



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 06:04 am GMT

PDB ID : 1R9T  
Title : RNA POLYMERASE II STRAND SEPARATED ELONGATION COM-  
PLEX, MISMATCHED NUCLEOTIDE  
Authors : Westover, K.D.; Bushnell, D.A.; Kornberg, R.D.  
Deposited on : 2003-10-30  
Resolution : 3.50 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](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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

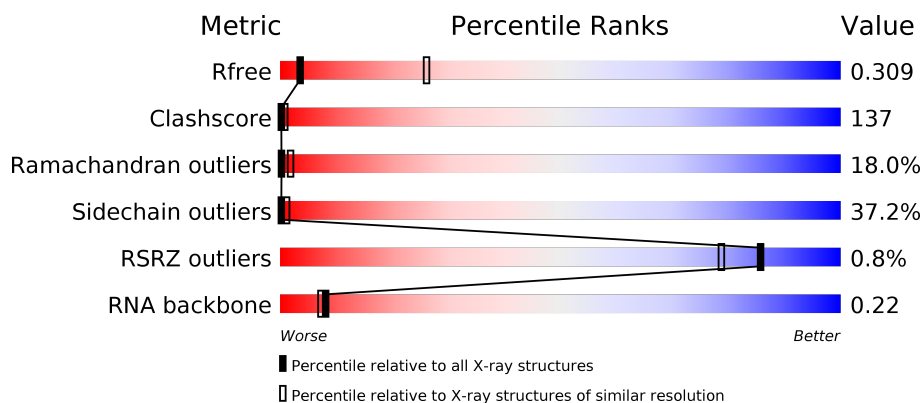
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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))
$R_{free}$	100719	1195 (3.60-3.40)
Clashscore	112137	1322 (3.60-3.40)
Ramachandran outliers	110173	1283 (3.60-3.40)
Sidechain outliers	110143	1284 (3.60-3.40)
RSRZ outliers	101464	1226 (3.60-3.40)
RNA backbone	2435	1024 (4.10-2.86)

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	R	10	 10% 50% 40%
2	T	28	 18% 50% 43% 7%
3	N	14	 36% 71% 29%
4	A	1733	 11% 35% 32% 20%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	B	1224	
6	C	318	
7	E	215	
8	F	155	
9	H	146	
10	I	122	
11	J	70	
12	K	120	
13	L	70	

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
15	MG	A	2002	-	-	-	X
16	ATP	B	1308	-	-	-	X

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 29248 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 RNA chain called RNA strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	10	Total	C	N	O	P	0	0	0
			217	98	45	65	9			

- Molecule 2 is a DNA chain called DNA template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	28	Total	C	N	O	P	0	0	0
			566	271	104	164	27			

- Molecule 3 is a DNA chain called DNA nontemplate strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	14	Total	C	N	O	P	0	0	0
			284	137	49	85	13			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1395	Total	C	N	O	S	0	0	0
			10969	6917	1923	2068	61			

- Molecule 5 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1106	Total	C	N	O	S	0	0	0
			8792	5568	1538	1631	55			

- Molecule 6 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

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

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	84	Total	C	N	O	S	0	0	0
			679	434	115	127	3			

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 10 is a protein called DNA-directed RNA polymerase II 14.2 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	I	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide.

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

- Molecule 12 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

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

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

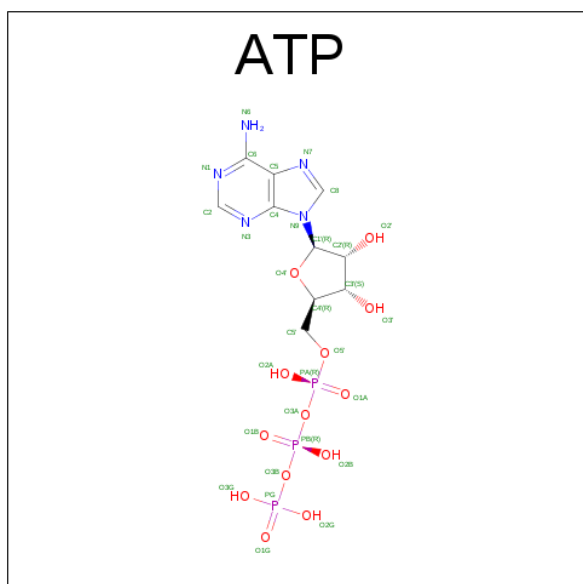
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	J	1	Total	Zn	0	0
			1	1		
14	B	1	Total	Zn	0	0
			1	1		
14	I	2	Total	Zn	0	0
			2	2		
14	C	1	Total	Zn	0	0
			1	1		
14	A	2	Total	Zn	0	0
			2	2		
14	L	1	Total	Zn	0	0
			1	1		

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
15	A	2	Total	Mg	0	0
			2	2		

- Molecule 16 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).

