



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 04:46 AM GMT

PDB ID : 2NVZ
Title : RNA Polymerase II elongation complex with UTP, updated 11/2006
Authors : Wang, D.; Bushnell, D.A.; Westover, K.D.; Kaplan, C.D.; Kornberg, R.D.
Deposited on : 2006-11-14
Resolution : 4.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

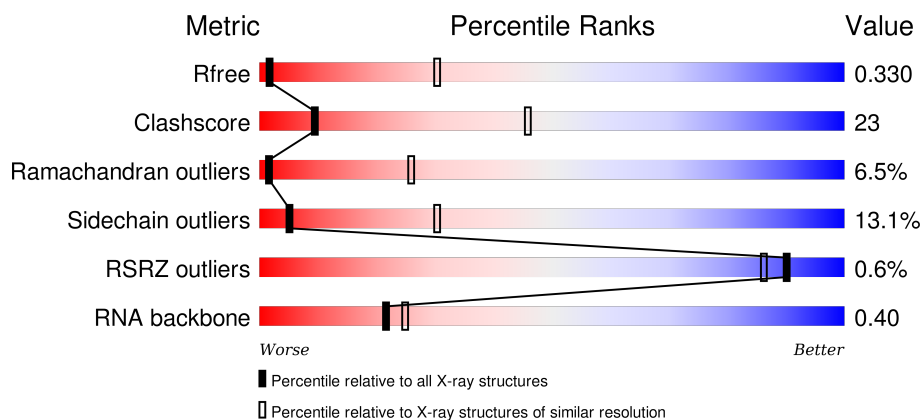
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.30 Å.

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




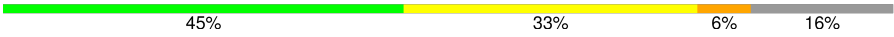




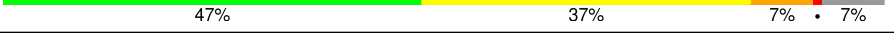

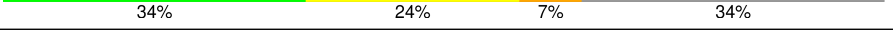
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1059 (5.00-3.60)
Clashscore	102246	1166 (5.00-3.60)
Ramachandran outliers	100387	1106 (5.00-3.60)
Sidechain outliers	100360	1089 (5.00-3.60)
RSRZ outliers	91569	1062 (5.00-3.60)
RNA backbone	2183	1087 (5.60-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	
2	T	28	
3	N	14	
4	A	1733	

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	

2 Entry composition [i](#)

There are 16 unique types of molecules in this entry. The entry contains 29002 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 5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	10	Total	C	N	O	P	0	0	0
			216	98	45	64	9			

- Molecule 2 is a DNA chain called 28-MER 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 5'-D(*CP*TP*GP*CP*TP*TP*AP*TP*CP*GP*GP*TP*AP*G)-3'.

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	1398	Total	C	N	O	S	0	0	0
			10984	6930	1924	2069	61			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1096	Total	C	N	O	S	0	0	0
			8701	5508	1518	1620	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	193	Total	C	N	O	S	0	0	0
			1594	1016	283	287	8			

- 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	83	Total	C	N	O	S	0	0	0
			670	428	114	125	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 subunit 9.

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/III subunit 10.

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 polypep-

tide.

