



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 28, 2020 – 09:29 pm BST

PDB ID : 2B8K  
Title : 12-subunit RNA Polymerase II  
Authors : Meyer, P.A.; Ye, P.; Zhang, M.; Suh, M.H.; Fu, J.  
Deposited on : 2005-10-07  
Resolution : 4.15 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  
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

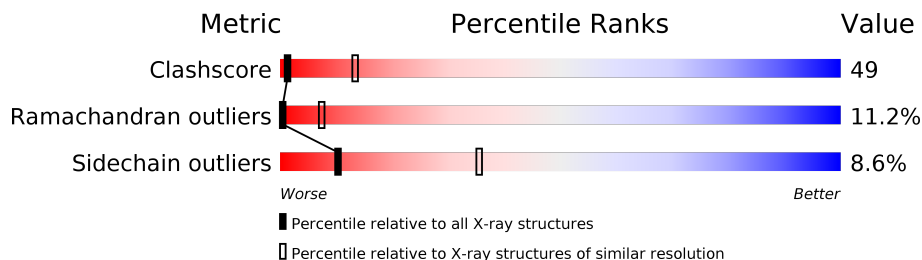
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.15 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))
Clashscore	141614	1028 (4.52-3.80)
Ramachandran outliers	138981	1005 (4.54-3.78)
Sidechain outliers	138945	1024 (4.54-3.76)

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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	1733	27% 43% 10% • 18%
2	B	1224	29% 48% 12% • 9%
3	C	318	23% 47% 12% • 16%
4	D	221	30% 42% 10% 18%
5	E	215	40% 53% 6%
6	F	155	17% 31% 6% • 46%
7	G	215	30% 42% 7% 20%
8	H	146	32% 49% 10% 9%
9	I	122	40% 43% 13% • •

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
10	J	70	
11	K	120	
12	L	70	

## 2 Entry composition [i](#)

There are 13 unique types of molecules in this entry. The entry contains 31040 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-directed RNA polymerase II largest subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1416	11140	7021	1946	2111	62	0	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1114	8800	5573	1540	1633	54	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

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

- Molecule 4 is a protein called DNA-directed RNA polymerase II 32 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	182	1373	851	243	277	2	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

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

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II 19 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	171	1339	861	222	248	8	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	172	SER	-	EXPRESSION TAG	UNP P34087
G	173	HIS	-	EXPRESSION TAG	UNP P34087
G	174	GLU	-	EXPRESSION TAG	UNP P34087
G	175	LYS	-	EXPRESSION TAG	UNP P34087
G	176	ARG	-	EXPRESSION TAG	UNP P34087
G	177	ARG	-	EXPRESSION TAG	UNP P34087
G	178	TRP	-	EXPRESSION TAG	UNP P34087
G	179	LYS	-	EXPRESSION TAG	UNP P34087
G	180	LYS	-	EXPRESSION TAG	UNP P34087
G	181	ASN	-	EXPRESSION TAG	UNP P34087
G	182	PHE	-	EXPRESSION TAG	UNP P34087
G	183	ILE	-	EXPRESSION TAG	UNP P34087
G	184	ALA	-	EXPRESSION TAG	UNP P34087
G	185	VAL	-	EXPRESSION TAG	UNP P34087
G	186	SER	-	EXPRESSION TAG	UNP P34087
G	187	ALA	-	EXPRESSION TAG	UNP P34087
G	188	ALA	-	EXPRESSION TAG	UNP P34087
G	189	ASN	-	EXPRESSION TAG	UNP P34087
G	190	ARG	-	EXPRESSION TAG	UNP P34087
G	191	PHE	-	EXPRESSION TAG	UNP P34087
G	192	LYS	-	EXPRESSION TAG	UNP P34087
G	193	LYS	-	EXPRESSION TAG	UNP P34087
G	194	ILE	-	EXPRESSION TAG	UNP P34087
G	195	SER	-	EXPRESSION TAG	UNP P34087
G	196	SER	-	EXPRESSION TAG	UNP P34087
G	197	SER	-	EXPRESSION TAG	UNP P34087
G	198	GLY	-	EXPRESSION TAG	UNP P34087
G	199	ALA	-	EXPRESSION TAG	UNP P34087
G	200	LEU	-	EXPRESSION TAG	UNP P34087
G	201	ASP	-	EXPRESSION TAG	UNP P34087
G	202	TYR	-	EXPRESSION TAG	UNP P34087