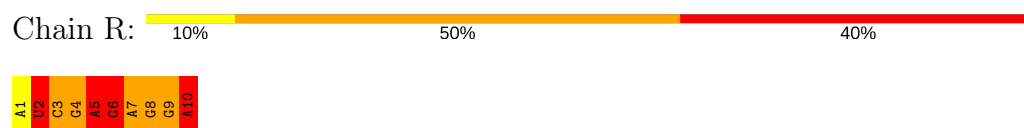


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
16	B	1	Total	C	N	O	P	0	0
			31	10	5	13	3		

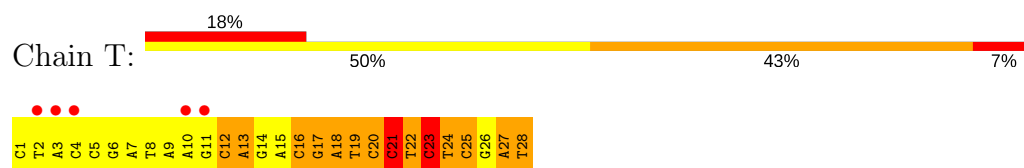
### 3 Residue-property plots

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.

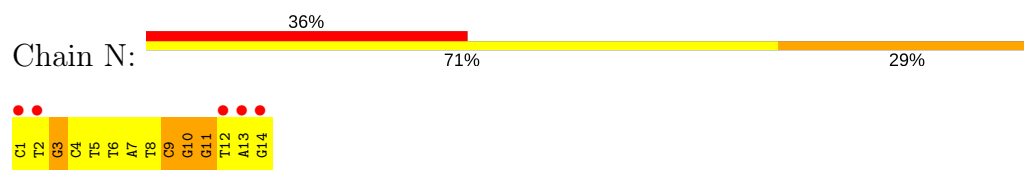
- Molecule 1: RNA strand



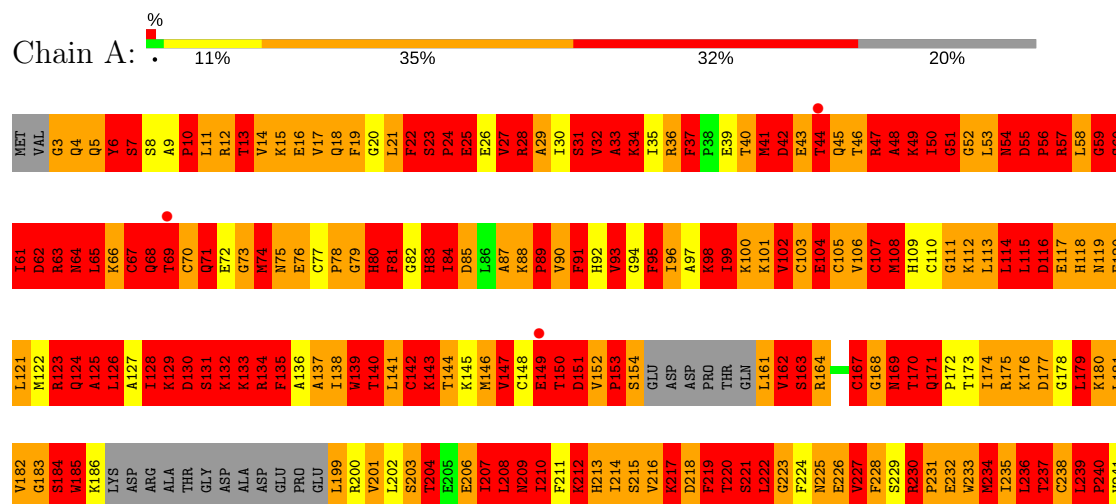
- Molecule 2: DNA template strand



- Molecule 3: DNA nontemplate strand



- Molecule 4: DNA-directed RNA polymerase II largest subunit

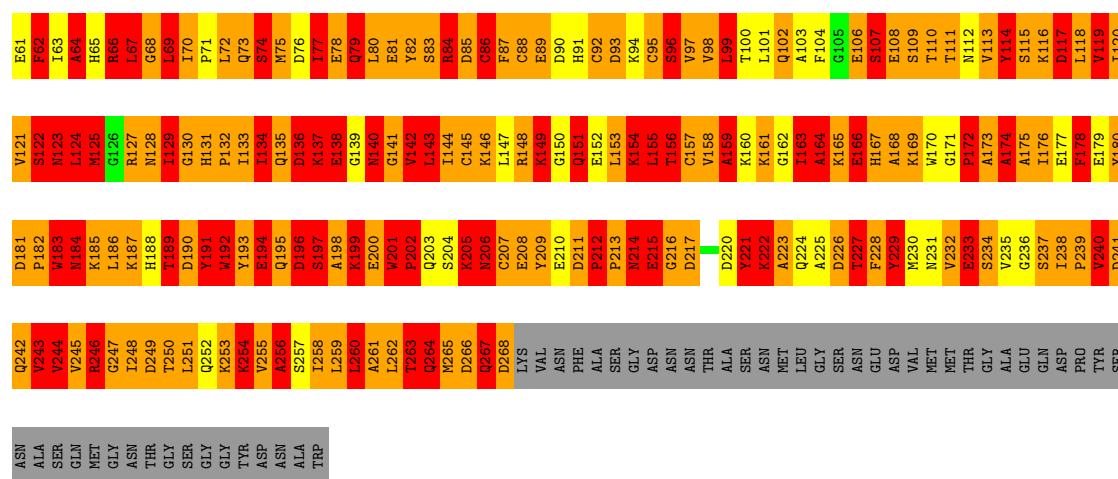




T1142	L1022	R862	L902	V842	T782	L722	R662	D602	F482	G422	D362	T302	P242
L1143	R1023	R963	R903	K843	T783	N723	S663	M603	D483	D423	Q363	Y303	P243
K1144	PHE S1024	I964	T904	A844	L784	A724	T664	G604	G484	I424	Q364	M304	P244
S1145	R1025	Q965	T905	L845	P785	A725	G665	M605	D485	Q425	G365	D305	P245
V1146	L1026	N966	H906	E846	H786	R726	I666	L606	E486	L426	P366	N306	V246
T1147	A1027	A967	T907	D847	F787	D727	G667	I607	M487	Q427	P367	D307	R247
L1148	T1028	Q968	L908	I848	S788	K728	D668	I608	K488	Y428	K368	I308	P248
A1149	R1029	Q969	D909	M849	K789	A729	T669	D609	L489	G429	S369	A309	S249
S1150	R1030	T970	P910	V850	D790	G730	I670	G610	L490	W430	I370	I309	I250
E1151	V1031	F971	S911	H851	D791	R731	A671	Q611	K431	W431	A371	G310	S251
Y1152	H972	H972	L912	H852	T792	L732	D672	I612	P492	K432	K372	P312	F252
Y1153	Q1033	I973	L913	D853	T793	A733	G673	I613	Q493	E433	T373	Q313	N253
Y1154	E914	D974	E914	N854	F794	E734	P674	F614	S494	E434	L374	A314	N254
D1155	Y1035	H975	S915	T855	E795	V735	T675	G615	E495	H435	T375	L315	S255
P1156	R1036	T976	G916	T856	S796	N736	M676	V616	E496	I436	Y376	Q316	Q256
