



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 11:45 am GMT

PDB ID : 1SFO
Title : RNA POLYMERASE II STRAND SEPARATED ELONGATION COMPLEX
Authors : Westover, K.D.; Bushnell, D.A.; Kornberg, R.D.
Deposited on : 2004-02-20
Resolution : 3.61 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : trunk31020
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk31020

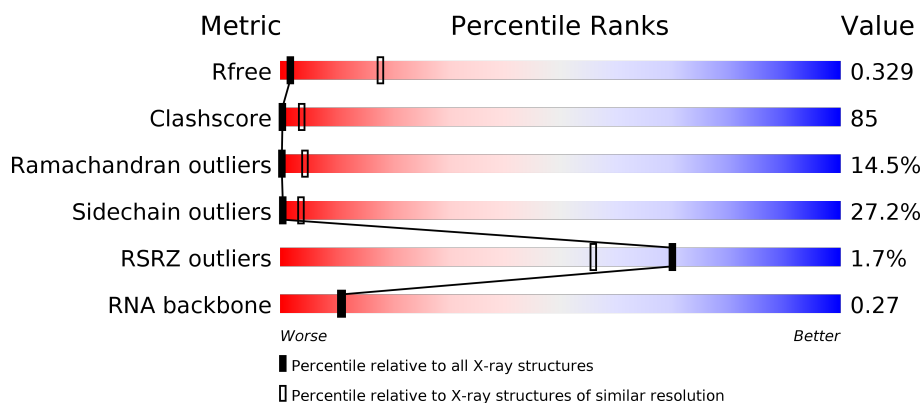
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.61 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}	111664	1075 (3.74-3.50)
Clashscore	122126	1163 (3.74-3.50)
Ramachandran outliers	120053	1122 (3.74-3.50)
Sidechain outliers	120020	1122 (3.74-3.50)
RSRZ outliers	108989	1043 (3.76-3.48)
RNA backbone	2636	1068 (4.30-2.90)

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	14	
3	A	1733	
4	B	1224	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	C	318	<div><div></div><div>19%44%18%16%</div></div>
6	E	215	<div><div></div><div>7%16%51%27%7%</div></div>
7	F	155	<div><div></div><div>10%31%12%46%</div></div>
8	H	146	<div><div></div><div>2%13%40%29%9%9%</div></div>
9	I	122	<div><div></div><div>18%40%30%10%</div></div>
10	J	70	<div><div></div><div>14%44%26%9%7%</div></div>
11	K	120	<div><div></div><div>14%61%18%5%</div></div>
12	L	70	<div><div></div><div>%6%21%31%7%34%</div></div>

2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 28647 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 STRAND.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	14	Total	C	N	O	P	0	0	0
			279	135	48	83	13			

- Molecule 3 is a protein called DNA-directed RNA polymerase II largest subunit.

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

- Molecule 4 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	1106	Total	C	N	O	S	0	0	0
			8793	5568	1538	1632	55			

- Molecule 5 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

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

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

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

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

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

- 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
8	H	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 9 is a protein called DNA-directed RNA polymerase II 14.2 kDa polypeptide.

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

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide.

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

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

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

- 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
12	L	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

- 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		

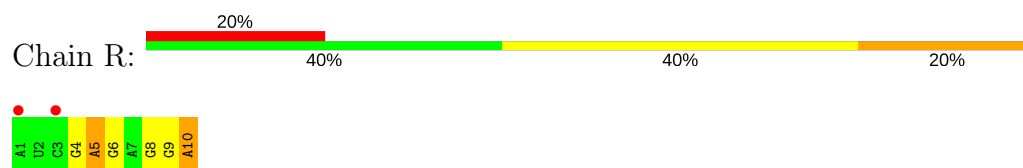
- Molecule 14 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	1	Total	Mg	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 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.