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	203	ASP	-	EXPRESSION TAG	UNP P34087
G	204	ILE	-	EXPRESSION TAG	UNP P34087
G	205	PRO	-	EXPRESSION TAG	UNP P34087
G	206	THR	-	EXPRESSION TAG	UNP P34087
G	207	THR	-	EXPRESSION TAG	UNP P34087
G	208	ALA	-	EXPRESSION TAG	UNP P34087
G	209	SER	-	EXPRESSION TAG	UNP P34087
G	210	GLU	-	EXPRESSION TAG	UNP P34087
G	211	ASN	-	EXPRESSION TAG	UNP P34087
G	212	LEU	-	EXPRESSION TAG	UNP P34087
G	213	TYR	-	EXPRESSION TAG	UNP P34087
G	214	PHE	-	EXPRESSION TAG	UNP P34087
G	215	GLN	-	EXPRESSION TAG	UNP P34087

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

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

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit 9.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

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

- Molecule 11 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	115	920	590	157	171	2	0	0	1

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

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

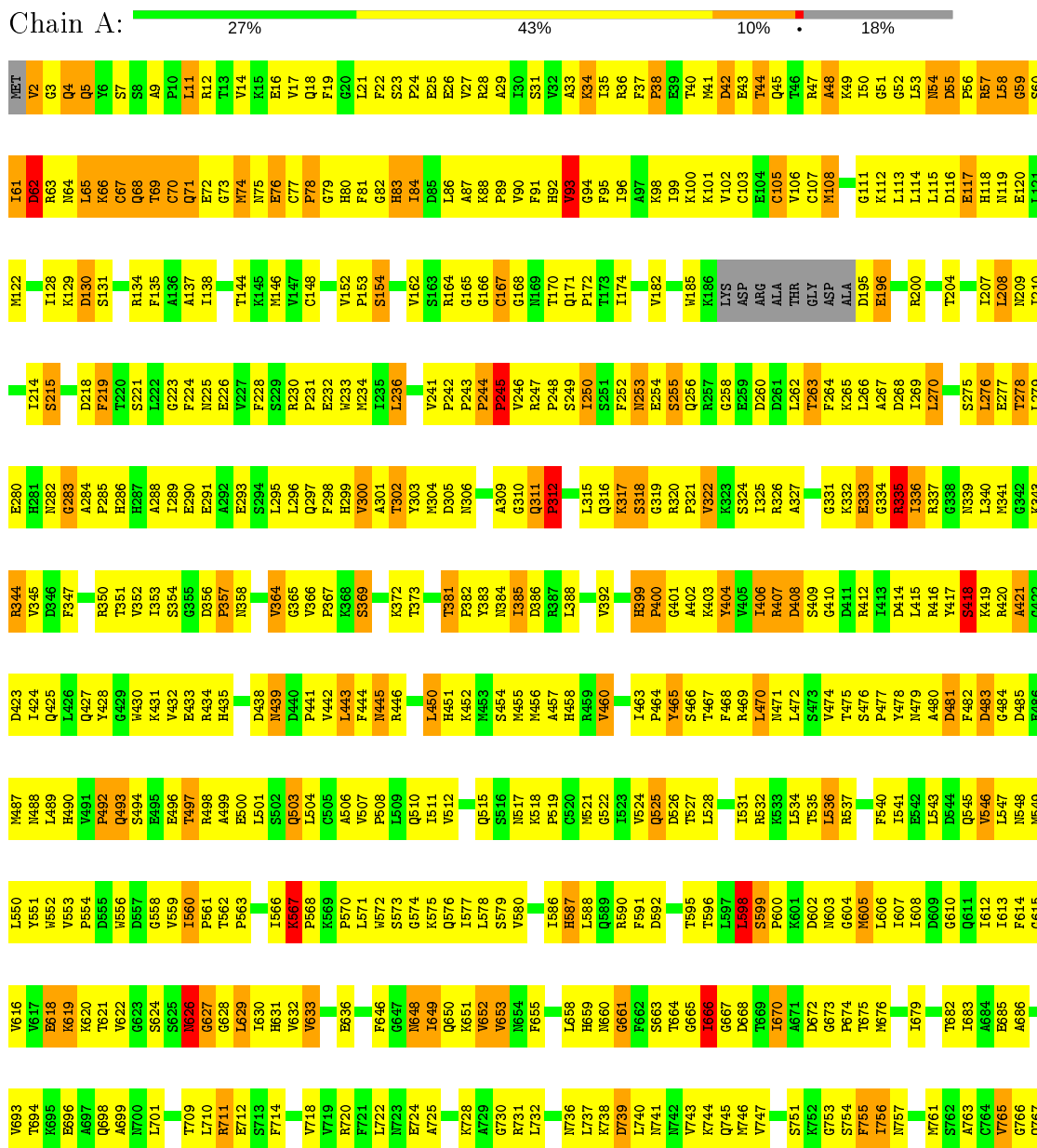
- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