I1157	L1037	K977	S917	R857	K797	L737	R677	V617	T497	K437	P377	Q317	R257
P1158	T1038	P978	E918	N858	G798	D738	E878	E618	R498	D438	E378	S318	G258
R1159	K1039	S979	I919	S859	F799	R739	I679	K619	A999	N439	V379	G319	E259
S1160	R1040	D980	L920	L860	V800	L740	T680	K620	E500	D440	V380	R320	D260
T1161	Q1041	L981	G921	G861	E801	N741	E681	T621	L501	P441	T381	P321	D261
V1162	F1042	T982	D922	N862	N902	N742	T682	V622	S502	W442	L262	V322	L262
E1163	I1043	I983	L923	V863	S803	V743	I683	G623	Q503	L443	Y383	K323	T263
P1164	W1044	K984	K924	I864	A804	K744	A684	S624	L504	F444	N384	S324	F264
I1165	V1045	D985	L925	Q865	L805	Q745	E585	S625	C505	N445	I385	I325	K265
D1166	L1046	I986	Q926	F866	R806	M746	A686	N626	A506	K446	D386	L266	L266
E1167	S1047	V987	V927	I867	G807	V747	K687	G627	V507	Q447	E387	A327	A267
E1168	A1048	L988	L928	Y868	L808	M748	R688	G628	P508	F448	L388	R328	D268
I1169	T1049	G989	L929	G869	T809	A749	K689	L629	F508	S449	L389	L329	L269
N1170	E1050	V990	D930	E870	P810	G750	V690	K630	L450	I460	Q390	K330	L270
Q1171	A1051	K991	E931	D871	Q811	S751	L691	H631	I511	H451	L391	G331	K271
L1172	L1052	D992	E932	G872	E812	K752	D692	V632	V612	K452	V392	K332	A272
H1173	F1053	L993	Y933	K873	F813	G753	V693	V633	S813	V453	N273	E333	N273
F1174	L1054	Q994	K934	D874	F814	S754	T694	T634	P514	S454	N394	G334	I274
S1175	S1115	E995	Q935	A875	F815	F755	E695	R635	Q515	M455	G395	R335	S275
L1176	L1055	N996	L936	A876	H816	I756	E696	E636	S516	M456	P396	L336	L276
LEU	V1057	L997	V937	H877	A817	N757	A697	K637	N517	A457	N397	R337	E277
ASP	V1058	K998	K938	I878	N818	I758	Q698	G638	K518	H458	E398	G338	T278