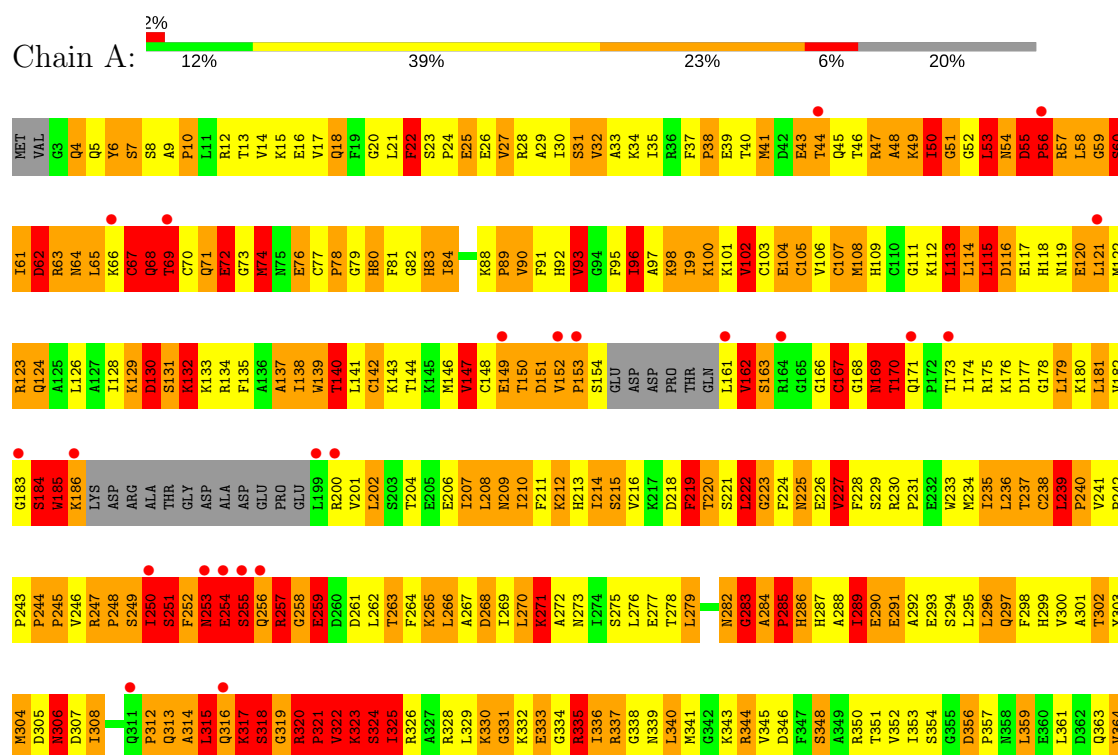
• Molecule 1: RNA STRAND



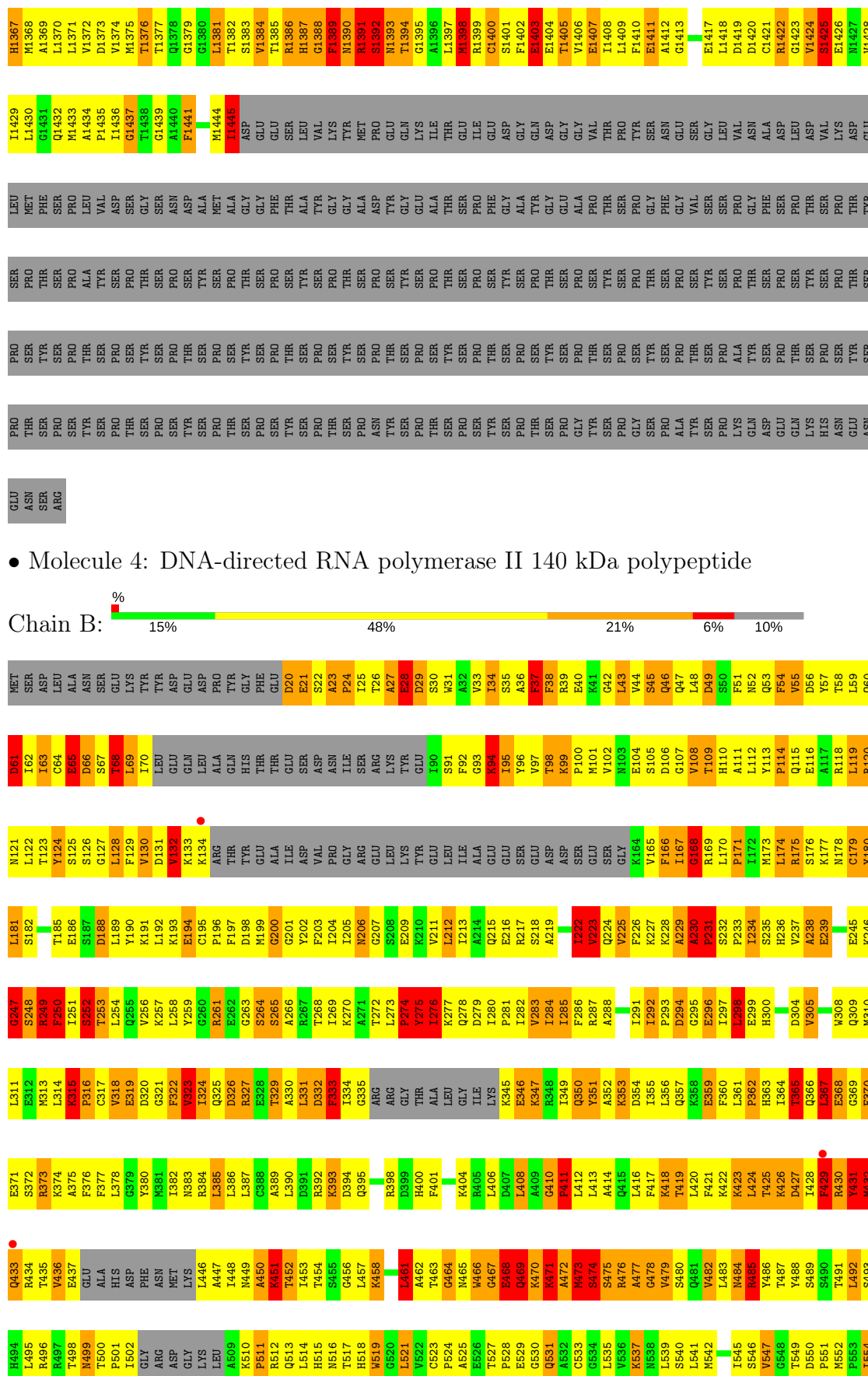
• Molecule 2: DNA STRAND



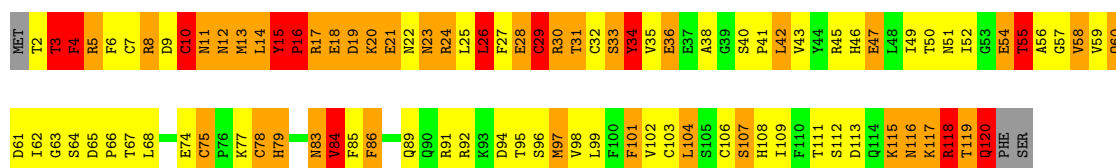
• Molecule 3: DNA-directed RNA polymerase II largest subunit