### 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. 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. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: DNA-directed RNA polymerase II largest subunit





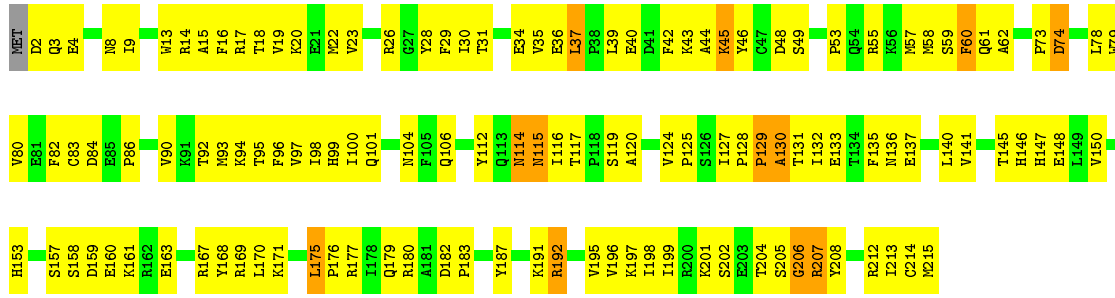
Q768	V842	P910	Y987	S1056	L1192	E1264	D1334	C1400	GLY	TYR	SER
K843	K843	S911	Y987	V1057	L1193	M1265	I1335	S1401	ALA	SER	PRO
A844	A844	L912	N996	V1058	R1194	M1336	M1336	S1401	TYR	PRO	SER
I775	I775	L913	L997	V1059	L1195	M1266	E1337	F1403	GLY	THR	TYR
F779	E846	E914	L998	P1060	E1196	L1268	V1338	E1404	GLY	PRO	SER
V780	D847	I919	V999	M1063	L1197	I1271	L1339	T1405	ALA	SER	PRO
D781	I848	G921	L1000	V1064	D1198	T1272	C1340	V1406	VAL	THR	THR
R782	Y852	G920	R1001	L1067	M1202	L1273	E1342	E1407	THR	THR	SER
T783	D853	D922	K1003	L1073	K1205	R1274	A1343	I1408	PRO	SER	PRO
L784	N854	D922	M1004	G1072	I1206	E1279	A1344	F1410	GLY	THR	TYR
P785	T855	Q926	E1005	E1073	I1207	E1280	R1345	E1411	SER	THR	SER
H786	T856	L929	I1006	E1074	T1208	R1281	A1346	A1412	GLY	PRO	ALA
K789	R857	E932	I1007	E1076	M1209	V1282	A1347	G1413	SER	THR	THR
P794	S859	E932	Q1008	A1076	M1209	V1282	L1348	A1414	SER	THR	SER
E795	S860	Y933	M1009	M1079	G1210	V1283	Y1349	S1415	PRO	PRO	PRO
S796	G861	K934	A1010	M1079	Q1211	K1350	A1416	A1416	ALA	PRO	ALA
K797	G862	Q835	Q1011	T1080	V1212	E1284	E1417	E1417	TYR	THR	GLN
K798	R863	L936	R1012	L1081	G1213	E1285	L1418	L1418	ASN	SER	SER
K799	R864	Y937	D1013	ASN	E1214	V1286	D1419	D1420	PRO	THR	THR
V800	Q865	K938	V1015	THR	R1215	Q1218	A1420	D1421	THR	THR	SER
E801	F866	R940	T1016	HIS	Q1219	I1219	C1421	R1422	SER	TYR	SER
V802	I867	K940	L1017	PHE	I1219	Y1153	C1422	R1422	VAL	THR	PRO
S803	Y868	F942	F1018	PHE	I1220	D1154	G1423	V1424	ASP	THR	TYR
V804	G869	L943	C1019	ALA	K1221	F1220	V1424	V1424	ASP	THR	SER
L805	E870	L943	C1020	ALA	M1222	F1220	M1362	A1434	GLY	THR	SER
R806	D871	F947	C1021	GLY	M1223	F1225	V1363	A1434	VAL	THR	PRO