GLU	H1059	V999	D939	E879	G819	A759	A699	P639	P519	R459	H399	N339	L279
GLU	L1120	L1000	R940	K880	G820	Q760	T700	Q640	C520	V460	P400	L340	E280
ALA	E1121	R1001	K941	Q881	R821	M761	L701	V641	M521	K461	G401	K341	N281
GLU	P1122	G1002	F942	S882	E822	S762	L702	C642	G522	V462	A402	G342	N282
GLN	G1123	M1003	L943	L883	G823	A763	T703	A643	I523	P463	K403	K343	G283
SER	H1124	E1004	R944	D884	L824	C764	A704	K644	M524	V464	Y404	K344	A284
PHE	A1125	E1005	E945	T885	T825	V765	K705	L645	Q525	V465	V405	V345	P285
ASP	A1126	V1066	V946	I886	D826	G766	H706	F646	D526	S466	I406	D346	H286
Q1187	D1127	I1007	F947	G887	T827	Q767	G707	G647	T527	T467	R407	F347	H287
Q1188	Q1128	Q1008	V948	G888	A828	Q768	M708	N648	L528	F468	D408	S348	A288
S1189	E1129	N1009	D949	S889	V829	S769	T709	I649	C529	R469	S409	A349	L289
P1190	Q1070	A1010	G950	D890	K830	V770	L710	Q650	G530	L470	G410	R350	E290
W1191	S1071	Q1011	E951	A891	T831	E771	R711	K651	I631	M471	D411	T351	E291
L1192	L1072	R1012	A952	A892	A832	G772	E712	V652	R532	L472	K412	V352	A292
L1193	G1073	D1013	N953	F893	E833	K773	S713	V653	K533	L473	I413	I353	E293
R1194	E1074	A1014	W954	E894	T834	R774	F714	N654	L534	V474	D414	S354	S294
L1195	P1075	K995	P955	K895	G835	I775	E715	F655	T595	T475	L415	G355	L295
E1196	A1076	R896	L956	R896	Y836	A776	D716	V656	L536	S476	R416	D356	L296
L1197	L1077	L1017	P957	Y837	T837	F777	N717	L657	R337	P477	Y417	P357	Q297
D1198	Q1078	F1018	V958	R898	Q838	G778	V718	L658	D338	V478	S418	N358	F298
R1199	M1079	C1019	N959	V899	R339	F779	V719	H659	T539	N479	K419	L359	S299
H1140	T1080	D1020	I960	D900	R840	Q760	R720	P660	F540	A480	R420	E360	V300
A1201	L1081	L1021	R961	L901	L941	D781	F721	G661	I541	D481	A421	L361	A301

