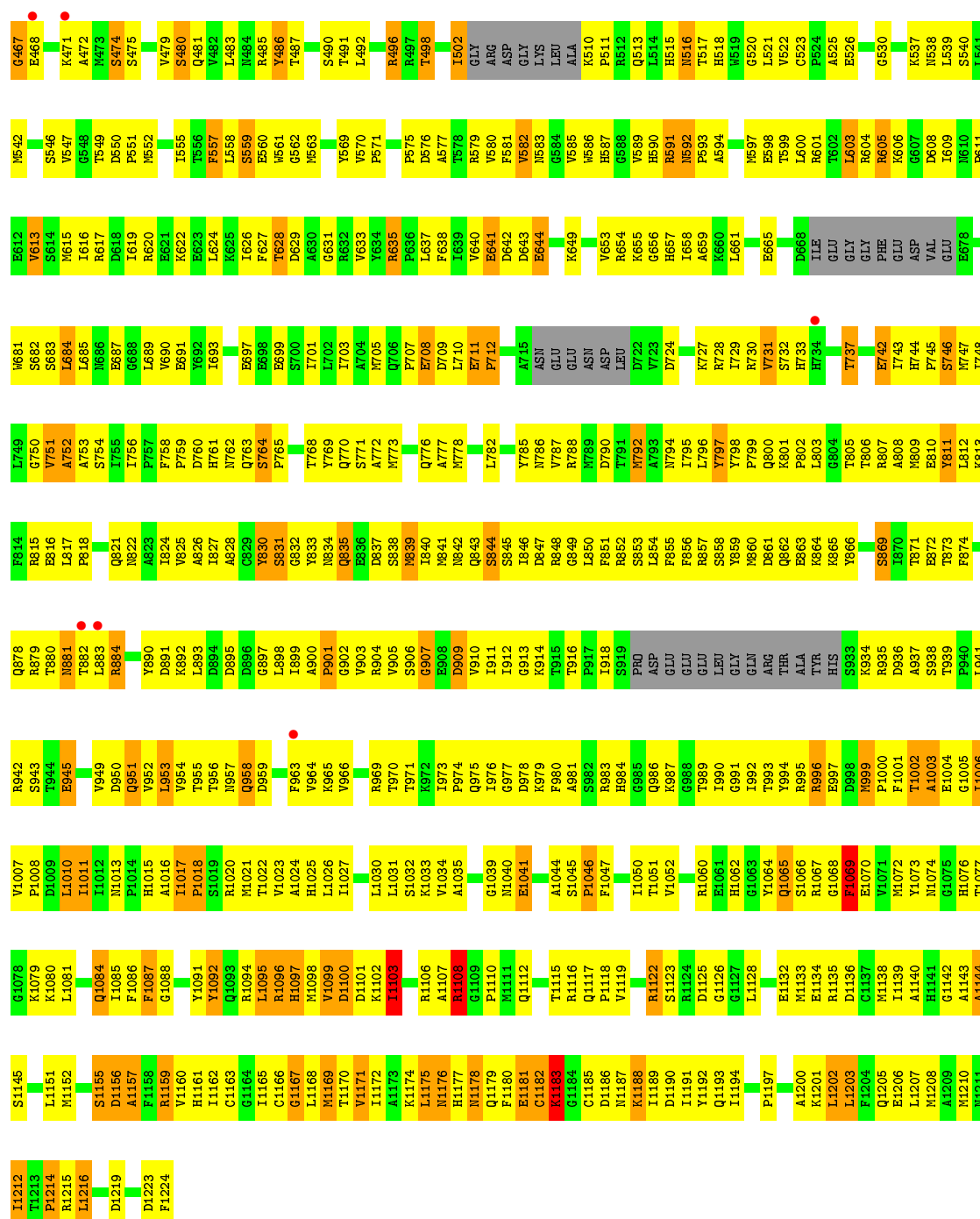


L1176	L1106	D1043	T976	D900	1837	V765	V690	E618	L547	F482	A421	R350	A284	L222
LEU	V1107	L1046	S979	L901	Q838	G766	T694	K619	M548	D483	G422	T351	P285	G223
ASP		L1047	D980	L902	R839	Q767	T694	D484	M549	D485	I423	V352	H286	F224
GLU	N1110	N1047	L981	T904	R840	Q768	T694	T621	M550		I424	I353	H287	N235
ALA	M1112	N1048	T982	D905	R842	R774	E696	E486	F551	E487	Q425	S354	A288	E226
GLU	T1113	L1049	I983	H906	R843	R775	Q698	G623	M552	G355	I426	G355	I289	V227
GLN	P1114	E1050	K984	H907	A844	I775	Q699	S624	M553	D356	Q427	D357	E290	F228
SER	S1115	A1051	D985	L908	L845	F779	A699	S625	T560	L488	Y428	P357	E291	S229
ASP	L1116	Q1052	I986		R846	V780	L701	R626	P661	H490	G429	N358	R230	R230
ASP	T1117	T1053	V987	L913	D847	D781	T709	V491	T562	V491	K430	D362	Q297	P231
Q1187	V1118	L1054	R848	R914	R848	R782	T709	K431	E232	K431	K431	D362	F298	E232
Q1188	Y1119	R1055	R849	S915	R849	T783	L710	Q493	K363	Q363	Q493	Q363	H299	M233
S1189	L1120	S1056				L784	R711	H631	A584	V364	V364	V364	H300	M234
P1190	E1121	V1057				R785	E712	V632	E585	G365	R434	G365	A301	L235
W1191	P1122	H1058				F786	S713	V633	I566	V366	P367	V366	T302	L236
L1192	G1123	H1059				F787	E715	E636	K567	K368	M304	K368	H299	
H1124	H1124	P1060				K788	E715	E500	P568	S369	D305	S369	V300	P240
R1194	G1061	T856				K789			K569	I370	N306	I370	N306	V241
L1195	E1062	R857							P570	D440	D440	A371	D367	P242
L1196	M1063	R858				F794			M572	V442	V442	K372	I308	P243
L1197	V1064	S859				E795			S573	L443	L443	T381	A309	P244
D1198	G1085	L860				S796			G574	F444	F444		G310	P245
R1199	V1066	R861				K797			K575	N445	N445	P382	Q311	V246
	L1067	R862				G798			L509	R446	R446	Y383	R312	R247
		V863				F799			Q510	Q447	Q447	N384	Q313	P248
		R864				R800			I511	P448	P448	I385	Q314	S249
		R865				E801			V512	S449	S449	D386	L315	I250
		R866				R802			S513	L450	L450	K316	L251	
		R867				S803			P514	H481	H481	K317	F252	
		R868				G804			Q515	K452	K452	S318	N253	
		R869				L805			S516	M453	M453	L391	G319	E254
		R870				R806			L588	M517	M517	V392	S285	
		R871				G807			Q589	K518	M455	R393	P321	Q256
		R872				L808			F591	R590	M456	N394	V322	R257
		R873				T809			G520	A457	A457	G395	K323	G258
		R874							M521	H458	H458	S324	S324	E259
		R875				E812			D592	R459	R459	I325	I325	D260
		R876				F813			T595	V460	V460	R399	R326	D261
		R877				F814			T596	K461	K461	P400	A327	L262
		R878				F815			L597	V462	V462	G401		T263
						H816			L598	I463	I463	A402	F264	
						A817			S599	P464	P464	K403	K265	
						M818			P600	Y465	Y465	Y404	L266	
									K601	S466	S466	V405	A267	
						R821			D602	T467	T467	I406	R335	D268
						E822			N603	F468	F468	R407	I336	I269
						G823			G604	D469	D469	D408	R337	L270
						L824			M605	L470	L470	S409	G338	
						T825			L606	N471	N471	G410	N339	K271
						D826			T607	L472	L472	D411	N339	
						T827			L608	S473	S473	R412	L340	I274
						A828			D609	V474	V474	I413	M341	S275
						R829			G610	T475	T475	D414	K343	L276
						K830				S476	S476	L415	R344	E277
						T831				P477	P477	L416	V345	T278
						R834				Y417	Y417	R416	V345	L279
						G835				S418	S418	D346	F347	E280
						H836				A480	A480	K419	S348	N282
						V836				V546	V546	R420	A349	G283