L807	E871	Y948	L1022	VAL	D1223	F1225	M1364	T1429	LEU	PRO	PRO
G807	G872	K948	R1023	VAL	D1224	F1225	Y1298	L1430	LEU	PRO	PRO
L808	M873	D949	A1024	SER	L1224	F1225	E1299	G1431	LEU	PRO	PRO
T809	D874	D949	S1024	K1092	E1226	V1226	K1300	Q1432	PRO	PRO	PRO
P810	A875	W954	A1027	K1093	I1227	M1227	E1303	M1433	PRO	PRO	PRO
Q811	A876	W954	A1027	K1093	M1228	M1228	M1303	A1434	VAL	THR	TYR
E812	H877	P957	T1028	V1094	M1228	M1228	M1304	A1435	VAL	THR	TYR
F813	I878	P957	R1029	T1095	D1233	D1233	V1305	I1436	ASP	PRO	PRO
F814	Q881	W959	R1030	S1096	D1236	L1236	L1306	G1437	ASP	PRO	PRO
F815	S882	N959	V1031	G1097	I1237	I1237	L1307	T1438	GLY	THR	TYR
H816	L883	I960	L1032	P1099	I1238	I1238	V1374	T1439	GLY	THR	TYR
M818	L884	R961	Q1033	M1000	R1239	R1239	M1375	A1440	ASN	PRO	PRO
G819	T885	R962	E1034	L1101	R1241	R1241	T1376	F1441	ASP	PRO	TYR
G820	T885	R963	Y1035	L1105	R1241	R1241	T1377	D1442	ALA	TYR	TYR
R821	G887	I964	L1037	M1106	V1242	V1242	T1377	D1442	ALA	TYR	TYR
E822	D890	Q965	T1038	V1107	ARG	ARG	T1377	D1446	GLY	PRO	PRO
G823	D890	Q966	K1039	M1107	PRO	PRO	V1384	E1447	PHE	PRO	PRO
I825	R896	Q969	Q1040	M1110	LYS	LYS	T1385	L1450	ALA	TYR	TYR
D826	Y897	Q969	F1042	M1111	SER	SER	R1386	V1451	TYR	PRO	PRO
V829	R898	F971	D1043	T1113	LEU	LEU	H1387	K1452	GLY	PRO	PRO
R830	R898	R971	W1044	P1114	ASP	ASP	G1388	Y1453	GLY	THR	THR
T831	V899	K977	V1045	S1115	ALA	ALA	F1389	M1454	ALA	SER	PRO
T834	D900	P978	M1048	L1116	GLU	GLU	M1390	P1455	ASP	PRO	PRO
T835	L901	P978	I1049	L1117	THR	THR	M1393	GLY	ASP	PRO	PRO
G836	N903	S979	I1049	T1117	THR	THR	T1394	GLN	THR	THR	THR
D837	L902	D980	E1050	V1119	GLU	GLU	T1394	GLN	THR	THR	THR
T838	T904	L981	E1050	V1119	GLU	GLU	G1395	ILE	ALA	PRO	PRO
D839	D905	L981	Q1051	L1120	ASP	ASP	A1396	THR	THR	THR	THR
Y836	H966	I983	Q1052	E1121	E1255	E1255	A1397	THR	THR	THR	THR
R840	T907	R984	F1053	P1122	E1256	E1256	L1397	GLY	PRO	PRO	PRO
L841	L909	L986	G1123	H1124	K1261	K1261	M1398	GLY	PRO	PRO	PRO
	D909	I986	R1055	H1124	M1191	M1191	R1399	GLY	PRO	PRO	PRO

