



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 21, 2017 – 10:51 PM EDT

PDB ID : 5M5W
EMDB ID: : EMD-3446
Title : RNA Polymerase I open complex
Authors : Tafur, L.; Sadian, Y.; Hoffmann, N.A.; Jakobi, A.J.; Wetzell, R.; Hagen, W.J.H.; Sachse, C.; Muller, C.W.
Deposited on : unknown
Resolution : 3.80 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

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

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

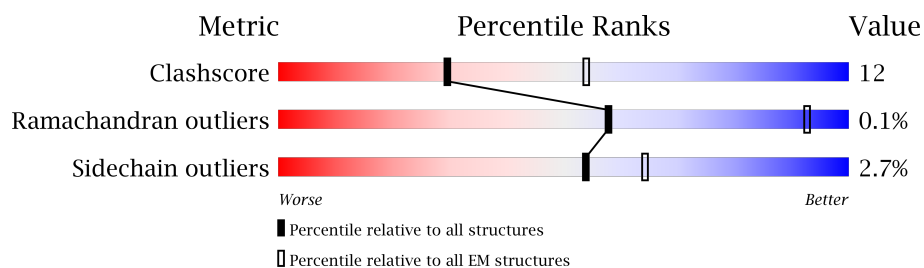
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.80 Å.

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










Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	1664	65% 24% • 10%
2	B	1203	67% 30% ..
3	C	335	72% 19% 9%
4	D	137	30% 12% • 57%
5	E	215	73% 24% •
6	F	155	45% 20% 35%
7	G	326	44% 18% 37%
8	H	146	66% 23% • 8%
9	I	125	70% 20% • 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	70	
11	K	142	
12	L	70	
13	M	415	
14	N	233	
15	S	70	
16	T	70	

2 Entry composition

There are 17 unique types of molecules in this entry. The entry contains 34780 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 I subunit RPA190.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1493	Total	C	N	O	S	0	0
			11760	7424	2048	2226	62		

- Molecule 2 is a protein called DNA-directed RNA polymerase I subunit RPA135.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1183	Total	C	N	O	S	0	0
			9389	5936	1642	1760	51		

- Molecule 3 is a protein called DNA-directed RNA polymerases I and III subunit RPAC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	304	Total	C	N	O	S	0	0
			2418	1536	414	460	8		

- Molecule 4 is a protein called DNA-directed RNA polymerase I subunit RPA14.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	D	59	Total	C	N	O	0	0
			467	293	80	94		

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	214	Total	C	N	O	S	0	0
			1751	1111	309	320	11		

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	100	Total	C	N	O	S	0	0
			823	522	144	154	3		

- Molecule 7 is a protein called DNA-directed RNA polymerase I subunit RPA43.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	205	Total	C	N	O	S	0	0
			1624	1040	280	299	5		

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	134	Total	C	N	O	S	0	0
			1075	677	182	212	4		

- Molecule 9 is a protein called DNA-directed RNA polymerase I subunit RPA12.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	114	Total	C	N	O	S	0	0
			869	541	146	173	9		

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	69	Total	C	N	O	S	0	0
			569	362	101	100	6		

- Molecule 11 is a protein called DNA-directed RNA polymerases I and III subunit RPAC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	100	Total	C	N	O	S	0	0
			785	491	129	160	5		

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	43	Total	C	N	O	S	0	0
			344	211	69	60	4		

- Molecule 13 is a protein called DNA-directed RNA polymerase I subunit RPA49.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	107	Total	C	N	O		0	0
			850	540	141	169			

- Molecule 14 is a protein called DNA-directed RNA polymerase I subunit RPA34.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	146	Total	C	N	O	S	0	0
			1164	743	188	229	4		

- Molecule 15 is a DNA chain called Non-template strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	S	18	Total	C	N	O	P	0	0
			376	178	77	103	18		

- Molecule 16 is a DNA chain called Template strand.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	T	25	Total	C	N	O	P	0	0
			509	244	86	154	25		