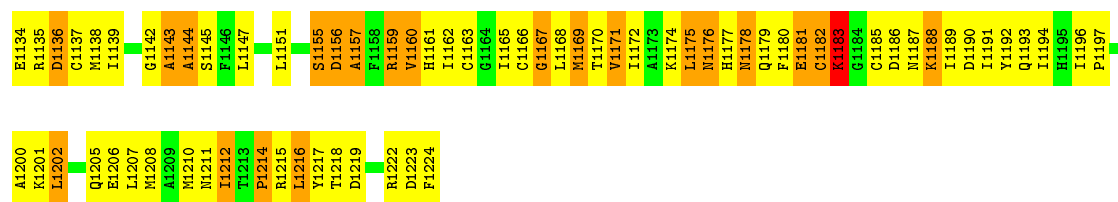


• Molecule 5: DNA-DIRECTED RNA POLYMERASE II 140 KDA POLYPEPTIDE

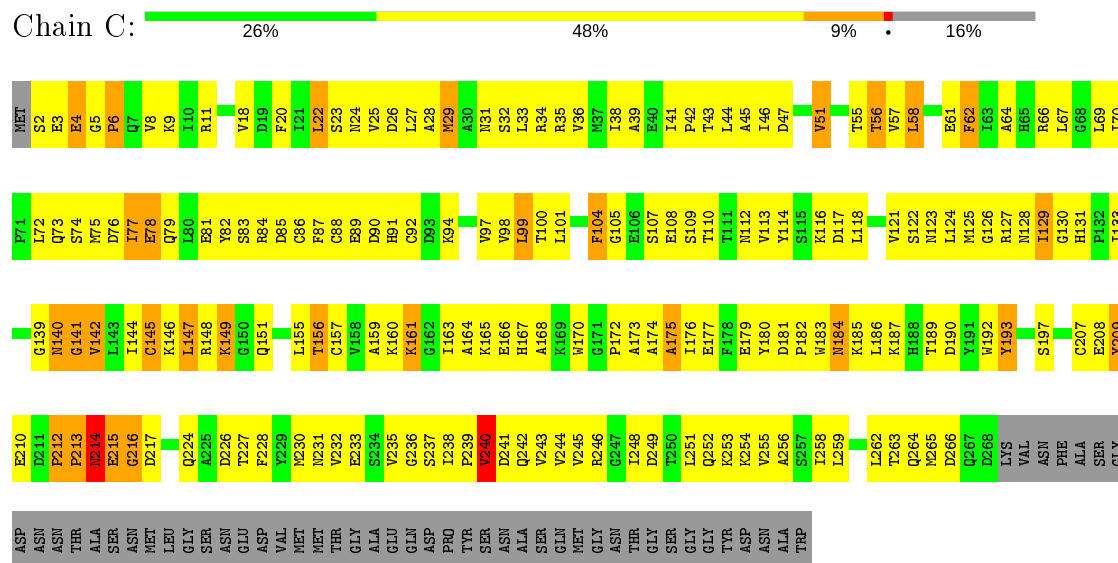




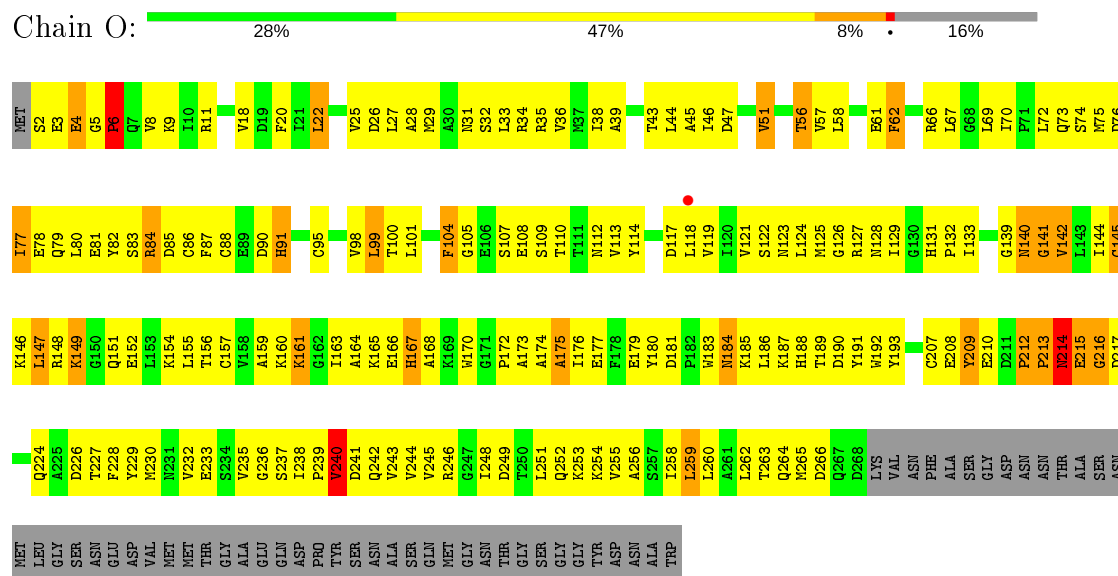
R1067	G1068	F1069	F1070	Y1071	F1001	T1002	A1003	E1004	G1005	V1007	P1008	D1009	L1010	Q1084	F1085	F1086	F1087	G1088	Y1091	H1092	Q1093	R1094	L1095	A1024	H1025	H1096	H1097	M1098	D1100	D1101	K1102	T1103	R1106	A1107	R1108	G1109	P1110	M1111	Q1112	T1115	R1116	Q1117	G1121	R1122	S1123	R1124	D1125	G1126	L1127	L1128	H1062	F1130	E1131	G1132	Y1064	Q1065	S1066																																																																											
HIS	S933	K934	R935	D936	A937	S938	T939	P940	L941	R942	S943	T944	E945	T948	V949	D950	Q951	L952	L953	V954	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																					
P799	Q800	K801	P802	L803	T806	R807	A808	M809	E810	Y811	L812	R815	P818	L882	L883	R884	D891	R892	L893	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																						
R730	V731	S732	H733	H734	T737	F738	E742	I743	H744	P745	S746	M747	I748	L749	G750	V751	A752	F753	S754	T755	T756	N757	Q758	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																						
Q800	K801	P802	L803	T806	R807	A808	M809	E810	Y811	L812	R815	P818	L882	L883	R884	D891	R892	L893	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																							
R730	V731	S732	H733	H734	T737	F738	E742	I743	H744	P745	S746	M747	I748	L749	G750	V751	A752	F753	S754	T755	T756	N757	Q758	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																						
Q800	K801	P802	L803	T806	R807	A808	M809	E810	Y811	L812	R815	P818	L882	L883	R884	D891	R892	L893	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																							
K864	R865	Y866	S869	T870	T871	E872	T873	F874	P877	R878	R879	M881	L882	L883	R884	D891	R892	L893	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																							
HIS	S933	K934	R935	D936	A937	S938	T939	P940	L941	R942	S943	T944	E945	T948	V949	D950	Q951	L952	L953	V954	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																					
R996	E997	M998	D999	F1000	F1001	T1002	A1003	E1004	G1005	V1007	P1008	D1009	L1010	Q1084	F1085	F1086	F1087	G1088	Y1091	H1092	Q1093	R1094	L1095	A1024	H1025	H1096	H1097	M1098	D1100	D1101	K1102	T1103	R1106	A1107	R1108	G1109	P1110	M1111	Q1112	T1115	R1116	Q1117	G1121	R1122	S1123	R1124	D1125	G1126	L1127	L1128	H1062	F1130	E1131	G1132	Y1064	Q1065	S1066																																																																											
R996	E997	M998	D999	F1000	F1001	T1002	A1003	E1004	G1005	V1007	P1008	D1009	L1010	Q1084	F1085	F1086	F1087	G1088	Y1091	H1092	Q1093	R1094	L1095	A1024	H1025	H1096	H1097	M1098	D1100	D1101	K1102	T1103	R1106	A1107	R1108	G1109	P1110	M1111	Q1112	T1115	R1116	Q1117	G1121	R1122	S1123	R1124	D1125	G1126	L1127	L1128	H1062	F1130	E1131	G1132	Y1064	Q1065	S1066																																																																											
S126	G127	L128	F129	V130	D131	Y132	K133	L134	K135	L136	L137	L138	L139	L140	L141	L142	L143	L144	L145	L146	L147	L148	L149	L150	L151	L152	L153	L154	L155	L156	L157	L158	L159	L160	L161	L162	L163	L164	L165	L166	L167	L168	L169	L170	L171	L172	L173	L174	L175	L176	L177	L178	L179	L180	L181	L182	L183	L184	L185	L186																																																																								
S187	D188	L189	Y190	K191	L192	K193	E194	C195	Y196	G197	T198	D199	M199	F203	I204	I205	N206	G207	S208	E209	K210	L211	L212	L213	L214	L215	L216	L217	L218	L219	L220	L221	L222	L223	L224	L225	L226	L227	L228	L229	L230	L231	L232	L233	L234	L235	L236	L237	L238	L239	L240	L241	L242	L243	L244	L245	L246	L247	L248																																																																									