PRO	PHE	H1124	V1063	K1003	K938	I878	A817	M748	V617	W566	Q493	Q429	G366
LYS	ASP	A1125	V1064	N1004	D939	E879	M818	M748	V617	W566	Q493	Q429	G366
SER	Q1187	A1126	G1065	I1006	R940	K880	G819	S751	E818	D557	S494	W430	V366
LEU	Q1188	D1127	V1066	I1007	K941	K881	G820	K752	K619	G588	E495	K431	P367
ASP	S1189	Q1128	S1189	I1007	F942	S882	R821	G753	K620	E496	E496	V432	K368
ALA	P1190	Q1129	A1068	Q1008	L943	L883	G822	S754	V622	I560	R498	E433	S369
GLU	W1191	Q1130	A1069	N1009	R944	D884	G823	G755	V622	P561	R498	R434	I370
THR	L1192	A1131	Q1070	A1010	E945	T886	L824	F756	S625	T562	A499	H435	I374
GLU	L1193	K1132	G1073	Q1011	V946	G887	R825	M757	L629	P563	E500	I436	T375
E1254	L1194	L1133	E1074	R1012	F947	G888	D826	F758	L629	I566	L501	M437	T376
E1255	L1195	L1134	P1075	D1013	F947	S889	T827	A759	L630	I566	L504	D438	P377
D1257	L1196	S1136	A1076	A1014	A952	R890	A828	G764	V633	P563	C505	V442	E378
H1258	L1197	L1137	T1077	V1015	N953	D890	R829	V765	V633	K567	C505	V442	V379
M1259	L1198	A1137	T1078	L1016	W954	A891	K830	V765	T634	K569	V507	L443	V380
L1260	E1138	I1138	Q1078	L1017	P955	A892	T831	Q768	R635	P570	L511	F444	T381
L1261	A1200	E1139	M1079	F1018	L956	F893	A832	Q768	P639	L571	V512	N445	P382
K1262	M1201	H1140	T1080	C1019	P957	E894	E833	S769	Q640	W572	S513	R446	V383
L1263	M1202	T1141	L1081	G1020	V958	K895	T834	V770	Q640	S573	P514	Q447	N384
E1264	L1242	L1142	ASN	L1021	N959	R896	G835	E771	V641	G574	P514	P448	I385
M1265	L1143	L1143	THR	L1022	T960	Y897	R836	G778	C642	K575	Q515	S449	D386
L1266	V1146	V1146	PHE	R1023	R961	R898	R837	R774	A643	Q576	L515	L459	R387
D1267	H1206	HIS	HIS	S1024	R962	V899	Q838	I775	R644	I577	N517	K451	L388
M1267	L1207	PHE	PHE	R1025	L963	D900	R839	A776	L645	L578	K518	K452	T389
L1268	I1148	I1148	ALA	L1026	I964	L901	R840	F777	L645	S579	P519	M453	Q390
M1269	A1149	A1149	GLY	L1027	Q965	L802	L841	G778	N648	V580	C520	S454	L391
G1210	S1150	S1150	VAL	T1028	N966	N903	V842	F779	T649	A581	N521	M455	K394
Q1211	E1151	E1151	ALA	R1029	A967	T904	K843	V780	Q650	I582	G522	M456	K395
T1212	I1152	I1152	SER	R1030	Q968	D905	A844	D781	R651	P583	L523	K457	G396
G1213	Y1153	Y1153	K1092	V1031	Q969	H906	E845	R782	V652	P583	V524	H458	P396
L1214	Y1154	Y1154	K1093	L1032	T970	T907	E846	R783	V653	I586	Q525	V462	N397
R1215	D1155	D1155	V1094	Q1033	F971	L908	R847	L784	N654	H587	D526	V462	E398
L1216	P1156	P1156	T1095	E1034	H972	D909	L848	R785	P655	L588	T527	L463	R399
K1217	D1157	D1157	S1096	L1036	I973	P910	M849	H786	W656	Q589	C529	P464	P400
Q1218	P1158	P1158	G1097	R1037	D974	S911	Y852	F787	L657	R590	C529	Y465	G401
T1219	V1098	V1098	L1036	L1037	D974	S911	Y852	S788	L658	F591	G530	S466	A402
F1220	S1160	S1160	P1099	T1038	P978	L913	D853	K789	H659	D592	L531	R469	K403
T1161	R1100	R1100	K1039	L1039	S979	E914	N854	Y792	N660	E593	R532	L470	Y404
V1162	L1101	L1101	Q1040	Q1040	D980	S915	T855	S793	G663	G594	R533	L470	Y405
I1163	K1102	K1102	A1041	A1041	L981	G916	T856	F794	P662	T596	L534	M471	L406
E1103	E1103	E1103	F1042	F1042	T982	S917	R857	S794	Q665	L597	L536	L472	R407
L1104	L1104	L1104	D1043	D1043	K984	E918	N858	E795	Q665	L597	L536	S473	D408
L1105	D1166	D1166	L1044	L1044	K984	I919	S859	S796	L666	L598	R537	V474	S409
E1167	E1167	E1167	V1045	V1045	D985	L920	L860	Y796	G667	S599	D538	T475	Y409
V1107	L1107	L1107	L1046	L1046	I986	G921	G861	V800	D668	P600	T539	S476	R412
A1108	A1108	A1108	S1047	S1047	V967	D922	N862	E801	T669	K601	F540	P477	L413
K1109	K1109	K1109	I1048	I1048	L988	L923	V863	N802	L670	D602	L541	P477	L414
N1110	N1110	N1110	I1049	I1049	G989	K924	L864	S803	A671	N603	E542	M479	L415
M1111	M1111	M1111	E1050	E1050	V990	L925	Q865	Y804	D672	G604	L544	A480	R416
H1173	H1173	H1173	A1051	A1051	K991	Q926	F866	L805	G673	M605	D544	D481	R416
F1174	T1113	T1113	Q1052	Q1052	D992	V927	L867	R806	P674	L606	O545	F482	S418
S1175	P1114	P1114	F1053	F1053	L993	L928	Y868	G807	T675	I607	V546	D483	K419
L1176	L1176	L1176	L1054	L1054	Q994	L929	G869	L308	M676	L608	L547	G484	R420
LEU	L1115	L1115	R1055	R1055	E995	D930	E870	T809	Q676	D609	L547	G484	R420
ASP	T1117	T1117	S1056	S1056	N996	E931	D871	N741	L679	D609	L547	D485	A421
GLU	GLU	GLU	V1057	V1057	L997	E932	M872	Q811	T680	G610	N549	E486	G422
GLU	GLU	GLU	V1058	V1058	L998	Y933	M873	E812	E681	Q611	L550	M487	D423
ALA	L1120	L1120	H1059	H1059	V999	K934	D874	F813	T682	I613	N552	M488	L424
GLU	P1060	P1060	T1060	T1060	L1000	Q935	A875	F914	L683	P614	V553	L489	Q425
GLN	P1122	P1122	E1301	E1301	R1001	L936	A876	F915	A684	G615	P554	H490	L426
SER	SER	SER	E1062	E1062	G1002	V937	H877	H816	E885	V616	D555	P492	Y428