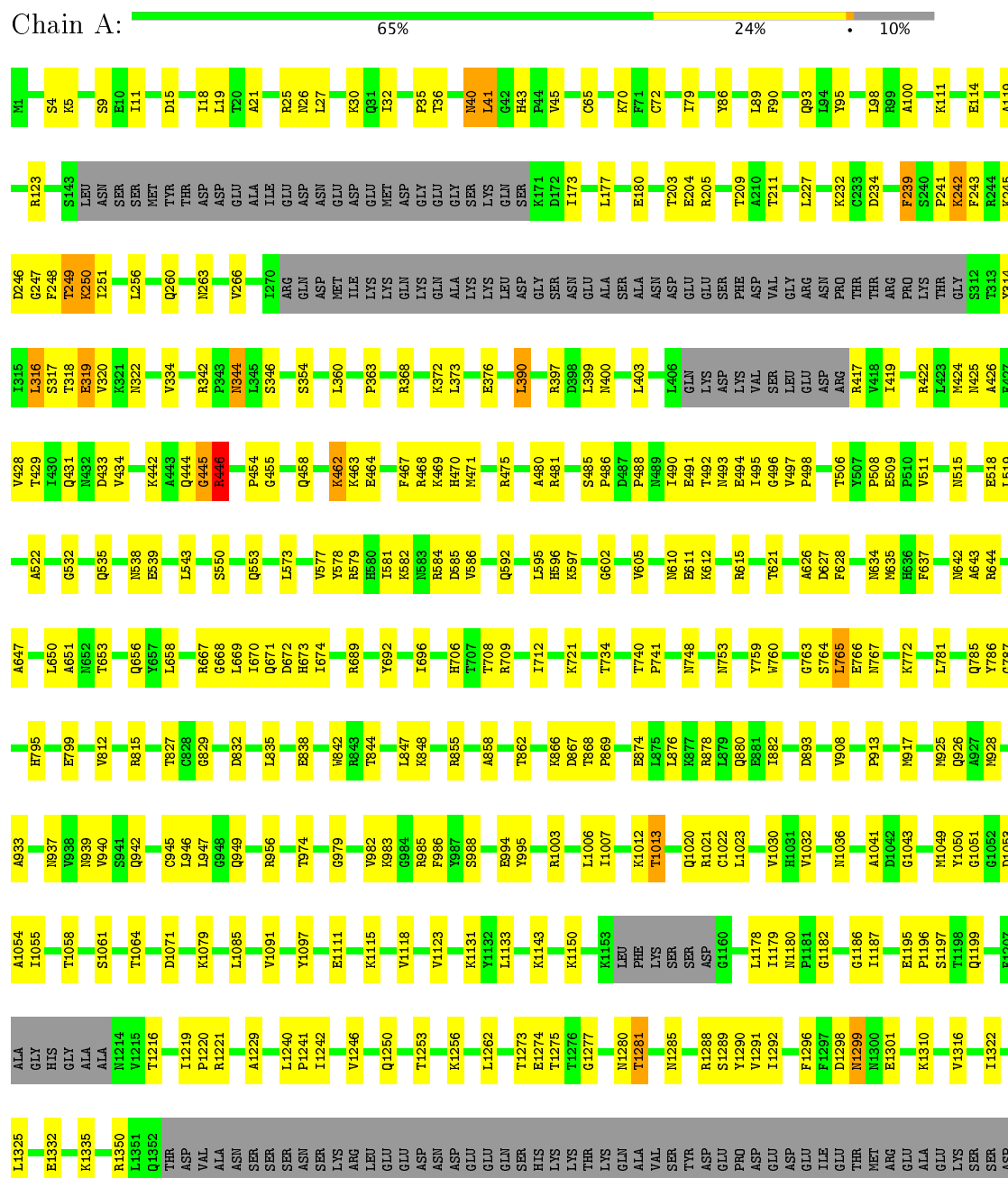
- Molecule 17 is ZINC ION (three-letter code: ZN) (formula: Zn).

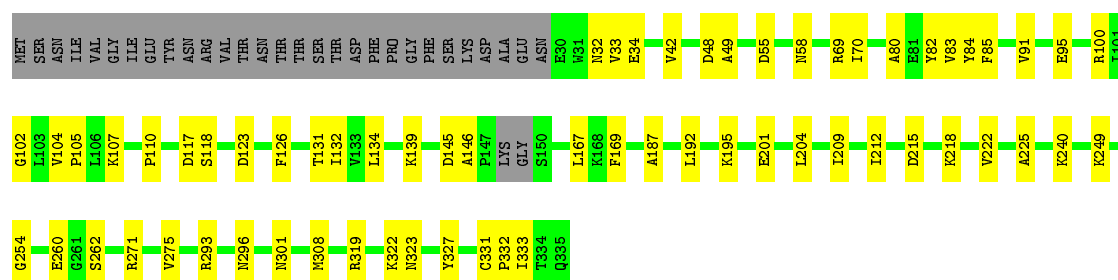
Mol	Chain	Residues	Atoms		AltConf
17	B	1	Total	Zn	0
			1	1	
17	A	2	Total	Zn	0
			2	2	
17	L	1	Total	Zn	0
			1	1	
17	J	1	Total	Zn	0
			1	1	
17	I	2	Total	Zn	0
			2	2	

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.