R249	F250	I251	L254	K257	L258	Y259	G260	R261	S264	S265	A266	R267	T268	I269	K270	L273	Y274	Y275	Y276	L277	K278	Q278	D279	L280	K281	L282	L283	L284	L285	L286	L287	L288	L289	L290	L291	L292	L293	L294	L295	L296	L297	L298	L299	L300	L301	L302	L303	L304	L305	L306	L307	L308	L309	L310	L311	L312	L313	L314	L315																																																																									
P316	C317	F322	L324	R327	R328	R329	R330	R331	R332	R333	R334	R335	R336	R337	R338	R339	R340	R341	R342	R343	R344	R345	R346	R347	R348	R349	R350	R351	R352	R353	R354	R355	R356	R357	R358	R359	R360	R361	R362	R363	R364	R365	R366	R367	R368	R369	R370	R371	R372	R373	R374	R375	R376	R377	R378	R379	R380	R381	R382	R383	R384																																																																							
L385	L386	L387	A388	L389	L390	R391	R392	R393	R394	R395	R396	D399	R400	R401	R402	R403	R404	R405	L408	L409	C410	P411	L412	L413	L414	L415	L416	L417	L418	L419	L420	L421	L422	L423	L424	L425	L426	L427	L428	L429	L430	L431	L432	L433	L434	L435	L436	L437	L438	L439	L440	L441	L442	L443	L444	L445	L446	L447	L448	L449	L450	L451	L452	L453	L454	L455																																																																		
L457	R458	Y459	A460	L461	A462	T463	R464	R465	R466	R467	R468	R469	R470	R471	R472	R473	R474	R475	S476	S477	S478	S479	S480	S481	S482	S483	S484	S485	S486	S487	S488	S489	S490	S491	S492	S493	S494	S495	S496	S497	S498	S499	S500	S501	S502	S503	S504	S505	S506	S507	S508	S509	S510	S511	S512	S513	S514	S515	S516	S517	S518	S519	S520	S521	S522	S523	S524	S525	S526																																																															
T527	P528	E529	C530	Q531	K537	N538	L539	M542	S546	V547	G548	T549	D550	P551	M552	T555	T556	F557	L558	S559	K560	R561	R562	R563	R564	R565	R566	R567	R568	R569	R570	R571	R572	R573	R574	R575	R576	R577	R578	R579	R580	R581	R582	R583	R584	R585	R586	R587	R588	R589	R590	R591	R592	R593	R594	R595	R596	R597	R598	R599	R600																																																																							
R601	T602	R603	R604	R605	K606	G607	D608	L609	P610	P611	E612	V613	S614	M615	I616	D617	D618	D619	D620	E621	L622	L623	L624	L625	L626	L627	L628	L629	L630	L631	L632	L633	L634	L635	L636	L637	L638	L639	L640	L641	L642	L643	L644	L645	L646	L647	L648	L649	L650	L651	L652	L653	L654	L655	L656	L657	L658	L659	L660	L661	L662	L663	L664	L665																																																																				
D668	I669	G670	G671	P672	P673	P674	P675	P676	P677	P678	P679	P680	P681	P682	P683	P684	P685	P686	P687	P688	P689	P690	P691	P692	P693	P694	P695	P696	P697	P698	P699	P700	P701	P702	P703	P704	P705	P706	P707	P708	P709	P710	P711	P712	P713	P714	P715	P716	P717	P718	P719	P720	P721	P722	P723	P724	P725	P726	P727	P728	P729	P730	P731	P732	P733	P734	P735	P736	P737	P738	P739	P740	P741	P742	P743	P744	P745	P746	P747	P748	P749	P750	P751	P752	P753	P754	P755	P756	P757	P758	P759	P760	P761	P762	P763	P764	P765	P766	P767	P768	P769	P770	P771	P772	P773	P774	P775	P776	P777	P778	P779	P780	P781	P782	P783	P784	P785	P786	P787	P788	P789	P790	P791	P792	P793	P794	P795	P796	P797	P798	P799	P800
P799	Q800	K801	P802	L803	T806	R807	A808	M809	E810	Y811	L812	R815	P818	L882	L883	R884	D891	R892	L893	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																						
K864	R865	Y866	S869	T870	T871	E872	T873	F874	P877	R878	R879	M881	L882	L883	R884	D891	R892	L893	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T970	E908	F972	T973	P974	Q975	T976	G977	D978	K979	F980	A981	S982	R983	H984	G987	G988	T989	L1060	E1061	G1062	G1063	L1092	G1093	T993	L994	Q1095	M1133																																																																							
HIS	S933	K934	R935	D936	A937	S938	T939	P940	L941	R942	S943	T944	E945	T948	V949	D950	Q951	L952	L953	V954	T955	T956	N957	Q958	H1026	L898	L899	G960	F961	G962	G963	Y964	K965	V966	R969	T97																																																																																																