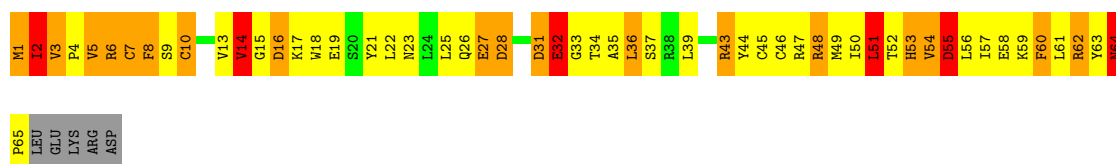






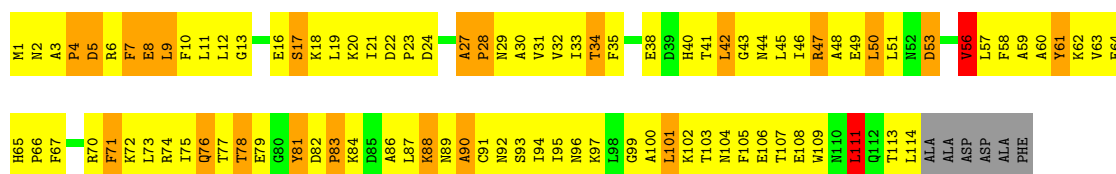
- Molecule 10: DNA-directed RNA polymerases I, II, and III 8.3 kDa polypeptide

Chain J: 14% 44% 26% 9% 7%



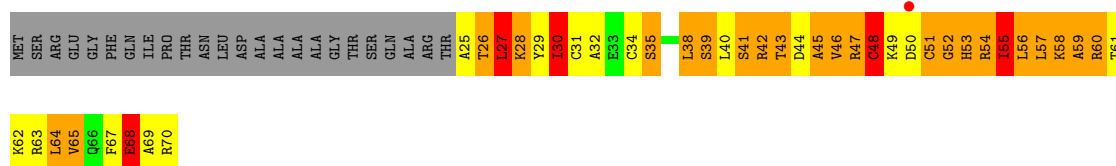
- Molecule 11: DNA-directed RNA polymerase II 13.6 kDa polypeptide