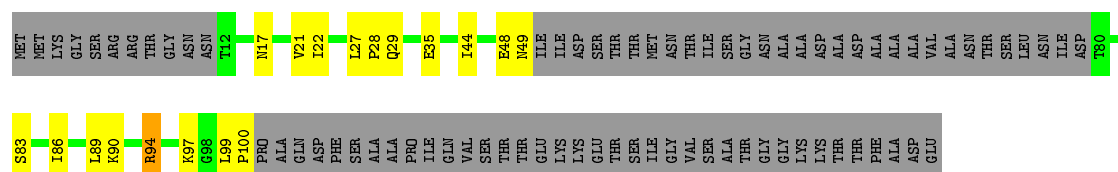
- Molecule 1: DNA-directed RNA polymerase I subunit RPA190





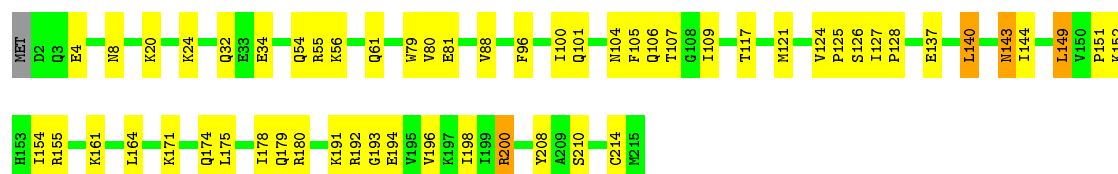
- Molecule 4: DNA-directed RNA polymerase I subunit RPA14

Chain D: 30% 12% 57%



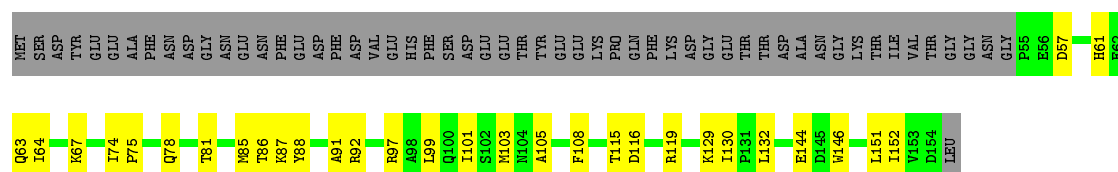
- Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E: 73% 24%



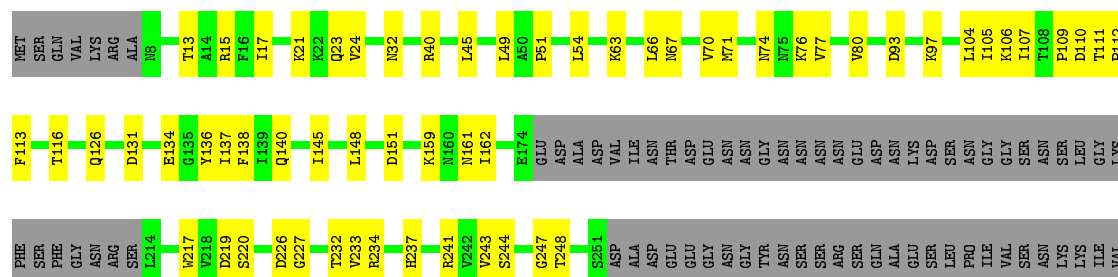
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 45% 20% 35%



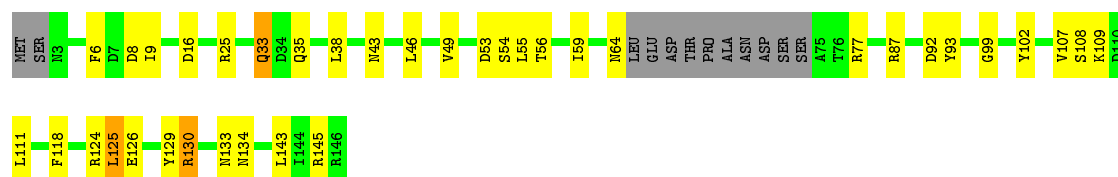
- Molecule 7: DNA-directed RNA polymerase I subunit RPA43