• Molecule 6: DNA-DIRECTED RNA POLYMERASE II 45KDA POLYPEPTIDE



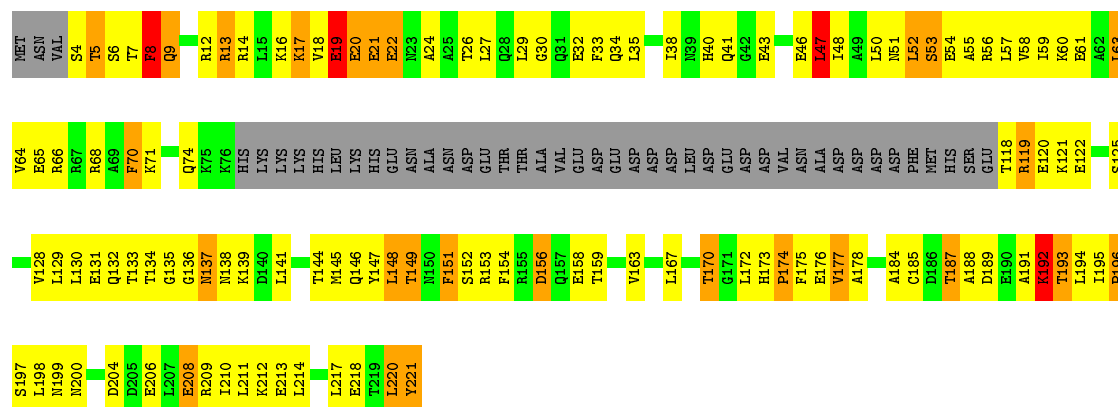
• Molecule 6: DNA-DIRECTED RNA POLYMERASE II 45KDA POLYPEPTIDE



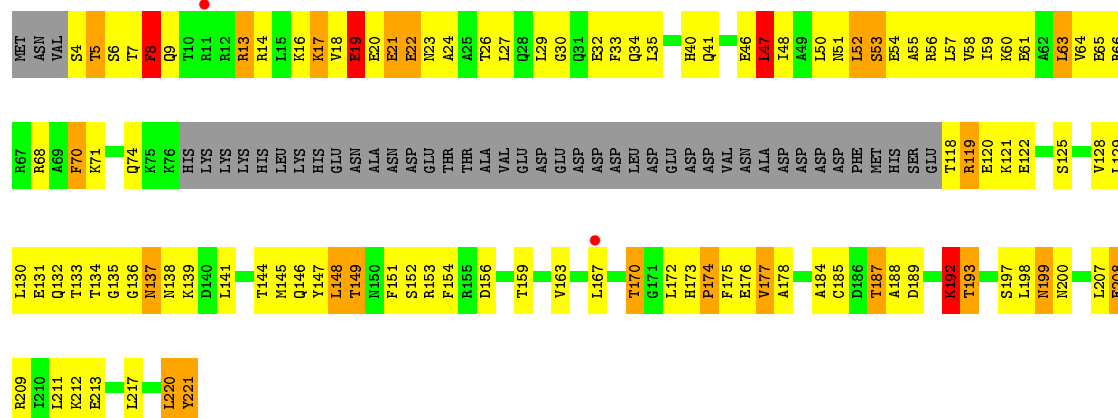
• Molecule 7: DNA-DIRECTED RNA POLYMERASE II 32KDA POLYPEPTIDE



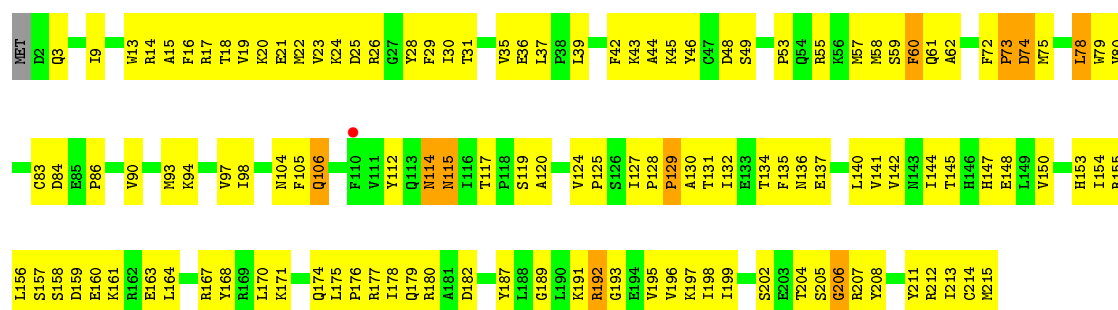




• Molecule 7: DNA-DIRECTED RNA POLYMERASE II 32KDA POLYPEPTIDE

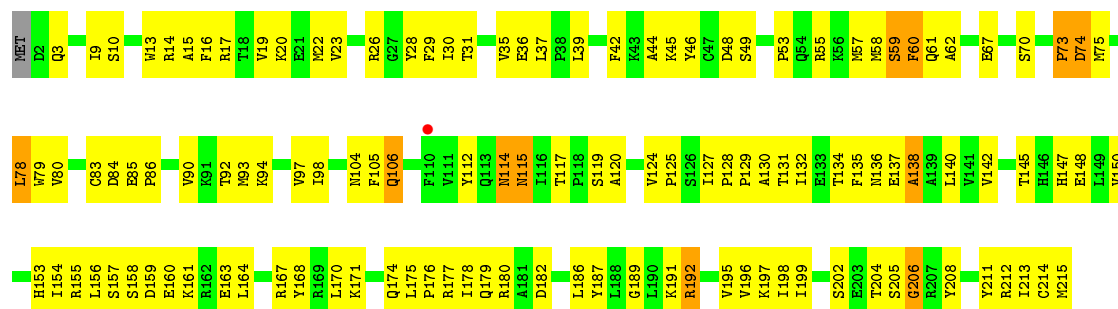


• Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III 27 KDA POLYPEPTIDE



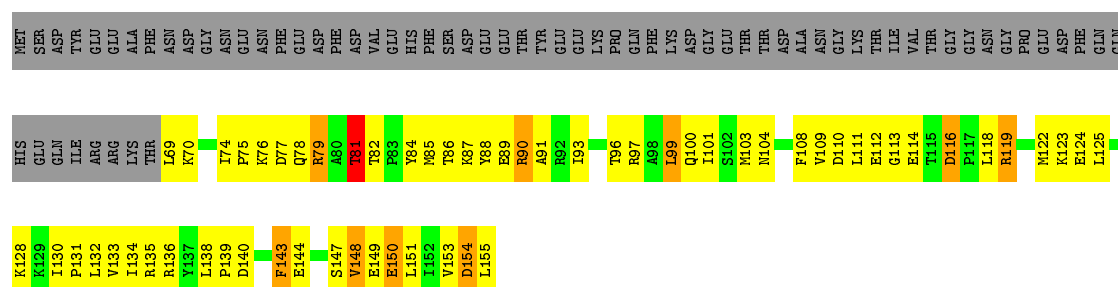
• Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III 27 KDA POLYPEPTIDE