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

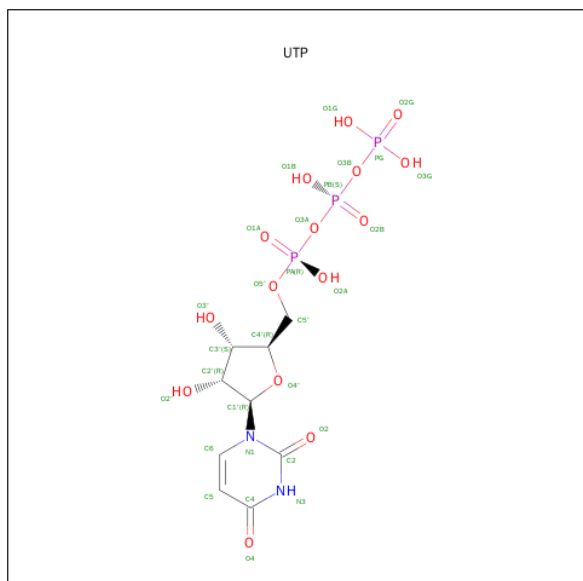
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

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 URIDINE 5'-TRIPHOSPHATE (three-letter code: UTP) (formula: C₉H₁₅N₂O₁₅P₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
16	B	1	29	9	2	15	3	0	0

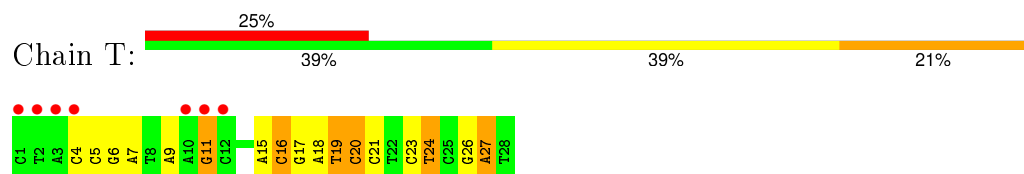
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

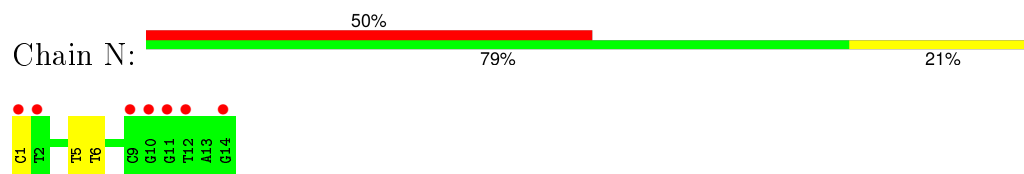
- Molecule 1: 5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*A)-3'



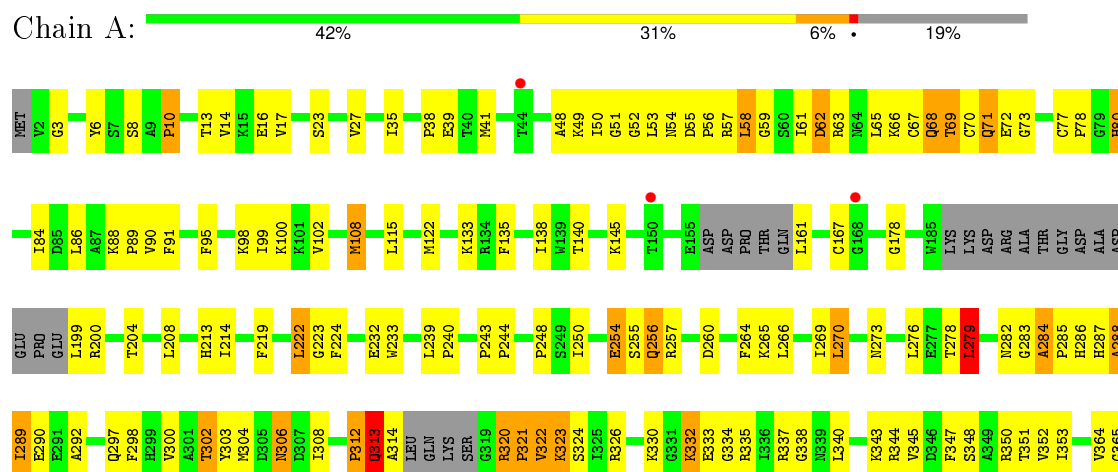
- Molecule 2: 28-MER DNA template strand