Chain G: 44% 18% 37%



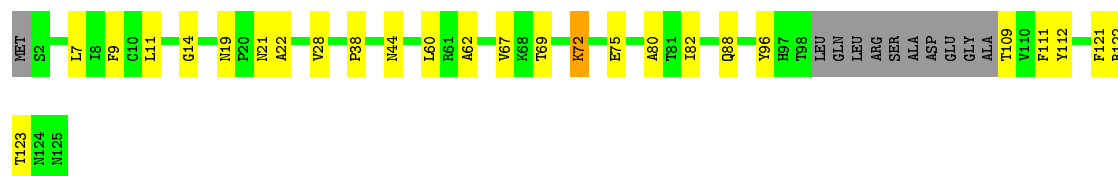
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:  66% 23% 8%



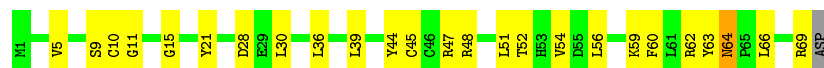
- Molecule 9: DNA-directed RNA polymerase I subunit RPA12

Chain I:  70% 20% 9%



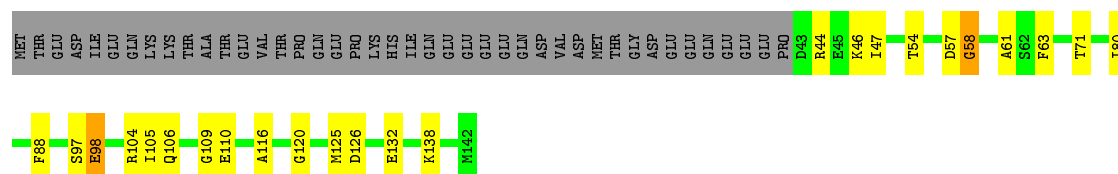
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:  63% 34% ...



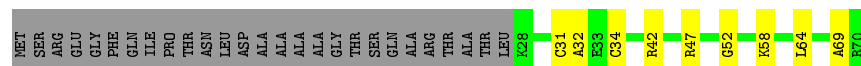
- Molecule 11: DNA-directed RNA polymerases I and III subunit RPAC2

Chain K:  54% 15% 30%



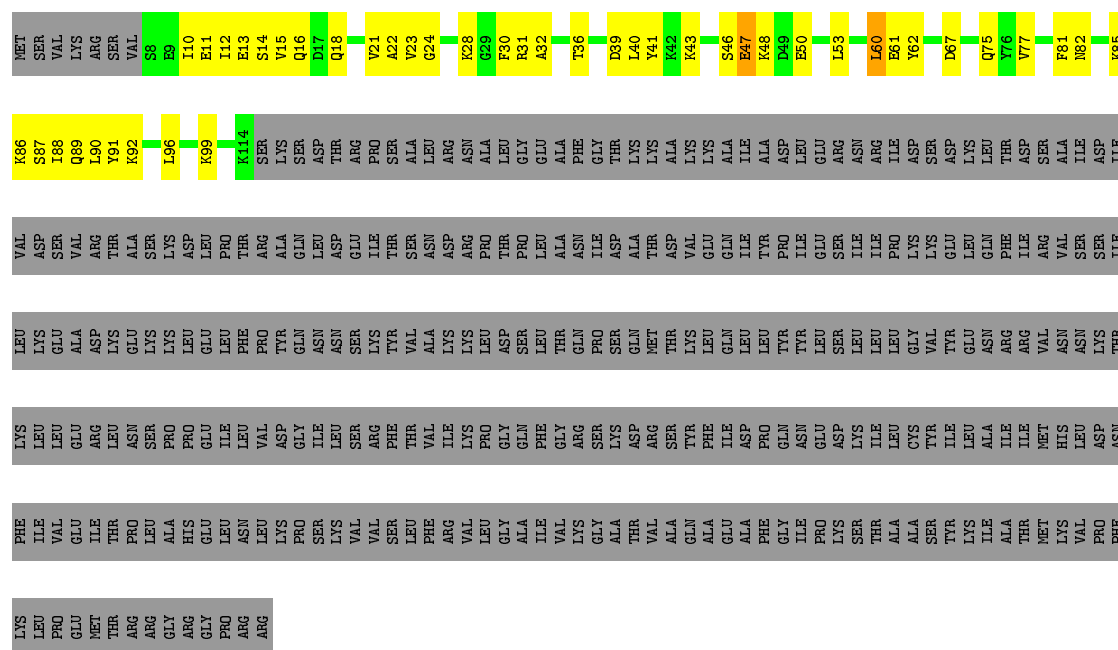
- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4