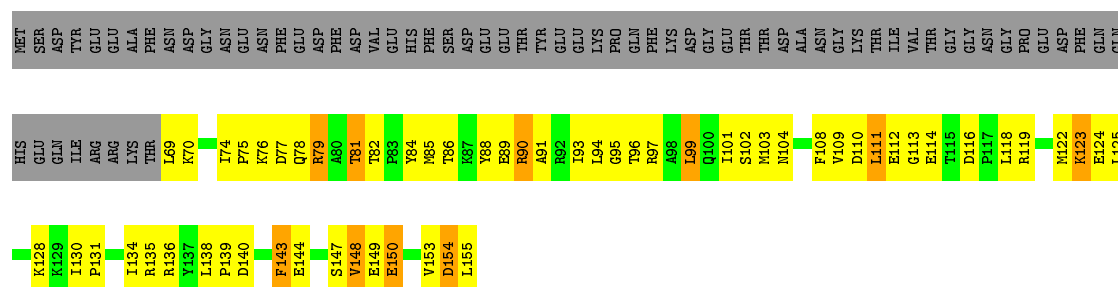
- Molecule 9: DNA-DIRECTED RNA POLYMERASES I, II, AND III 23 KDA POLYPEPTIDE

Chain F: 17% 33% 6% 44%



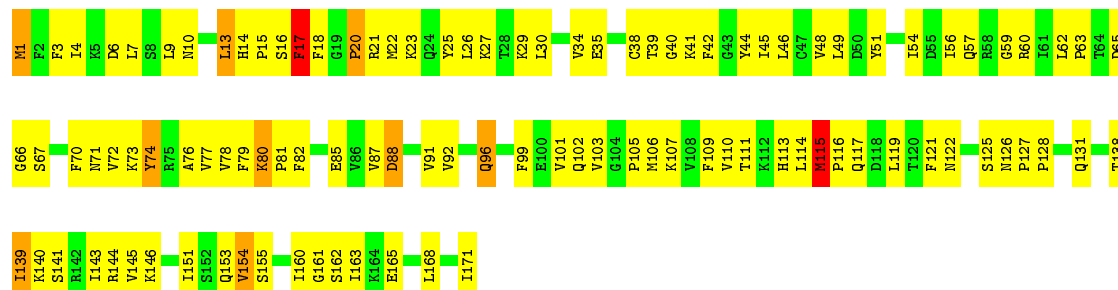
- Molecule 9: DNA-DIRECTED RNA POLYMERASES I, II, AND III 23 KDA POLYPEPTIDE

Chain R: 18% 32% 6% 44%

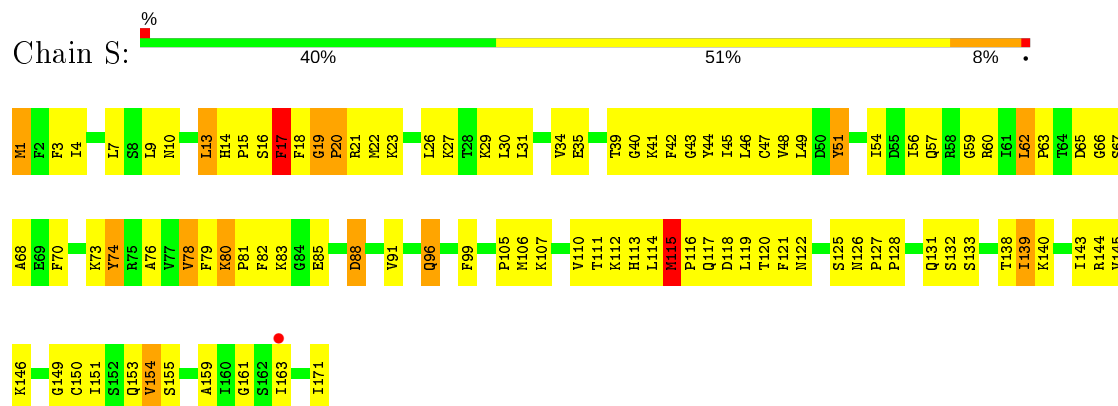


- Molecule 10: DNA-DIRECTED RNA POLYMERASE II 19KDA POLYPEPTIDE

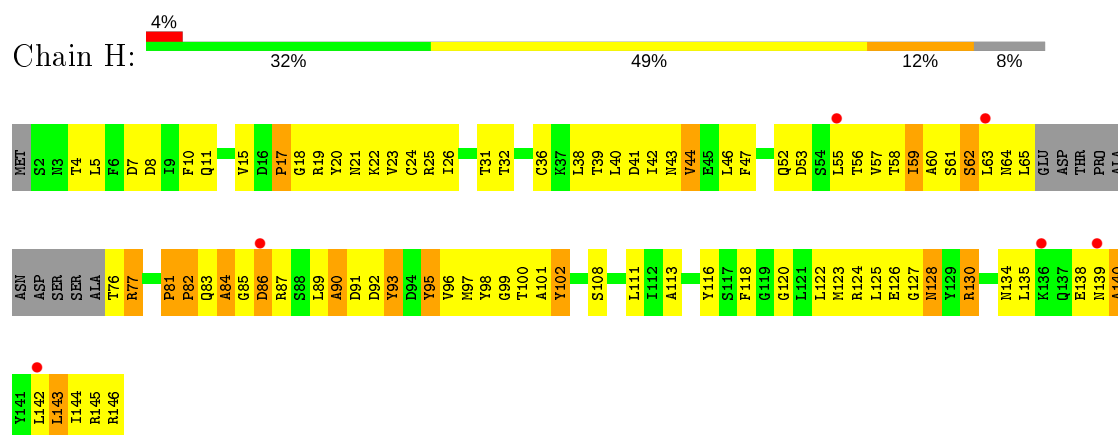
Chain G: 39% 55% 5%



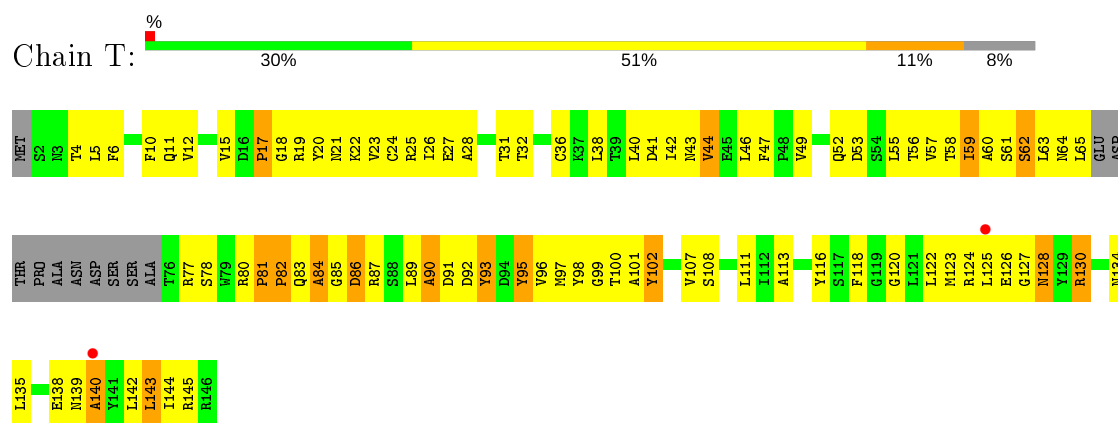
• Molecule 10: DNA-DIRECTED RNA POLYMERASE II 19KDA POLYPEPTIDE