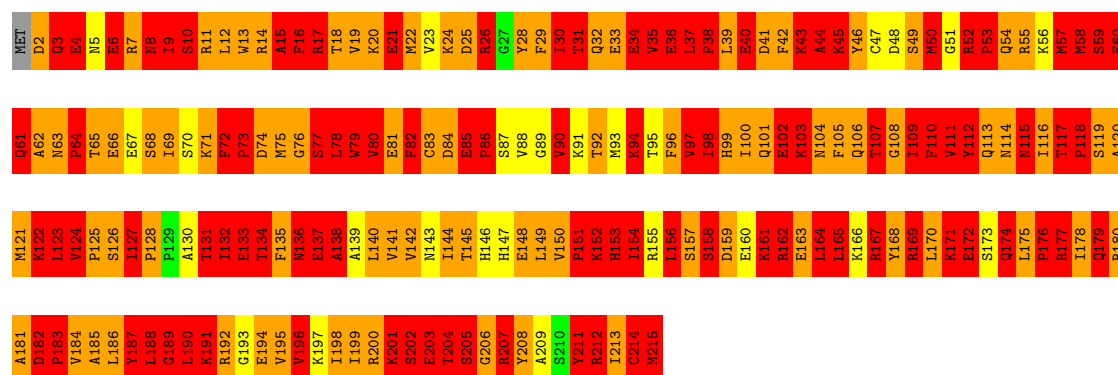






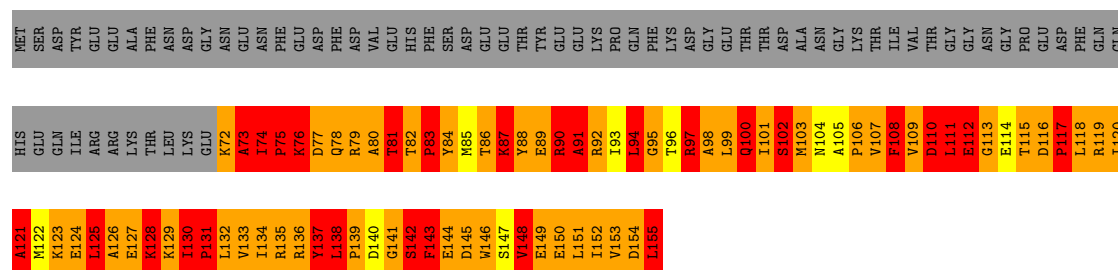
• Molecule 7: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

Chain E: 12% 39% 47%



• Molecule 8: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

Chain F: 6% 30% 19% 46%



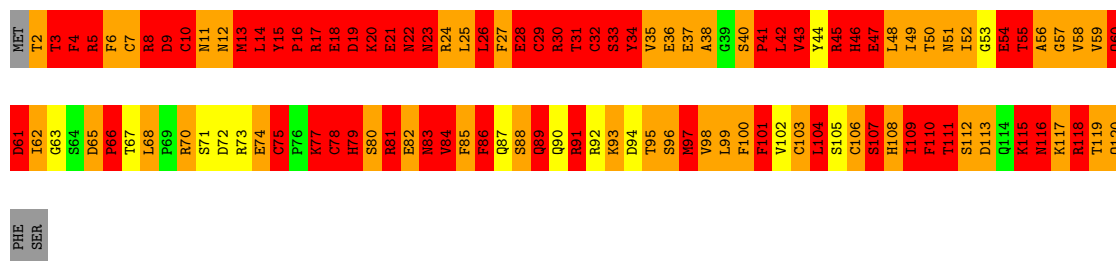
• Molecule 9: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

Chain H: 10% 42% 38% 9%

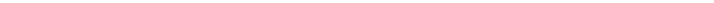


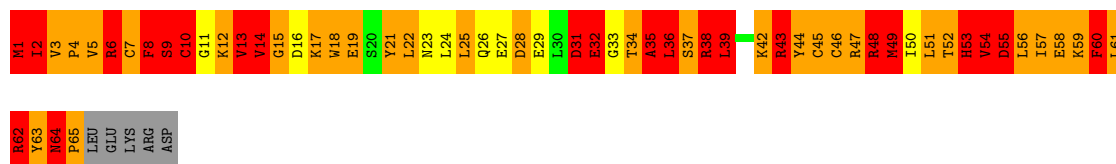
- Molecule 10: DNA-directed RNA polymerase II 14.2 kDa polypeptide