Chain L: 49% 13% 39%

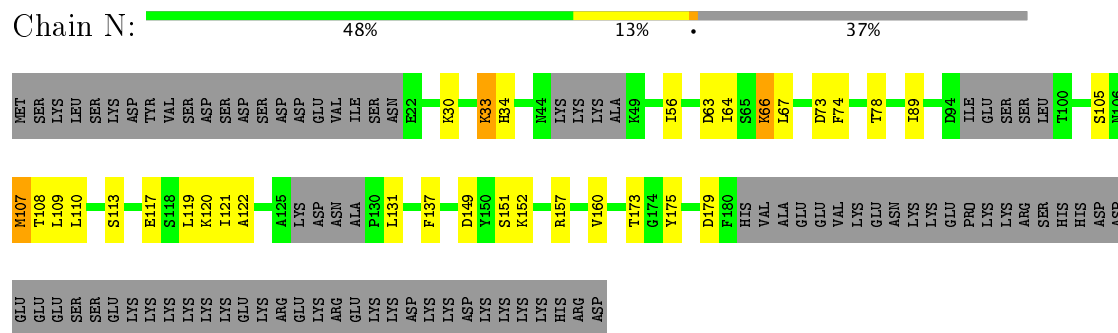


- Molecule 13: DNA-directed RNA polymerase I subunit RPA49

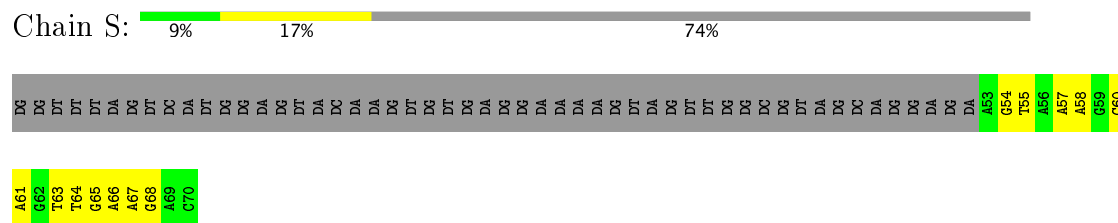
Chain M:  15% 10% 74%



• Molecule 14: DNA-directed RNA polymerase I subunit RPA34



• Molecule 15: Non-template strand



• Molecule 16: Template strand



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	98430	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	40	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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 >2	RMSZ	# Z >2
1	A	0.25	1/11976 (0.0%)	0.43	3/16180 (0.0%)
10	J	0.22	0/578	0.40	0/775
11	K	0.50	3/795 (0.4%)	0.81	5/1072 (0.5%)
12	L	0.22	0/346	0.45	0/457
13	M	0.24	0/866	0.46	0/1162
14	N	0.24	0/1186	0.44	0/1596
15	S	0.47	0/424	0.81	0/653
16	T	0.48	0/568	0.94	0/874
2	B	0.24	0/9596	0.46	4/12971 (0.0%)
3	C	0.24	0/2469	0.41	0/3347
4	D	0.23	0/473	0.44	0/641
5	E	0.24	0/1787	0.41	0/2406
6	F	0.22	0/838	0.38	0/1129
7	G	0.24	0/1662	0.44	0/2260
8	H	0.24	0/1093	0.48	0/1480
9	I	0.24	0/881	0.47	0/1187
All	All	0.26	4/35538 (0.0%)	0.47	12/48190 (0.0%)

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
11	K	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	K	57	ASP	N-CA	10.18	1.66	1.46
1	A	250	LYS	CA-C	-5.97	1.37	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	K	57	ASP	CA-CB	5.24	1.65	1.53
11	K	58	GLY	CA-C	-5.23	1.43	1.51

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	K	57	ASP	N-CA-C	-14.03	73.12	111.00
11	K	57	ASP	CB-CA-C	-11.85	86.70	110.40
2	B	1064	LYS	N-CA-C	-11.72	79.37	111.00
1	A	444	GLN	C-N-CA	-8.56	104.32	122.30
2	B	359	LEU	CA-CB-CG	8.24	134.25	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
11	K	58	GLY	Mainchain