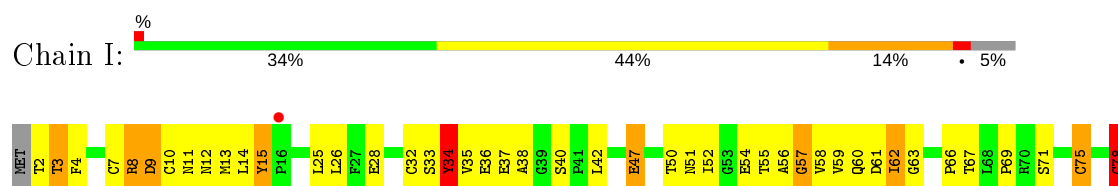
• Molecule 11: DNA-DIRECTED RNA POLYMERASES I, II, AND III 14.5 KDA POLYPEPTIDE

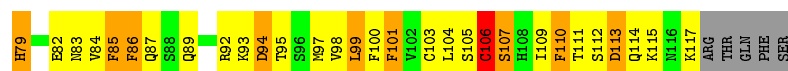


• Molecule 11: DNA-DIRECTED RNA POLYMERASES I, II, AND III 14.5 KDA POLYPEPTIDE



• Molecule 12: DNA-DIRECTED RNA POLYMERASE II SUBUNIT 9

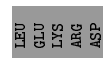
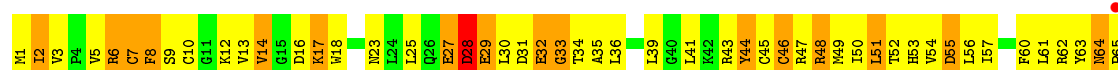
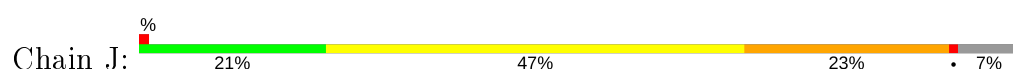




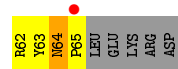
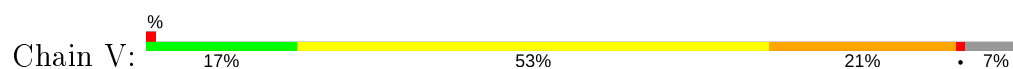
• Molecule 12: DNA-DIRECTED RNA POLYMERASE II SUBUNIT 9



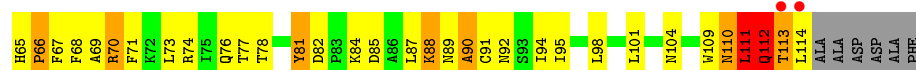
• Molecule 13: DNA-DIRECTED RNA POLYMERASES I/II/III SUBUNIT 10



• Molecule 13: DNA-DIRECTED RNA POLYMERASES I/II/III SUBUNIT 10

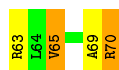


• Molecule 14: DNA-DIRECTED RNA POLYMERASE II 13.6 KDA POLYPEPTIDE



• Molecule 14: DNA-DIRECTED RNA POLYMERASE II 13.6 KDA POLYPEPTIDE





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	394.36Å 221.86Å 283.11Å 90.00° 90.56° 90.00°	Depositor
Resolution (Å)	50.00 – 3.80 48.96 – 3.80	Depositor EDS
% Data completeness (in resolution range)	98.1 (50.00-3.80) 96.7 (48.96-3.80)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.76 (at 3.77Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.257 , 0.275 0.216 , 0.229	Depositor DCC
$R_{free}$ test set	9042 reflections (1.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	116.2	Xtriage
Anisotropy	0.457	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 51.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	0.024 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.024 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.024 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.020 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.308 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	63924	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.93% 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, 0.333 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: MG, ZN, BRU, TT

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	1	1.37	0/158	1.57	3/242 (1.2%)
1	4	1.35	0/158	1.58	3/242 (1.2%)
2	2	1.47	1/357 (0.3%)	1.42	3/544 (0.6%)
2	5	1.46	1/357 (0.3%)	1.40	4/544 (0.7%)
3	3	1.55	4/237 (1.7%)	2.20	8/368 (2.2%)
3	6	1.59	4/237 (1.7%)	2.10	8/368 (2.2%)
4	A	0.48	0/11385	0.73	2/15393 (0.0%)
4	M	0.48	0/11385	0.73	2/15393 (0.0%)
5	B	0.47	0/9037	0.70	3/12181 (0.0%)
5	N	0.46	0/9037	0.70	2/12181 (0.0%)
6	C	0.48	0/2138	0.71	0/2896
6	O	0.50	0/2138	0.71	0/2896
7	D	0.44	0/1437	0.67	0/1925
7	P	0.46	0/1437	0.68	0/1925
8	E	0.43	0/1788	0.63	0/2406
8	Q	0.43	0/1788	0.63	0/2406
9	F	0.55	0/716	0.77	0/964
9	R	0.55	0/716	0.75	0/964
10	G	0.52	0/1368	0.74	0/1844
10	S	0.52	0/1368	0.74	0/1844
11	H	0.40	0/1102	0.65	0/1492
11	T	0.40	0/1102	0.65	0/1492
12	I	0.39	0/962	0.69	0/1295
12	U	0.41	0/962	0.69	0/1295
13	J	0.49	0/541	0.77	0/727
13	V	0.52	0/541	0.79	1/727 (0.1%)
14	K	0.92	6/937 (0.6%)	1.00	11/1265 (0.9%)
14	W	0.93	6/937 (0.6%)	0.99	11/1265 (0.9%)
15	L	0.43	0/366	0.71	0/485
15	X	0.45	0/366	0.72	0/485
All	All	0.53	22/65058 (0.0%)	0.76	61/88054 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	2	0	2
2	5	0	1
5	B	0	1
5	N	0	1
All	All	0	5

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	W	112	GLN	CA-C	10.22	1.79	1.52
14	K	112	GLN	CA-C	10.06	1.79	1.52
14	W	113	THR	N-CA	9.11	1.64	1.46
14	K	113	THR	N-CA	9.06	1.64	1.46
14	W	112	GLN	N-CA	8.71	1.63	1.46