Chain K: 14% 61% 18% 5%



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

Chain L: 6% 21% 31% 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	167.08Å 221.26Å 193.69Å 90.00° 100.10° 90.00°	Depositor
Resolution (Å)	39.86 – 3.61 39.86 – 3.60	Depositor EDS
% Data completeness (in resolution range)	92.7 (39.86-3.61) 92.7 (39.86-3.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.13 (at 3.57Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.315 , 0.343 0.309 , 0.329	Depositor DCC
R_{free} test set	8031 reflections (10.04%)	wwPDB-VP
Wilson B-factor (Å ²)	47.6	Xtriage
Anisotropy	0.357	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , -1.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.76	EDS
Total number of atoms	28647	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality

5.1 Standard geometry

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

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

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	R	0.67	0/244	0.83	0/380
2	T	0.75	0/311	1.39	3/477 (0.6%)
3	A	0.88	15/11163 (0.1%)	1.45	167/15091 (1.1%)
4	B	0.84	7/8964 (0.1%)	1.38	114/12086 (0.9%)
5	C	0.78	0/2133	1.24	13/2891 (0.4%)
6	E	0.90	2/1788 (0.1%)	1.40	14/2406 (0.6%)
7	F	0.83	0/691	1.28	7/933 (0.8%)
8	H	0.85	1/1086 (0.1%)	1.59	20/1470 (1.4%)
9	I	1.03	2/989 (0.2%)	1.64	23/1331 (1.7%)
10	J	0.78	0/541	1.44	7/727 (1.0%)
11	K	0.74	0/937	1.20	5/1265 (0.4%)
12	L	0.99	1/366 (0.3%)	1.78	12/485 (2.5%)
All	All	0.86	28/29213 (0.1%)	1.41	385/39542 (1.0%)

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	T	0	5
3	A	1	6
4	B	0	7
5	C	0	2
6	E	0	1
9	I	0	1
All	All	1	22

The worst 5 of 28 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	319	GLY	C-O	10.97	1.41	1.23
3	A	255	SER	CA-CB	8.59	1.65	1.52
3	A	320	ARG	CA-CB	8.38	1.72	1.53
3	A	320	ARG	CG-CD	7.83	1.71	1.51
4	B	595	ARG	CG-CD	7.49	1.70	1.51

The worst 5 of 385 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	321	PRO	N-CA-C	-18.56	63.84	112.10
3	A	322	VAL	N-CA-C	14.54	150.26	111.00
10	J	10	CYS	CA-CB-SG	12.01	135.61	114.00
3	A	315	LEU	CA-CB-CG	11.82	142.50	115.30
4	B	478	GLY	N-CA-C	-11.76	83.71	113.10

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	322	VAL	CA

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	T	10	DT	Sidechain
2	T	11	DC	Sidechain
2	T	13	DA	Sidechain
2	T	6	DC	Sidechain
2	T	8	DT	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	18	0
2	T	279	0	160	43	0
3	A	10969	0	11070	2106	0
4	B	8793	0	8823	1592	0
5	C	2095	0	2051	337	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	E	1752	0	1776	303	0
7	F	679	0	701	127	0
8	H	1068	0	1040	193	0
9	I	971	0	929	164	0
10	J	532	0	542	125	0
11	K	919	0	929	175	0
12	L	364	0	387	65	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
14	A	1	0	0	0	0
All	All	28647	0	28518	4858	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 85.

The worst 5 of 4858 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:A:853:ASP:OD1	3:A:855:THR:HB	1.32	1.25
3:A:90:VAL:HG12	3:A:297:GLN:NE2	1.49	1.24
4:B:635:ARG:HB2	4:B:636:PRO:CD	1.65	1.21
3:A:321:PRO:O	3:A:322:VAL:HG22	1.41	1.18
3:A:351:THR:HG23	4:B:1103:ILE:HD12	1.23	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	
3	A	1383/1733 (80%)	851 (62%)	315 (23%)	217 (16%)	0	4
4	B	1088/1224 (89%)	730 (67%)	214 (20%)	144 (13%)	0	5
5	C	264/318 (83%)	187 (71%)	49 (19%)	28 (11%)	0	8
6	E	212/215 (99%)	142 (67%)	41 (19%)	29 (14%)	0	5
7	F	82/155 (53%)	49 (60%)	26 (32%)	7 (8%)	1	12
8	H	129/146 (88%)	87 (67%)	16 (12%)	26 (20%)	0	2
9	I	117/122 (96%)	74 (63%)	24 (20%)	19 (16%)	0	3
10	J	63/70 (90%)	42 (67%)	11 (18%)	10 (16%)	0	3
11	K	112/120 (93%)	81 (72%)	20 (18%)	11 (10%)	1	10
12	L	44/70 (63%)	20 (46%)	10 (23%)	14 (32%)	0	0
All	All	3494/4173 (84%)	2263 (65%)	726 (21%)	505 (14%)	0	4