● Molecule 2: DNA-directed RNA polymerase II 140 kDa polypeptide

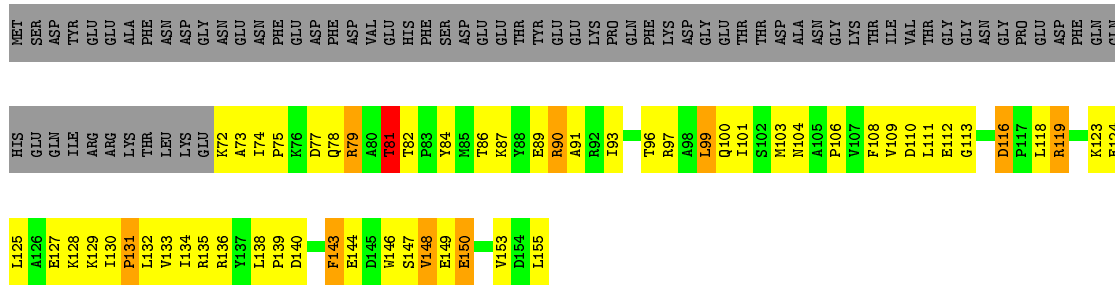




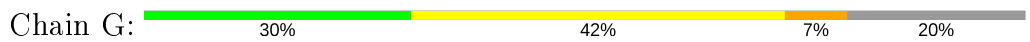




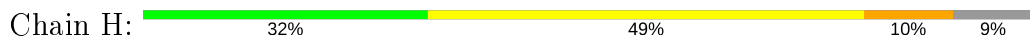
• Molecule 6: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide



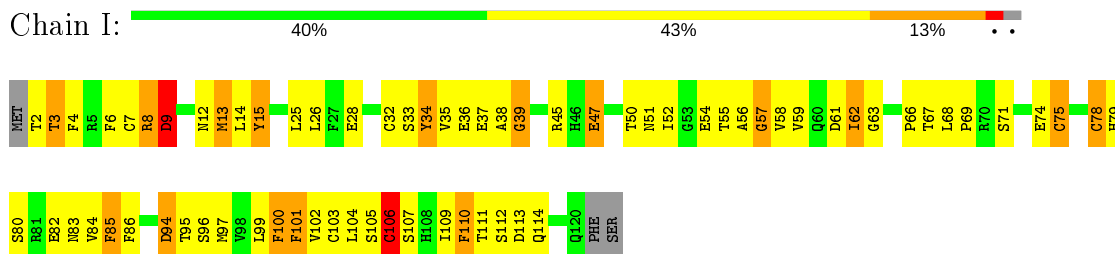
• Molecule 7: DNA-directed RNA polymerase II 19 kDa polypeptide



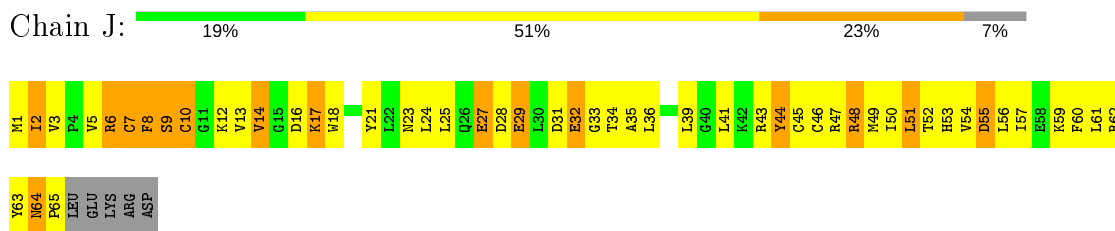
• Molecule 8: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide