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



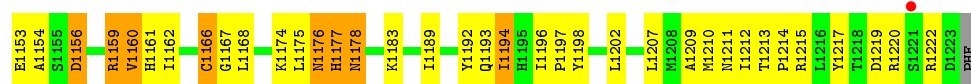
- Molecule 4: DNA-directed RNA polymerase II largest subunit



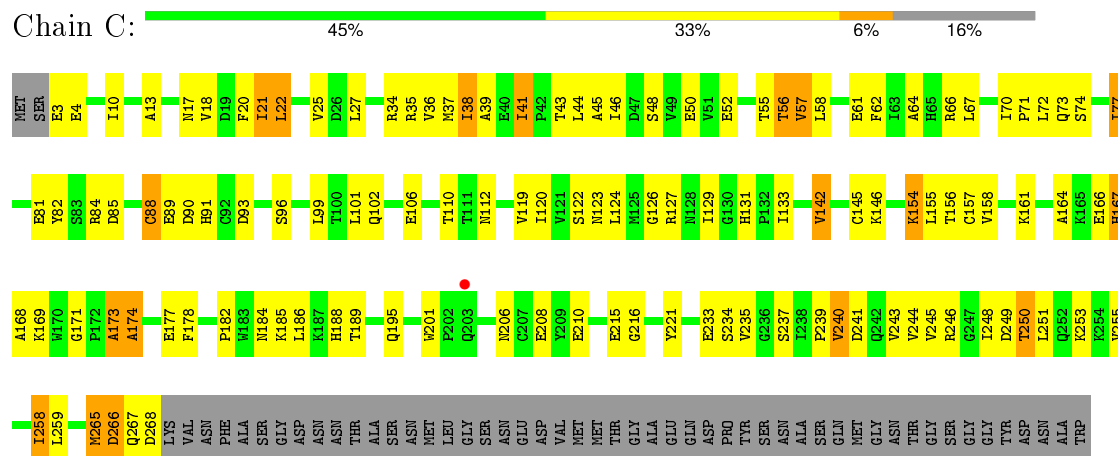
PRO	THR	GLY	LYS	E1264	SER	E1103	L1017	Q926	B346	C764	I582	P508	R441	V366
SER	THR	GLY	TYR	D1373	PHE	I1104	F1018	V927	D847	V765	Q673	P508	P441	P367
PRO	SER	ALA	MET	M1267	ASP	L1105	C1019	L928	R348	G766	Q689	Q510	V442	
THR	PRO	THR	PRO	L1268	Q1187	N1106	R1025	L929	B349	Q767	R590	R591	R444	I370
SER	THR	GLY	GLN	E1269	Q1188	V1107	R1025	D930	D853	Q768	F591	S513	R445	
PRO	SER	GLY	LYS	M1270	S1189	T1113	T1028	B331	B354	R774	T680	Q515	R446	T373
THR	THR	ALA	ILE	I1271	P1190	P1114	R1029	B332	B355	I775	B881	Q516	Q447	T374
GLY	SER	GLU	THR	V1276	L1192	S1115	R1030	Y933	R355	A776	L683	N517	P448	T375
GLU	PRO	GLU	GLU	E1277	R1194	L1116	E1034	L936	B357	F777	A684	K518	L450	E378
PRO	PRO	ILE	GLU	R1281	L1195	T1117	Y1035	T941	B358	R782	K687	N521	R451	V379
THR	THR	GLU	GLU	V1282	L1196	Y1118	Y1036	B941	S559	T783	T680	G522	R452	V380
PRO	PRO	GLY	ASP	Y1283	L1197	Y1119	R1036	F942	B360	T783	D692	I523	R453	T381
GLY	THR	GLN	GLN	M1284	D1198	L1120	E1050	R943	B362	S788	L612	V524	S454	P382
THR	THR	GLY	ASP	T1394	R1199	D1127	E1050	B945	B364	K789	I613	Q525	N455	Y383
SER	THR	GLY	GLY	Y1287	A1200	Q1128	L1054	T949	B367	Y792	V617	D526	N456	R387
PRO	PRO	ALA	GLY	M1398	A1201	Q1128	L1054	D949	B368	T792	Q688	T527	N457	T388
THR	THR	PRO	VAL	T1295	M1202	A1131	V1057	B969	B369	S793	B918	L528	R459	T389
SER	THR	THR	THR	F1402	N1203	I1134	V1058	L956	B370	T794	K619	T528	V460	Q390
PRO	PRO	SER	PRO	V1304	D1206	T1134	M1063	P957	D871	E795	L701	L534	R461	I391
PRO	PRO	GLY	TYR	L1306	V1305	T1141	V1064	P958	B372	S796	L702	T535	I461	P396
THR	THR	GLY	SER	E1307	G1213	T1142	G1065	N959	B373	K797	T703	I536	T463	
PRO	PRO	PHE	ASN	T1308	E1214	L1143	G1065	T960	B373	T797	T703	I536	T463	
THR	THR	GLU	GLU	E1411	R1215	L1143	V1066	R961	B380	V800	A704	D538	P464	
SER	THR	GLY	GLY	V1311	R1216	S1150	L1067	Q961	B381	E801	G707	D538	Y465	R399
PRO	PRO	ALA	LEU	N1312	D1223	E1151	A1069	I964	S882	N802	M708	I541	S466	P400
ALA	ALA	GLY	VAL	E1415	D1223	I1152	Q1070	Q965	L683	L808	T710	L543	T467	K403