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	3	2	C	OP2-P-O3'	-22.50	55.69	105.20
3	3	3	G	O5'-P-OP2	22.31	137.47	110.70
3	6	3	G	O5'-P-OP2	21.67	136.70	110.70
3	6	2	C	OP2-P-O3'	-20.00	61.20	105.20
14	W	113	THR	N-CA-C	9.88	137.68	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	2	20	DC	Sidechain
2	2	26	DC	Sidechain
2	5	20	DC	Sidechain
5	B	486	TYR	Sidechain
5	N	486	TYR	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	1	141	0	81	11	0
1	4	141	0	81	12	0
2	2	380	0	218	36	0
2	5	380	0	218	33	0
3	3	212	0	110	14	0
3	6	212	0	110	10	0
4	A	11186	0	11266	1277	0
4	M	11186	0	11266	1263	0
5	B	8866	0	8898	968	0
5	N	8866	0	8898	1006	0
6	C	2101	0	2055	256	0
6	O	2101	0	2055	237	0
7	D	1427	0	1451	142	0
7	P	1427	0	1451	144	0
8	E	1752	0	1776	131	0
8	Q	1752	0	1776	124	0
9	F	705	0	730	82	0
9	R	705	0	730	80	0
10	G	1340	0	1357	154	0
10	S	1340	0	1357	164	0
11	H	1084	0	1057	122	0
11	T	1084	0	1057	122	0
12	I	944	0	899	112	0
12	U	944	0	899	113	0
13	J	532	0	542	98	0
13	V	532	0	542	110	0
14	K	919	0	929	113	0
14	W	919	0	929	99	0
15	L	364	0	386	41	0
15	X	364	0	386	41	0
16	A	1	0	0	0	0
16	M	1	0	0	0	0
17	A	2	0	0	0	0
17	B	1	0	0	0	0
17	C	1	0	0	0	0
17	I	2	0	0	0	0
17	J	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	L	1	0	0	0	0
17	M	2	0	0	0	0
17	N	1	0	0	0	0
17	O	1	0	0	0	0
17	U	2	0	0	0	0
17	V	1	0	0	0	0
17	X	1	0	0	0	0
All	All	63924	0	63510	6519	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 51.

The worst 5 of 6519 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)
14:W:112:GLN:CB	14:W:112:GLN:CG	1.76	1.59
14:K:112:GLN:CB	14:K:112:GLN:CG	1.75	1.56
14:K:112:GLN:CA	14:K:112:GLN:C	1.79	1.49
14:W:112:GLN:C	14:W:112:GLN:CA	1.79	1.47
4:A:855:THR:HG21	4:A:857:ARG:HE	1.09	1.18

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	1410/1733 (81%)	968 (69%)	288 (20%)	154 (11%)	0	8
4	M	1410/1733 (81%)	964 (68%)	291 (21%)	155 (11%)	0	8
5	B	1096/1224 (90%)	762 (70%)	222 (20%)	112 (10%)	0	9
5	N	1096/1224 (90%)	765 (70%)	217 (20%)	114 (10%)	0	9

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	C	264/318 (83%)	172 (65%)	64 (24%)	28 (11%)	0	8
6	O	264/318 (83%)	171 (65%)	63 (24%)	30 (11%)	0	7
7	D	173/221 (78%)	125 (72%)	27 (16%)	21 (12%)	0	6
7	P	173/221 (78%)	124 (72%)	32 (18%)	17 (10%)	0	10
8	E	212/215 (99%)	155 (73%)	42 (20%)	15 (7%)	1	17
8	Q	212/215 (99%)	156 (74%)	42 (20%)	14 (7%)	1	19
9	F	84/155 (54%)	67 (80%)	11 (13%)	6 (7%)	1	17
9	R	84/155 (54%)	67 (80%)	12 (14%)	5 (6%)	1	20
10	G	169/171 (99%)	125 (74%)	37 (22%)	7 (4%)	3	27
10	S	169/171 (99%)	132 (78%)	28 (17%)	9 (5%)	2	23
11	H	131/146 (90%)	87 (66%)	27 (21%)	17 (13%)	0	5
11	T	131/146 (90%)	83 (63%)	30 (23%)	18 (14%)	0	4
12	I	114/122 (93%)	83 (73%)	19 (17%)	12 (10%)	0	8
12	U	114/122 (93%)	81 (71%)	22 (19%)	11 (10%)	0	10
13	J	63/70 (90%)	35 (56%)	14 (22%)	14 (22%)	0	1
13	V	63/70 (90%)	36 (57%)	15 (24%)	12 (19%)	0	2
14	K	112/120 (93%)	86 (77%)	14 (12%)	12 (11%)	0	8
14	W	112/120 (93%)	86 (77%)	17 (15%)	9 (8%)	1	14
15	L	44/70 (63%)	16 (36%)	21 (48%)	7 (16%)	0	3
15	X	44/70 (63%)	17 (39%)	19 (43%)	8 (18%)	0	2
All	All	7744/9130 (85%)	5363 (69%)	1574 (20%)	807 (10%)	0	9

5 of 807 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	44	THR
4	A	48	ALA
4	A	57	ARG
4	A	62	ASP
4	A	65	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	
4	A	1244/1520 (82%)	1138 (92%)	106 (8%)	10	40
4	M	1244/1520 (82%)	1142 (92%)	102 (8%)	11	40
5	B	967/1061 (91%)	888 (92%)	79 (8%)	11	40
5	N	967/1061 (91%)	886 (92%)	81 (8%)	11	40
6	C	235/274 (86%)	216 (92%)	19 (8%)	11	41
6	O	235/274 (86%)	215 (92%)	20 (8%)	10	40
7	D	159/200 (80%)	136 (86%)	23 (14%)	3	19
7	P	159/200 (80%)	138 (87%)	21 (13%)	4	22
8	E	196/197 (100%)	191 (97%)	5 (3%)	46	69
8	Q	196/197 (100%)	191 (97%)	5 (3%)	46	69
9	F	77/137 (56%)	68 (88%)	9 (12%)	5	27
9	R	77/137 (56%)	69 (90%)	8 (10%)	7	30
10	G	152/152 (100%)	141 (93%)	11 (7%)	14	45
10	S	152/152 (100%)	140 (92%)	12 (8%)	12	42
11	H	119/128 (93%)	112 (94%)	7 (6%)	19	51
11	T	119/128 (93%)	112 (94%)	7 (6%)	19	51
12	I	110/116 (95%)	96 (87%)	14 (13%)	4	23
12	U	110/116 (95%)	95 (86%)	15 (14%)	3	22
13	J	60/65 (92%)	53 (88%)	7 (12%)	5	27
13	V	60/65 (92%)	54 (90%)	6 (10%)	7	32
14	K	99/102 (97%)	89 (90%)	10 (10%)	7	32
14	W	99/102 (97%)	88 (89%)	11 (11%)	6	29
15	L	40/57 (70%)	36 (90%)	4 (10%)	7	32
15	X	40/57 (70%)	36 (90%)	4 (10%)	7	32
All	All	6916/8018 (86%)	6330 (92%)	586 (8%)	10	40

5 of 586 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	I	78	CYS
4	M	470	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
10	S	126	ASN
12	I	110	PHE
4	M	37	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 149 such sidechains are listed below:

Mol	Chain	Res	Type
12	I	60	GLN
4	M	631	HIS
8	Q	147	HIS
13	J	64	ASN
4	M	68	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	3	9/11 (81%)	1 (11%)	0
3	6	9/11 (81%)	1 (11%)	0
All	All	18/22 (81%)	2 (11%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	3	3	G
3	6	3	G

There are no RNA pucker outliers to report.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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
2	BRU	5	22	3,2	15,21,22	1.87	3 (20%)	17,30,33	4.29	6 (35%)
2	BRU	2	22	3,2	15,21,22	1.74	3 (20%)	17,30,33	4.18	5 (29%)
2	TT	2	18	2	40,43,44	4.61	8 (20%)	59,69,72	2.09	12 (20%)
2	TT	5	18	2	40,43,44	4.60	7 (17%)	59,69,72	2.09	13 (22%)