Chain I:  11% 37% 46%




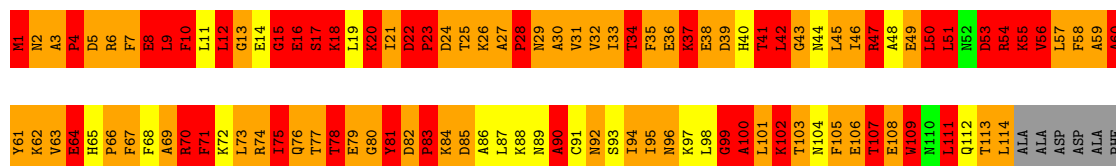
- Molecule 11: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide

Chain J: 



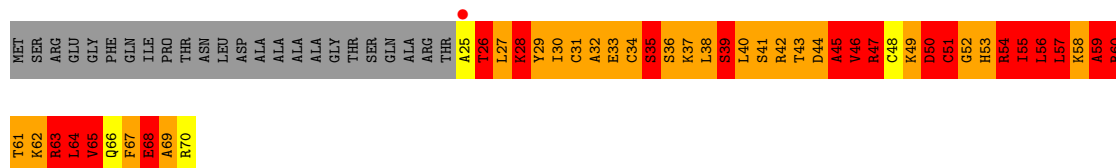
- Molecule 12: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K:  16% 44% 33% 5%



- Molecule 13: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide

Chain L: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	168.85Å 222.96Å 193.60Å 90.00° 101.17° 90.00°	Depositor
Resolution (Å)	40.00 – 3.50 39.98 – 3.50	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-3.50) 93.7 (39.98-3.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.26 (at 3.48Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.230 , 0.317 0.226 , 0.309	Depositor DCC
$R_{free}$ test set	8299 reflections (10.03%)	DCC
Wilson B-factor (Å <sup>2</sup> )	80.5	Xtriage
Anisotropy	0.219	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.20 , 80.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	29248	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% 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: ZN, MG, ATP

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	R	3.34	30/244 (12.3%)	3.47	35/380 (9.2%)
2	T	2.72	42/633 (6.6%)	2.95	63/971 (6.5%)
3	N	1.33	4/316 (1.3%)	1.10	6/484 (1.2%)
4	A	3.73	1529/11163 (13.7%)	3.57	2038/15091 (13.5%)
5	B	3.68	1206/8963 (13.5%)	3.62	1655/12086 (13.7%)
6	C	3.67	281/2133 (13.2%)	3.42	376/2891 (13.0%)
7	E	3.85	252/1788 (14.1%)	3.52	348/2406 (14.5%)
8	F	3.81	105/691 (15.2%)	3.78	136/933 (14.6%)
9	H	4.17	178/1086 (16.4%)	3.65	243/1470 (16.5%)
10	I	3.97	150/989 (15.2%)	3.89	212/1331 (15.9%)
11	J	3.68	72/541 (13.3%)	3.56	89/727 (12.2%)
12	K	3.73	149/937 (15.9%)	3.38	150/1265 (11.9%)
13	L	4.22	68/365 (18.6%)	3.67	86/485 (17.7%)
All	All	3.71	4066/29849 (13.6%)	3.55	5437/40520 (13.4%)

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	R	0	1
2	T	1	2
4	A	2	121
5	B	4	114
6	C	1	11
7	E	0	17
8	F	0	5
9	H	0	16
10	I	1	18
11	J	0	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
12	K	0	9
13	L	0	5
All	All	9	321

The worst 5 of 4066 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1110	ASN	CB-CG	39.65	2.42	1.51
4	A	820	GLY	C-O	-37.66	0.63	1.23
7	E	117	THR	CA-CB	34.64	2.43	1.53
4	A	437	MET	SD-CE	30.90	3.50	1.77
4	A	322	VAL	CA-CB	-27.33	0.97	1.54

The worst 5 of 5437 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	941	LEU	CB-CG-CD2	37.01	173.92	111.00
4	A	337	ARG	NE-CZ-NH2	-34.93	102.83	120.30
4	A	980	ASP	CB-CG-OD2	33.77	148.69	118.30
5	B	466	TRP	CA-C-N	-31.53	53.14	116.20
4	A	337	ARG	NE-CZ-NH1	31.48	136.04	120.30