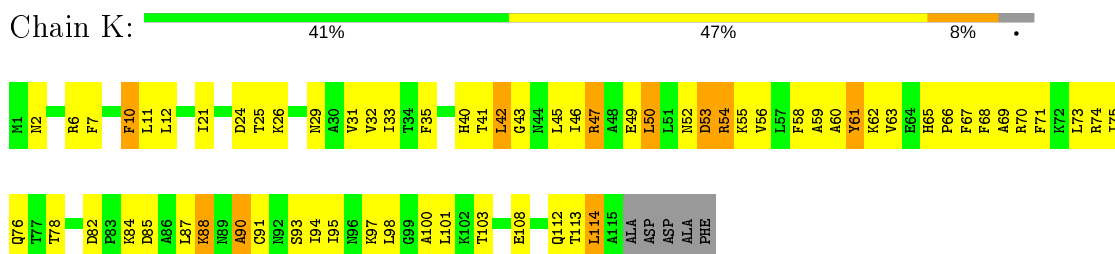
- Molecule 9: DNA-directed RNA polymerase II subunit 9



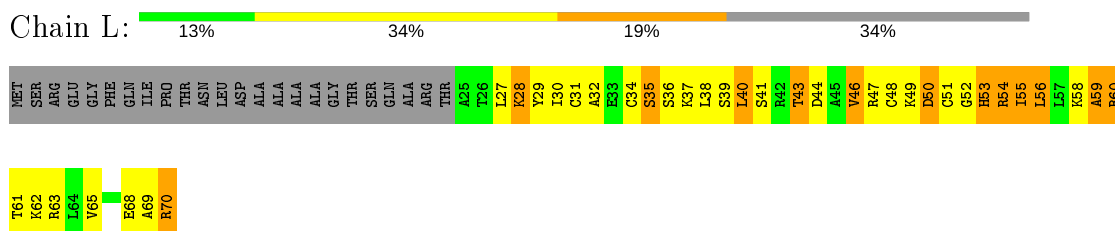
- Molecule 10: DNA-directed RNA polymerases I/II/III subunit 10



- Molecule 11: DNA-directed RNA polymerase II 13.6 kDa polypeptide



- Molecule 12: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	220.69Å 394.33Å 281.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	158.11 – 4.15 161.46 – 4.15	Depositor EDS
% Data completeness (in resolution range)	(Not available) (158.11-4.15) 83.4 (161.46-4.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 4.15Å)	Xtriage
Refinement program		Depositor
R, $R_{free}$	0.387 , (Not available) 0.306 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	118.6	Xtriage
Anisotropy	0.584	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 156.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.023 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.023 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.82	EDS
Total number of atoms	31040	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	169.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.01% 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 i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section:  
ZN

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.49	3/11339 (0.0%)	0.75	9/15334 (0.1%)
2	B	0.53	6/8971 (0.1%)	0.97	25/12103 (0.2%)
3	C	0.52	0/2133	0.76	0/2891
4	D	0.49	0/1382	0.81	3/1862 (0.2%)
5	E	0.44	0/1788	0.64	0/2406
6	F	0.53	0/691	0.78	0/933
7	G	0.53	0/1367	0.74	0/1844
8	H	0.40	0/1086	0.66	0/1470
9	I	0.83	1/989 (0.1%)	0.94	3/1331 (0.2%)
10	J	0.54	0/541	0.90	1/727 (0.1%)
11	K	0.50	0/938	0.68	0/1267
12	L	0.55	0/365	0.79	0/485
All	All	0.52	10/31590 (0.0%)	0.82	41/42653 (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
1	A	0	2
2	B	1	10
3	C	0	1
9	I	0	1
All	All	1	14

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	I	39	GLY	C-N	-21.36	0.84	1.34
2	B	442	PHE	C-N	-8.69	1.14	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	446	LEU	N-CA	-8.12	1.30	1.46
2	B	439	ALA	C-N	7.40	1.51	1.34
1	A	1274	ARG	C-N	-6.45	1.21	1.33

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	475	SER	CB-CA-C	-51.20	12.82	110.10
9	I	39	GLY	O-C-N	-18.15	93.66	122.70
2	B	439	ALA	N-CA-CB	-16.49	87.01	110.10
2	B	442	PHE	C-N-CA	15.77	161.13	121.70
2	B	476	ARG	C-N-CA	-15.07	84.02	121.70