5 of 505 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	44	THR
3	A	50	ILE
3	A	55	ASP
3	A	56	PRO
3	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	
3	A	1218/1520 (80%)	859 (70%)	359 (30%)	0	3
4	B	960/1061 (90%)	718 (75%)	242 (25%)	0	5
5	C	234/274 (85%)	175 (75%)	59 (25%)	0	5
6	E	196/197 (100%)	141 (72%)	55 (28%)	0	3
7	F	74/137 (54%)	57 (77%)	17 (23%)	1	6
8	H	117/128 (91%)	80 (68%)	37 (32%)	0	2

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
9	I	113/116 (97%)	84 (74%)	29 (26%)	0	5
10	J	60/65 (92%)	42 (70%)	18 (30%)	0	3
11	K	99/102 (97%)	81 (82%)	18 (18%)	2	12
12	L	40/57 (70%)	27 (68%)	13 (32%)	0	2
All	All	3111/3657 (85%)	2264 (73%)	847 (27%)	0	4

5 of 847 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	177	LYS
4	B	635	ARG
9	I	31	THR
4	B	234	ILE
4	B	408	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 101 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	46	GLN
4	B	515	HIS
9	I	12	ASN
4	B	121	ASN
4	B	366	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	9/10 (90%)	2 (22%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	5	A
1	R	10	A

There are no RNA pucker outliers to report.

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 9 ligands modelled in this entry, 9 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](#)

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%)	1.35	2 (20%) 1 0	30, 82, 200, 200	0
2	T	14/14 (100%)	1.04	4 (28%) 0 0	30, 107, 200, 200	0
3	A	1395/1733 (80%)	-0.15	26 (1%) 66 52	30, 32, 134, 200	0
4	B	1106/1224 (90%)	-0.23	11 (0%) 82 70	30, 30, 118, 198	0
5	C	266/318 (83%)	-0.35	0 100 100	30, 30, 84, 140	0
6	E	214/215 (99%)	0.28	15 (7%) 16 10	30, 67, 153, 200	0
7	F	84/155 (54%)	-0.08	0 100 100	30, 31, 93, 141	0
8	H	133/146 (91%)	-0.08	3 (2%) 60 45	30, 47, 138, 190	0
9	I	119/122 (97%)	-0.09	0 100 100	30, 32, 114, 163	0
10	J	65/70 (92%)	-0.41	0 100 100	30, 30, 96, 141	0
11	K	114/120 (95%)	-0.36	0 100 100	30, 30, 73, 107	0
12	L	46/70 (65%)	-0.24	1 (2%) 62 46	30, 45, 132, 158	0
All	All	3566/4197 (84%)	-0.16	62 (1%) 70 56	30, 31, 130, 200	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	866	TYR	5.8
3	A	316	GLN	5.0
2	T	1	DA	4.8
3	A	1175	SER	4.5
3	A	149	GLU	4.4

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
14	MG	A	2000	1/1	0.69	0.20	22,22,22,22	0
13	ZN	A	1734	1/1	0.84	0.29	22,22,22,22	0
13	ZN	A	1735	1/1	0.87	0.10	22,22,22,22	0
13	ZN	B	1307	1/1	0.91	0.11	22,22,22,22	0
13	ZN	I	203	1/1	0.92	0.21	22,22,22,22	0
13	ZN	J	101	1/1	0.93	0.11	22,22,22,22	0
13	ZN	I	204	1/1	0.96	0.17	22,22,22,22	0
13	ZN	C	319	1/1	0.98	0.06	22,22,22,22	0
13	ZN	L	105	1/1	0.98	0.06	22,22,22,22	0

6.5 Other polymers [i](#)

There are no such residues in this entry.