5 of 9 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	T	21	DC	C3'
4	A	317	LYS	CA
4	A	324	SER	CA
5	B	636	PRO	CA
5	B	637	LEU	CA

5 of 321 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	3	GLY	Mainchain
4	A	6	TYR	Peptide
1	R	10	A	Sidechain
2	T	21	DC	Sidechain
2	T	23	DC	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	R	217	0	110	34	0
2	T	566	0	314	82	4
3	N	284	0	162	44	0
4	A	10969	0	11061	3443	0
5	B	8792	0	8821	2369	0
6	C	2095	0	2051	490	0
7	E	1752	0	1776	523	0
8	F	679	0	701	194	0
9	H	1068	0	1040	377	0
10	I	971	0	929	310	0
11	J	532	0	543	160	0
12	K	919	0	929	259	0
13	L	363	0	387	98	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	A	2	0	0	0	0
16	B	31	0	12	7	0
All	All	29248	0	28836	7970	4

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 137.

The worst 5 of 7970 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:N:12:DT:P	7:E:117:THR:HG21	1.32	1.67
4:A:128:ILE:CB	4:A:128:ILE:CA	1.74	1.66
5:B:422:LYS:CB	5:B:422:LYS:CG	1.74	1.65
5:B:866:TYR:CB	5:B:866:TYR:CG	1.80	1.65
4:A:37:PHE:CG	4:A:37:PHE:CB	1.74	1.65

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:T:1:DC:C4'	2:T:1:DC:C4'[2_656]	1.63	0.57
2:T:1:DC:C5'	2:T:1:DC:O4'[2_656]	1.69	0.51
2:T:1:DC:O4'	2:T:1:DC:O4'[2_656]	1.78	0.42
2:T:1:DC:C4'	2:T:1:DC:O4'[2_656]	1.80	0.40

## 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	
4	A	1383/1733 (80%)	816 (59%)	306 (22%)	261 (19%)	0	2
5	B	1088/1224 (89%)	708 (65%)	199 (18%)	181 (17%)	0	2
6	C	264/318 (83%)	170 (64%)	58 (22%)	36 (14%)	0	4
7	E	212/215 (99%)	119 (56%)	47 (22%)	46 (22%)	0	1
8	F	82/155 (53%)	45 (55%)	22 (27%)	15 (18%)	0	2
9	H	129/146 (88%)	75 (58%)	28 (22%)	26 (20%)	0	1
10	I	117/122 (96%)	60 (51%)	32 (27%)	25 (21%)	0	1
11	J	63/70 (90%)	41 (65%)	10 (16%)	12 (19%)	0	2
12	K	112/120 (93%)	78 (70%)	18 (16%)	16 (14%)	0	3
13	L	44/70 (63%)	21 (48%)	11 (25%)	12 (27%)	0	0
All	All	3494/4173 (84%)	2133 (61%)	731 (21%)	630 (18%)	0	2

5 of 630 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	34	LYS
4	A	44	THR
4	A	48	ALA
4	A	51	GLY

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	A	55	ASP

### 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	1218/1520 (80%)	759 (62%)	459 (38%)	0	1
5	B	960/1061 (90%)	614 (64%)	346 (36%)	0	1
6	C	234/274 (85%)	151 (64%)	83 (36%)	0	1
7	E	196/197 (100%)	116 (59%)	80 (41%)	0	1
8	F	74/137 (54%)	47 (64%)	27 (36%)	0	1
9	H	117/128 (91%)	71 (61%)	46 (39%)	0	1
10	I	113/116 (97%)	70 (62%)	43 (38%)	0	1
11	J	60/65 (92%)	38 (63%)	22 (37%)	0	1
12	K	99/102 (97%)	63 (64%)	36 (36%)	0	1
13	L	40/57 (70%)	24 (60%)	16 (40%)	0	1
All	All	3111/3657 (85%)	1953 (63%)	1158 (37%)	0	1

5 of 1158 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	257	LYS
5	B	682	SER
10	I	79	HIS
5	B	319	GLU
5	B	436	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 93 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	1364	ASN
5	B	383	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
10	I	60	GLN
4	A	1390	ASN
5	B	115	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	9/10 (90%)	4 (44%)	0

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	2	U
1	R	5	A
1	R	6	G
1	R	10	A

There are no RNA pucker outliers to report.