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	475	SER	CA

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	303	TYR	Sidechain
1	A	807	GLY	Mainchain
2	B	217	ARG	Mainchain
2	B	405	ARG	Mainchain
2	B	438	GLU	Peptide

## 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11140	0	11218	1169	0
2	B	8800	0	8777	947	0
3	C	2095	0	2053	241	0
4	D	1373	0	1312	144	0
5	E	1752	0	1776	149	0
6	F	679	0	701	94	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	G	1339	0	1357	145	0
8	H	1068	0	1040	107	0
9	I	971	0	929	91	0
10	J	532	0	542	93	0
11	K	920	0	929	83	0
12	L	363	0	387	47	0
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
All	All	31040	0	31021	3039	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 49.

The worst 5 of 3039 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:469:GLN:CA	2:B:474:SER:CB	1.80	1.51
4:D:119:ARG:N	4:D:121:LYS:HB2	1.46	1.30
4:D:113:PHE:CB	4:D:156:ASP:OD1	1.78	1.30
4:D:118:THR:HA	4:D:121:LYS:CB	1.64	1.27
2:B:435:THR:CG2	2:B:439:ALA:HB2	1.63	1.26

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	1406/1733 (81%)	949 (68%)	293 (21%)	164 (12%)	0	6
2	B	1096/1224 (90%)	744 (68%)	226 (21%)	126 (12%)	0	6
3	C	264/318 (83%)	159 (60%)	66 (25%)	39 (15%)	0	3
4	D	178/221 (80%)	124 (70%)	35 (20%)	19 (11%)	0	8
5	E	212/215 (99%)	147 (69%)	50 (24%)	15 (7%)	1	17
6	F	82/155 (53%)	64 (78%)	14 (17%)	4 (5%)	2	23
7	G	169/215 (79%)	131 (78%)	26 (15%)	12 (7%)	1	17
8	H	129/146 (88%)	84 (65%)	29 (22%)	16 (12%)	0	5
9	I	117/122 (96%)	79 (68%)	31 (26%)	7 (6%)	1	20
10	J	63/70 (90%)	37 (59%)	10 (16%)	16 (25%)	0	1
11	K	113/120 (94%)	89 (79%)	18 (16%)	6 (5%)	2	22
12	L	44/70 (63%)	19 (43%)	14 (32%)	11 (25%)	0	1
All	All	3873/4609 (84%)	2626 (68%)	812 (21%)	435 (11%)	0	7

5 of 435 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	48	ALA
1	A	54	ASN
1	A	55	ASP
1	A	57	ARG

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1239/1520 (82%)	1135 (92%)	104 (8%)	11	36
2	B	952/1061 (90%)	865 (91%)	87 (9%)	9	32
3	C	234/274 (85%)	212 (91%)	22 (9%)	8	30
4	D	138/200 (69%)	122 (88%)	16 (12%)	5	24

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
5	E	196/197 (100%)	187 (95%)	9 (5%)	27 53
6	F	74/137 (54%)	65 (88%)	9 (12%)	5 22
7	G	152/190 (80%)	142 (93%)	10 (7%)	16 43
8	H	117/128 (91%)	111 (95%)	6 (5%)	24 51
9	I	113/116 (97%)	99 (88%)	14 (12%)	4 22
10	J	60/65 (92%)	54 (90%)	6 (10%)	7 28
11	K	99/102 (97%)	92 (93%)	7 (7%)	14 41
12	L	40/57 (70%)	37 (92%)	3 (8%)	13 40
All	All	3414/4047 (84%)	3121 (91%)	293 (9%)	10 35

5 of 293 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	466	TRP
2	B	956	THR
9	I	85	PHE
2	B	516	ASN
2	B	682	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 77 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	60	GLN
2	B	518	HIS
9	I	12	ASN
2	B	121	ASN
2	B	366	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 8 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
2	B	1
1	A	1
9	I	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	807:GLY	C	808:LEU	N	1.20
1	B	442:PHE	C	443:ASN	N	1.14
1	I	39:GLY	C	40:SER	N	0.84

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.