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
2	BRU	5	22	3,2	-	2/4/21/22	0/2/2/2
2	BRU	2	22	3,2	-	2/4/21/22	0/2/2/2
2	TT	2	18	2	-	10/18/105/106	0/5/6/6
2	TT	5	18	2	-	10/18/105/106	0/5/6/6

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	5	18	TT	C5-C6	-20.59	1.31	1.55
2	2	18	TT	C5-C6	-20.44	1.31	1.55
2	2	18	TT	C5T-C6T	-19.00	1.33	1.55
2	5	18	TT	C5T-C6T	-18.67	1.33	1.55
2	5	22	BRU	C4-C5	5.39	1.45	1.38

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	5	22	BRU	C4-N3-C2	14.83	127.66	115.14
2	2	22	BRU	C4-N3-C2	14.74	127.59	115.14
2	5	18	TT	C5T-C5-C6	7.53	97.75	88.38
2	5	22	BRU	C5-C4-N3	-7.21	115.01	123.64
2	2	22	BRU	C5-C4-N3	-7.12	115.11	123.64

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	2	18	TT	O4R-C4R-C5R-O5R
2	2	18	TT	C3R-C4R-C5R-O5R

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	5	18	TT	C3R-C4R-C5R-O5R
2	5	18	TT	O4R-C4R-C5R-O5R
2	2	18	TT	O4'-C4'-C5'-O5'

There are no ring outliers.

4 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	5	22	BRU	3	0
2	2	22	BRU	4	0
2	2	18	TT	4	0
2	5	18	TT	4	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 18 ligands modelled in this entry, 18 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
5	B	2
5	N	2
6	C	1
4	A	1
2	5	1
6	O	1
2	2	1
9	R	1
9	F	1
4	M	1

The worst 5 of 12 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	5	18:TT	O3'	20:DC	P	7.05
1	2	18:TT	O3'	20:DC	P	7.02
1	O	2:SER	C	3:GLU	N	4.13
1	C	2:SER	C	3:GLU	N	4.09
1	B	18:PHE	C	19:GLU	N	3.79



## 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	1	7/14 (50%)	-0.74	0 100 100	123, 130, 149, 149	0
1	4	7/14 (50%)	-0.68	0 100 100	120, 128, 147, 148	0
2	2	16/25 (64%)	-0.67	0 100 100	84, 130, 151, 155	0
2	5	16/25 (64%)	-0.57	0 100 100	87, 132, 150, 153	0
3	3	10/11 (90%)	-0.48	0 100 100	95, 100, 152, 155	0
3	6	10/11 (90%)	-0.44	0 100 100	95, 102, 153, 156	0
4	A	1421/1733 (81%)	-0.09	8 (0%) 89 85	22, 88, 163, 200	0
4	M	1421/1733 (81%)	-0.07	5 (0%) 92 89	20, 88, 163, 200	0
5	B	1115/1224 (91%)	-0.04	8 (0%) 87 83	24, 101, 175, 200	0
5	N	1115/1224 (91%)	-0.03	7 (0%) 89 85	23, 101, 174, 200	0
6	C	267/318 (83%)	-0.14	0 100 100	49, 88, 147, 173	0
6	O	267/318 (83%)	-0.07	1 (0%) 92 89	52, 87, 148, 170	0
7	D	177/221 (80%)	-0.06	0 100 100	72, 121, 166, 183	0
7	P	177/221 (80%)	-0.06	2 (1%) 80 74	71, 124, 166, 182	0
8	E	214/215 (99%)	-0.15	1 (0%) 91 87	60, 145, 193, 197	0
8	Q	214/215 (99%)	-0.14	1 (0%) 91 87	58, 145, 194, 197	0
9	F	87/155 (56%)	-0.07	0 100 100	31, 62, 108, 140	0
9	R	87/155 (56%)	-0.05	0 100 100	31, 63, 109, 138	0
10	G	171/171 (100%)	-0.10	0 100 100	64, 91, 136, 146	0
10	S	171/171 (100%)	-0.07	1 (0%) 89 85	65, 93, 136, 143	0
11	H	135/146 (92%)	-0.03	6 (4%) 34 29	101, 145, 182, 192	0
11	T	135/146 (92%)	0.00	2 (1%) 73 66	99, 146, 181, 191	0
12	I	116/122 (95%)	-0.10	1 (0%) 84 79	82, 139, 170, 195	0
12	U	116/122 (95%)	-0.14	1 (0%) 84 79	81, 138, 170, 194	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	J	65/70 (92%)	-0.14	1 (1%) 73 66	52, 83, 128, 133	0
13	V	65/70 (92%)	-0.10	1 (1%) 73 66	46, 81, 130, 136	0
14	K	114/120 (95%)	-0.00	2 (1%) 68 61	48, 92, 120, 170	0
14	W	114/120 (95%)	-0.05	2 (1%) 68 61	47, 92, 118, 167	0
15	L	46/70 (65%)	0.06	1 (2%) 62 54	86, 155, 179, 186	0
15	X	46/70 (65%)	0.06	0 100 100	84, 156, 179, 185	0
All	All	7922/9230 (85%)	-0.07	51 (0%) 89 85	20, 99, 173, 200	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
14	K	114	LEU	8.6
14	K	113	THR	8.0
14	W	114	LEU	6.6
14	W	113	THR	6.2
4	M	1092	LYS	5.7

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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, 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	B-factors(Å <sup>2</sup> )	Q<0.9
2	BRU	2	22	20/21	0.90	0.17	69,76,80,84	0
2	BRU	5	22	20/21	0.92	0.14	77,84,87,89	0
2	TT	2	18	38/39	0.94	0.21	93,106,123,125	0
2	TT	5	18	38/39	0.94	0.16	95,108,126,127	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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, 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	B-factors( $\text{\AA}^2$ )	Q<0.9
17	ZN	I	2457	1/1	0.97	0.24	199,199,199,199	0
17	ZN	B	2457	1/1	0.97	0.28	53,53,53,53	0
17	ZN	M	2457	1/1	0.98	0.27	87,87,87,87	0
17	ZN	X	2457	1/1	0.98	0.25	120,120,120,120	0
17	ZN	U	2458	1/1	0.98	0.19	178,178,178,178	0
17	ZN	L	2457	1/1	0.98	0.24	116,116,116,116	0
17	ZN	J	2457	1/1	0.99	0.25	70,70,70,70	0
16	MG	A	2457	1/1	0.99	0.19	50,50,50,50	0
17	ZN	I	2458	1/1	0.99	0.29	94,94,94,94	0
17	ZN	V	2457	1/1	0.99	0.24	67,67,67,67	0
17	ZN	M	2458	1/1	0.99	0.24	51,51,51,51	0
16	MG	M	2459	1/1	0.99	0.24	34,34,34,34	0
17	ZN	C	2457	1/1	0.99	0.30	48,48,48,48	0
17	ZN	U	2457	1/1	0.99	0.30	91,91,91,91	0
17	ZN	O	2457	1/1	0.99	0.30	39,39,39,39	0
17	ZN	A	2471	1/1	0.99	0.24	90,90,90,90	0
17	ZN	A	2472	1/1	1.00	0.22	52,52,52,52	0
17	ZN	N	2457	1/1	1.00	0.28	54,54,54,54	0

## 6.5 Other polymers

There are no such residues in this entry.