## 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 11 ligands modelled in this entry, 10 are monoatomic - leaving 1 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
16	ATP	B	1308	-	27,33,33	2.26	7 (25%)	25,52,52	3.38	6 (24%)

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
16	ATP	B	1308	-	-	0/18/38/38	0/3/3/3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	B	1308	ATP	O5'-C5'	-5.75	1.22	1.44
16	B	1308	ATP	PA-O5'	-4.46	1.40	1.59
16	B	1308	ATP	C5'-C4'	-2.76	1.42	1.51
16	B	1308	ATP	C2-N1	2.67	1.38	1.33
16	B	1308	ATP	C2-N3	3.97	1.38	1.32

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	B	1308	ATP	N3-C2-N1	-10.09	120.07	128.86
16	B	1308	ATP	C5'-C4'-C3'	-5.23	95.35	115.29
16	B	1308	ATP	C1'-N9-C4	-3.26	120.99	126.64
16	B	1308	ATP	O4'-C4'-C5'	-2.93	99.50	109.40
16	B	1308	ATP	O4'-C4'-C3'	4.02	113.16	105.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	B	1308	ATP	7	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

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

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	R	10/10 (100%)	0.36	0 <span>100</span> <span>100</span>	70, 83, 159, 164	0
2	T	28/28 (100%)	0.90	5 (17%) <span>2</span> <span>2</span>	84, 188, 200, 200	0
3	N	14/14 (100%)	1.69	5 (35%) <span>0</span> <span>0</span>	185, 198, 200, 200	0
4	A	1395/1733 (80%)	-0.71	11 (0%) <span>86</span> <span>79</span>	1, 51, 137, 186	0
5	B	1106/1224 (90%)	-0.73	8 (0%) <span>87</span> <span>82</span>	1, 43, 119, 194	0
6	C	266/318 (83%)	-0.83	0 <span>100</span> <span>100</span>	6, 44, 95, 151	0
7	E	214/215 (99%)	-0.54	0 <span>100</span> <span>100</span>	13, 79, 141, 165	0
8	F	84/155 (54%)	-0.62	0 <span>100</span> <span>100</span>	17, 58, 105, 114	0
9	H	133/146 (91%)	-0.60	0 <span>100</span> <span>100</span>	19, 74, 132, 154	0
10	I	119/122 (97%)	-0.72	0 <span>100</span> <span>100</span>	4, 56, 105, 146	0
11	J	65/70 (92%)	-0.81	0 <span>100</span> <span>100</span>	11, 40, 92, 116	0
12	K	114/120 (95%)	-0.74	0 <span>100</span> <span>100</span>	8, 43, 88, 131	0
13	L	46/70 (65%)	-0.44	1 (2%) <span>62</span> <span>53</span>	17, 84, 143, 163	0
All	All	3594/4225 (85%)	-0.68	30 (0%) <span>86</span> <span>79</span>	1, 50, 133, 200	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	316	GLN	5.4
4	A	255	SER	4.0
4	A	44	THR	3.8
3	N	14	DG	3.6
5	B	866	TYR	3.5

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
15	MG	A	2002	1/1	0.71	0.49	11.34	63,63,63,63	0
16	ATP	B	1308	31/31	0.75	0.35	6.10	64,71,77,78	0
14	ZN	A	1735	1/1	0.98	0.04	-1.42	75,75,75,75	0
14	ZN	I	203	1/1	0.98	0.05	-1.61	98,98,98,98	0
14	ZN	I	204	1/1	0.99	0.05	-1.69	31,31,31,31	0
14	ZN	J	101	1/1	0.97	0.10	-1.79	39,39,39,39	0
14	ZN	C	319	1/1	0.99	0.02	-1.89	31,31,31,31	0
14	ZN	B	1307	1/1	0.99	0.05	-2.32	80,80,80,80	0
14	ZN	A	1734	1/1	0.80	0.05	-2.73	115,115,115,115	0
14	ZN	L	105	1/1	0.98	0.04	-2.77	76,76,76,76	0
15	MG	A	2001	1/1	0.95	0.56	-	55,55,55,55	0

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

There are no such residues in this entry.