THR	THR	GLY	ASN	A1416	L1224	Y1153	I1072	Q968	B386	L808	T710	L543	R469	Y404
SER	SER	PHE	ALA	D1322	I1227	D1155	G1073	Q969	B387	F814	E712	D544	L470	V405
PRO	PRO	PRO	ASP	D1420	W1228	P1156	E1074	T970	B389	F815	S713	Q545	I471	T406
THR	THR	THR	LEU	P1324	D1228	D1157	P1075	B972	B394	R816	F714	V546	L472	R407
SER	PRO	SER	VAL	T1325	D1231	P1158	A1076	B973	B395	R818	F721	M549	S473	D411
PRO	PRO	PRO	LYS	R1326	ASN	R1159	T1077	D974	B396	G819	L722	P639	V474	
THR	THR	ASP	ASP	I1327	ASP	S1160	Q1078	B975	B397	B820	L722	P639	P477	I413
GLY	GLY	GLU	GLU	Y1328	GLU	T1161	M1079	R898	B398	R821	K728	W552	Y478	
PRO	PRO	LEU	LEU	T1329	LYS	V1162	T1080	K977	B399	E822	C642	W556	N479	Y416
THR	THR	PRO	THR	I1428	L1236	I1163	L1081	P978	D900	G823	R731	L560	A480	
SER	THR	PHE	PHE	V1429	I1237	P1164	M1082	S979	L901	L824	D481	L560	D481	R420
PRO	PRO	SER	SER	L1430	D1334	E1165	T1083	D980	L902	I825	C642	L560	P482	
THR	THR	PRO	PRO	M1335	R1241	D1166	F1084	B987	B399	D826	L736	P561	D483	R420
PRO	PRO	ALA	LEU	G1336	Y1242	E1167	H1085	V987	L903	G847	T637	P561	G484	D423
THR	THR	VAL	VAL	V1243	W1242	I1169	A1087	L988	B399	R827	N648	P563	I424	I424
SER	PRO	ASP	ASP	G1340	ARG	I1169	F1086	G989	T907	A828	Q650	I565	N488	I426
PRO	PRO	SER	SER	I1341	PRO	I1170	G1087	D909	D909	B929	L740	I565	N488	I426
THR	THR	THR	THR	E1342	LYS	Q1171	V1089	P910	P910	K830	K651	I566	N489	Q427
THR	THR	GLY	GLY	A1343	SER	L1172	A1090	T994	T994	T831	V652	K567	L490	Y428
PRO	PRO	SER	SER	L1348	LEU	H1173	S1091	Q994	S911	A832	L657	P568	V491	G429
THR	THR	ASP	ASP	L1442	ASP	F1174	K1092	N996	L913	E833	Q745	K569	Q492	W430
SER	THR	ALA	ALA	V1443	ALA	S1175	K1093	L997	B314	T834	T742	L566	K493	K431
PRO	PRO	GLY	GLY	S1358	GLY	L1176	V1084	L998	S915	G835	W572	W572	S494	V432
THR	THR	ALA	ALA	N1364	THR	LEU	T1095	L998	S915	I837	N660	S573	S494	E433
SER	THR	GLY	GLY	Y1365	GLU	ASP	S1096	K1003	B918	I837	S663	Q574	L501	R434
PRO	PRO	GLY	GLY	R1366	GLU	GLU	G1097	L919	T919	Q838	I666	Q575	S502	R435
PRO	PRO	PHE	PHE	K1261	ALA	ALA	V1098	D1013	R920	R839	I666	Q575	Q503	I436
THR	THR	THR	THR	L1261	GLU	GLU	P1099	A1014	L920	R840	G667	I577	L504	M437
SER	PRO	LEU	LEU	L1370	GLY	GLY	R1100	G921	G921	L841	N757	L578	C505	D438
PRO	PRO	THR	THR	L1371	VAL	VAL	T1262	D922	D922	T945	W569	S579	N439	I440