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	11760	0	11838	303	0
2	B	9389	0	9269	280	0
3	C	2418	0	2401	47	0
4	D	467	0	468	13	0
5	E	1751	0	1776	37	0
6	F	823	0	841	23	0
7	G	1624	0	1625	52	0
8	H	1075	0	1045	23	0
9	I	869	0	862	22	0
10	J	569	0	585	20	0
11	K	785	0	782	18	0
12	L	344	0	363	9	0
13	M	850	0	850	34	0
14	N	1164	0	1160	27	0
15	S	376	0	202	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	T	509	0	285	11	0
17	A	2	0	0	0	0
17	B	1	0	0	0	0
17	I	2	0	0	0	0
17	J	1	0	0	0	0
17	L	1	0	0	0	0
All	All	34780	0	34352	795	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 12.

The worst 5 of 795 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:1140:LYS:HE2	7:G:17:ILE:CD1	1.46	1.44
1:A:248:PHE:CE1	1:A:442:LYS:NZ	2.04	1.26
2:B:1140:LYS:CE	7:G:17:ILE:HD11	1.75	1.15
1:A:248:PHE:CD1	1:A:442:LYS:NZ	2.17	1.10
10:J:10:CYS:HB3	10:J:45:CYS:SG	1.94	1.07

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 PDB entries followed by that with respect to all EM entries.

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	1479/1664 (89%)	1413 (96%)	62 (4%)	4 (0%)	44	81
2	B	1177/1203 (98%)	1118 (95%)	59 (5%)	0	100	100
3	C	300/335 (90%)	292 (97%)	8 (3%)	0	100	100
4	D	55/137 (40%)	55 (100%)	0	0	100	100
5	E	212/215 (99%)	206 (97%)	6 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	F	98/155 (63%)	96 (98%)	2 (2%)	0	100	100
7	G	201/326 (62%)	193 (96%)	8 (4%)	0	100	100
8	H	130/146 (89%)	125 (96%)	5 (4%)	0	100	100
9	I	110/125 (88%)	105 (96%)	5 (4%)	0	100	100
10	J	67/70 (96%)	63 (94%)	4 (6%)	0	100	100
11	K	98/142 (69%)	92 (94%)	6 (6%)	0	100	100
12	L	41/70 (59%)	40 (98%)	1 (2%)	0	100	100
13	M	105/415 (25%)	99 (94%)	5 (5%)	1 (1%)	18	61
14	N	138/233 (59%)	129 (94%)	9 (6%)	0	100	100
All	All	4211/5236 (80%)	4026 (96%)	180 (4%)	5 (0%)	58	88

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	765	LEU
1	A	250	LYS
1	A	319	GLU
13	M	47	GLU
1	A	445	GLY

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 PDB entries followed by that with respect to all EM entries.

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	1313/1465 (90%)	1276 (97%)	37 (3%)	49	76
2	B	1034/1053 (98%)	1008 (98%)	26 (2%)	53	79
3	C	269/296 (91%)	266 (99%)	3 (1%)	78	89
4	D	56/116 (48%)	52 (93%)	4 (7%)	17	55
5	E	196/197 (100%)	189 (96%)	7 (4%)	40	72
6	F	90/137 (66%)	90 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	G	183/291 (63%)	176 (96%)	7 (4%)	38	70
8	H	117/128 (91%)	111 (95%)	6 (5%)	28	64
9	I	102/110 (93%)	100 (98%)	2 (2%)	60	83
10	J	64/65 (98%)	61 (95%)	3 (5%)	30	66
11	K	90/130 (69%)	89 (99%)	1 (1%)	78	89
12	L	38/57 (67%)	37 (97%)	1 (3%)	51	78
13	M	97/371 (26%)	95 (98%)	2 (2%)	59	82
14	N	136/220 (62%)	132 (97%)	4 (3%)	48	76
All	All	3785/4636 (82%)	3682 (97%)	103 (3%)	54	77

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	441	LYS
2	B	946	ASP
11	K	98	GLU
2	B	478	LEU
2	B	721	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 65 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	423	ASN
2	B	718	GLN
11	K	106	GLN
2	B	532	HIS
2	B	702	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 7 ligands modelled in this entry, 7 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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.