Chain B: 44% 38% 7% 10%

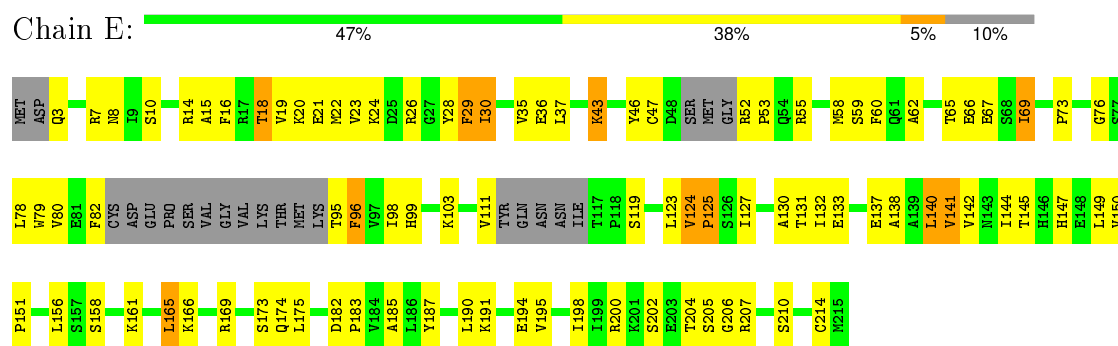




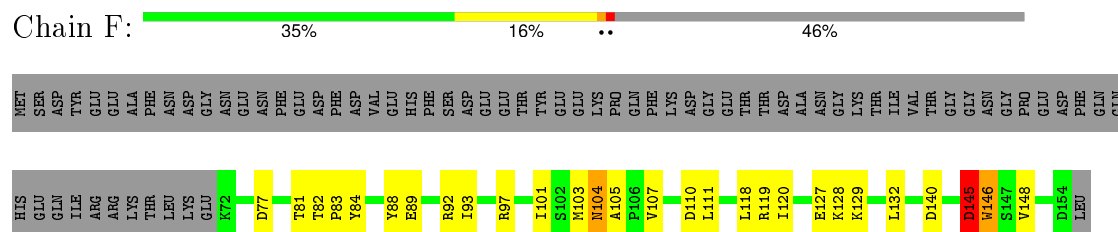
- Molecule 6: DNA-directed RNA polymerase II 45 kDa polypeptide



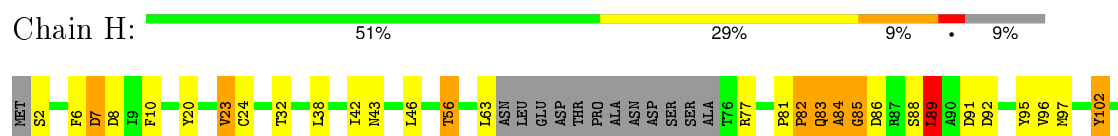
- Molecule 7: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide



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



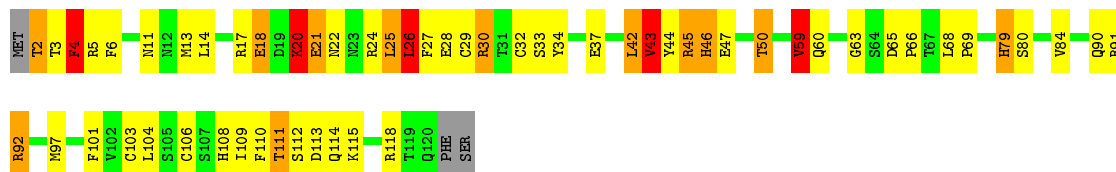
- Molecule 9: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide





• Molecule 10: DNA-directed RNA polymerase II subunit 9

Chain I: 50% 34% 10%



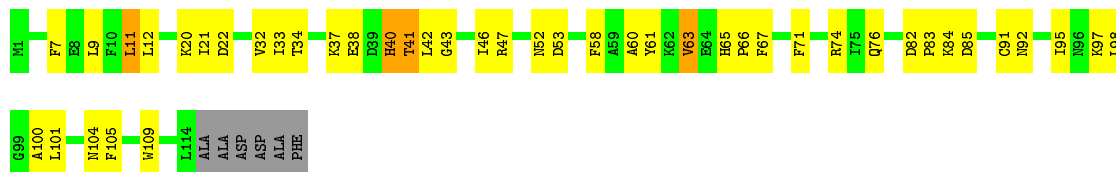
• Molecule 11: DNA-directed RNA polymerases I/II/III subunit 10

Chain J: 47% 37% 7%



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

Chain K: 58% 33% 5%



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

Chain L: 34% 24% 7% 34%



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	169.65Å 222.34Å 194.32Å 90.00° 101.67° 90.00°	Depositor
Resolution (Å)	40.00 – 4.30 39.94 – 4.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-4.30) 86.4 (39.94-4.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.67 (at 4.28Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.270 , 0.332 0.277 , 0.330	Depositor DCC
R_{free} test set	4179 reflections (11.22%)	DCC
Wilson B-factor (Å ²)	97.0	Xtriage
Anisotropy	0.409	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 60.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 41426 reflections	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	29002	wwPDB-VP
Average B, all atoms (Å ²)	126.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.04% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: UTP, MG, 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	R	0.81	0/243	1.33	0/378
2	T	0.87	0/634	1.63	17/975 (1.7%)
3	N	0.77	0/317	1.35	1/488 (0.2%)
4	A	0.45	1/11180 (0.0%)	0.67	3/15117 (0.0%)
5	B	0.45	0/8866	0.65	1/11956 (0.0%)
6	C	0.43	0/2133	0.60	0/2891
7	E	0.44	0/1625	0.61	0/2182
8	F	0.43	0/682	0.61	1/922 (0.1%)
9	H	0.47	0/1086	0.77	2/1470 (0.1%)
10	I	0.48	0/989	0.74	0/1331
11	J	0.47	0/541	0.62	0/727
12	K	0.46	0/937	0.60	0/1265
13	L	0.53	0/365	0.79	0/485
All	All	0.47	1/29598 (0.0%)	0.72	25/40187 (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
4	A	0	9
5	B	0	1
9	H	0	2
All	All	0	12

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1158	PRO	CG-CD	5.94	1.70	1.50

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	T	23	DC	O4'-C4'-C3'	-9.30	100.42	106.00
2	T	16	DC	O4'-C1'-N1	9.29	114.50	108.00
2	T	11	DG	O4'-C1'-N9	8.90	114.23	108.00
2	T	23	DC	O4'-C1'-N1	8.24	113.77	108.00
2	T	20	DC	O4'-C4'-C3'	-8.15	101.11	106.00

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	1068	ALA	Peptide
4	A	1069	ALA	Peptide
4	A	1070	GLN	Peptide
4	A	451	HIS	Peptide
4	A	974	ASP	Peptide

5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	R	216	0	109	6	0
2	T	566	0	316	18	0
3	N	284	0	161	2	0
4	A	10984	0	11069	568	0
5	B	8701	0	8728	479	0
6	C	2095	0	2051	92	0
7	E	1594	0	1622	52	0
8	F	670	0	690	15	0
9	H	1068	0	1040	52	0
10	I	971	0	928	63	0
11	J	532	0	542	38	0
12	K	919	0	929	32	0
13	L	363	0	387	13	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
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	29	0	11	2	0
All	All	29002	0	28583	1299	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 23.

The worst 5 of 1299 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)
4:A:701:LEU:HA	4:A:702:LEU:CB	1.68	1.20
4:A:975:HIS:HB3	4:A:976:THR:OG1	1.41	1.19
4:A:335:ARG:HD2	5:B:1202:LEU:HD12	1.27	1.16
4:A:1167:GLU:CB	4:A:1168:GLU:HA	1.75	1.15
5:B:1019:SER:HB2	5:B:1020:ARG:HB2	1.24	1.15

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	1384/1733 (80%)	1055 (76%)	224 (16%)	105 (8%)	1	20
5	B	1074/1224 (88%)	855 (80%)	161 (15%)	58 (5%)	2	30
6	C	264/318 (83%)	217 (82%)	35 (13%)	12 (4%)	3	34
7	E	185/215 (86%)	146 (79%)	26 (14%)	13 (7%)	1	23
8	F	81/155 (52%)	62 (76%)	15 (18%)	4 (5%)	3	32

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	H	129/146 (88%)	87 (67%)	30 (23%)	12 (9%)	1	16
10	I	117/122 (96%)	73 (62%)	32 (27%)	12 (10%)	1	12
11	J	63/70 (90%)	56 (89%)	4 (6%)	3 (5%)	3	32
12	K	112/120 (93%)	98 (88%)	13 (12%)	1 (1%)	21	67
13	L	44/70 (63%)	29 (66%)	11 (25%)	4 (9%)	1	16
All	All	3453/4173 (83%)	2678 (78%)	551 (16%)	224 (6%)	1	26

5 of 224 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	63	ARG
4	A	223	GLY
4	A	284	ALA
4	A	313	GLN
4	A	404	TYR

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%)	1055 (87%)	163 (13%)	5	30
5	B	951/1061 (90%)	827 (87%)	124 (13%)	5	31
6	C	234/274 (85%)	207 (88%)	27 (12%)	7	35
7	E	177/197 (90%)	160 (90%)	17 (10%)	10	44
8	F	73/137 (53%)	65 (89%)	8 (11%)	8	38
9	H	117/128 (91%)	103 (88%)	14 (12%)	6	33
10	I	113/116 (97%)	92 (81%)	21 (19%)	2	15
11	J	60/65 (92%)	51 (85%)	9 (15%)	3	25
12	K	99/102 (97%)	88 (89%)	11 (11%)	8	37
13	L	40/57 (70%)	31 (78%)	9 (22%)	1	10
All	All	3082/3657 (84%)	2679 (87%)	403 (13%)	5	30

5 of 403 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	242	SER
5	B	651	LEU
10	I	109	ILE
5	B	297	ILE
5	B	425	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
4	A	1390	ASN
5	B	515	HIS
9	H	33	GLN
5	B	121	ASN
5	B	366	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	8/10 (80%)	1 (12%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	8	G

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	UTP	B	3000	15	20,30,30	2.06	5 (25%)	30,47,47	1.75	3 (10%)

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	UTP	B	3000	15	-	0/18/38/38	0/2/2/2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	B	3000	UTP	PG-O3G	2.07	1.62	1.54
16	B	3000	UTP	PG-O2G	3.06	1.61	1.51
16	B	3000	UTP	O4'-C1'	4.18	1.46	1.41
16	B	3000	UTP	C6-N1	4.63	1.42	1.35
16	B	3000	UTP	C4-N3	4.71	1.41	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	B	3000	UTP	PB-O3B-PG	-3.40	121.26	132.67
16	B	3000	UTP	PB-O3A-PA	-2.97	124.38	132.73
16	B	3000	UTP	C4-N3-C2	6.49	120.57	114.14

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	B	3000	UTP	2	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	R	10/10 (100%)	-0.32	0 100 100	89, 111, 136, 141	0
2	T	28/28 (100%)	0.95	7 (25%) 1 2	81, 200, 285, 291	0
3	N	14/14 (100%)	1.78	7 (50%) 0 1	265, 273, 291, 292	0
4	A	1398/1733 (80%)	-0.35	5 (0%) 93 90	84, 118, 171, 184	0
5	B	1096/1224 (89%)	-0.36	1 (0%) 95 95	86, 114, 151, 165	0
6	C	266/318 (83%)	-0.40	1 (0%) 93 90	99, 118, 146, 150	0
7	E	193/215 (89%)	-0.30	0 100 100	98, 130, 163, 167	0
8	F	83/155 (53%)	-0.33	0 100 100	111, 127, 138, 144	0
9	H	133/146 (91%)	-0.24	0 100 100	120, 141, 176, 181	0
10	I	119/122 (97%)	-0.19	0 100 100	118, 149, 168, 173	0
11	J	65/70 (92%)	-0.54	0 100 100	103, 119, 134, 136	0
12	K	114/120 (95%)	-0.32	0 100 100	99, 118, 134, 136	0
13	L	46/70 (65%)	-0.02	0 100 100	137, 178, 191, 193	0
All	All	3565/4225 (84%)	-0.32	21 (0%) 90 86	81, 120, 168, 292	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	T	2	DT	4.4
2	T	3	DA	3.9
2	T	10	DA	3.6
2	T	1	DC	3.3
3	N	11	DG	3.2

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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
16	UTP	B	3000	29/29	0.96	0.16	-1.08	110,112,113,113	0
14	ZN	B	1307	1/1	0.97	0.05	-1.36	147,147,147,147	0
14	ZN	A	1735	1/1	0.95	0.05	-1.54	165,165,165,165	0
14	ZN	C	319	1/1	0.98	0.03	-1.68	117,117,117,117	0
14	ZN	L	105	1/1	0.95	0.11	-2.01	181,181,181,181	0
14	ZN	J	101	1/1	0.98	0.12	-2.20	112,112,112,112	0
14	ZN	A	1734	1/1	0.92	0.07	-2.33	177,177,177,177	0
14	ZN	I	203	1/1	0.95	0.09	-2.39	121,121,121,121	0
14	ZN	I	204	1/1	0.95	0.04	-3.60	157,157,157,157	0
15	MG	A	2001	1/1	0.84	0.14	-	97,97,97,97	0
15	MG	A	2002	1/1	0.88	0.19	-	102,102,102,102	0

6.5 Other polymers [i](#)

There are no such residues in this entry.