



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Dec 20, 2016 – 12:34 PM EST

PDB ID : 5M3F
EMDB ID: : EMD-4147
Title : Yeast RNA polymerase I elongation complex at 3.8Å
Authors : Neyer, S.; Kunz, M.; Geiss, C.; Hantsche, M.; Hodirnau, V.-V.; Seybert, A.;
Engel, C.; Scheffer, M.P.; Cramer, P.; Frangakis, A.S.
Deposited on : 2016-10-14
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
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20028442

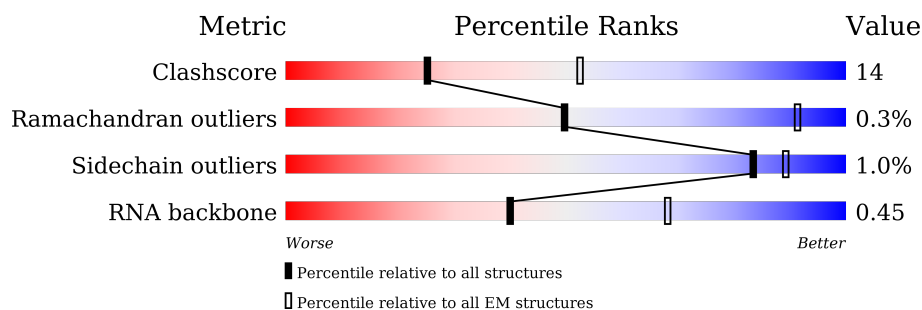
1 Overall quality at a glance

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.









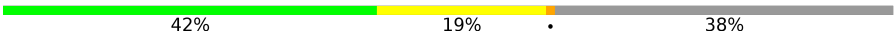

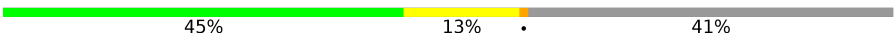
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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% 22% 12%
2	B	1203	71% 27% .
3	C	335	72% 19% 9%
4	E	215	81% 17% .
5	F	155	49% 15% 35%
6	H	146	71% 19% 10%
7	I	125	30% 20% 50%
8	J	70	76% 23% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	K	142	
10	L	70	
11	T	39	
12	U	39	
13	R	20	
14	M	415	
15	N	233	
16	D	137	
17	G	326	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
20	SO4	B	1301	-	-	X	-

2 Entry composition

There are 20 unique types of molecules in this entry. The entry contains 34032 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	1459	Total	C	N	O	S	0	0
			11526	7281	2004	2180	61		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1177	Total	C	N	O	S	0	0
			9350	5913	1639	1747	51		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	305	Total	C	N	O	S	0	0
			2423	1539	416	460	8		

- Molecule 4 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	212	Total	C	N	O	S	0	0
			1735	1102	306	316	11		

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

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

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	H	131	Total	C	N	O	S	0	0
			1052	664	176	208	4		

- Molecule 7 is a protein called DNA-directed RNA polymerase I subunit RPA12.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	I	63	Total	C	N	O	S	0	0
			466	292	77	93	4		

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 9 is a protein called DNA-directed RNA polymerases I and III subunit RPAC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	K	101	Total	C	N	O	S	0	0
			793	496	130	162	5		

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	L	44	Total	C	N	O	S	0	0
			352	217	70	61	4		

- Molecule 11 is a DNA chain called template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	T	25	Total	C	N	O	P	0	0
			509	244	95	146	24		

- Molecule 12 is a DNA chain called non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	U	14	Total	C	N	O	P	0	0
			285	138	51	83	13		

- Molecule 13 is a RNA chain called RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	R	8	Total	C	N	O	P	0	0
			173	77	34	54	8		

- Molecule 14 is a protein called DNA-directed RNA polymerase I subunit RPA49.

Mol	Chain	Residues	Atoms				AltConf	Trace
14	M	108	Total	C	N	O	0	0
			856	543	142	171		

- Molecule 15 is a protein called DNA-directed RNA polymerase I subunit RPA34.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	N	145	Total	C	N	O	S	0	0
			1151	735	188	224	4		

- Molecule 16 is a protein called DNA-directed RNA polymerase I subunit RPA14.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	D	54	Total	C	N	O	0	0
			431	270	73	88		

- Molecule 17 is a protein called DNA-directed RNA polymerase I subunit RPA43.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	G	193	Total	C	N	O	S	0	0
			1526	985	262	274	5		

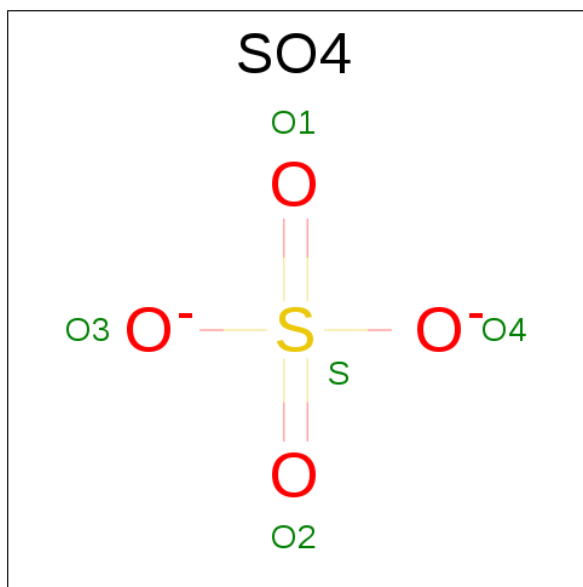
- Molecule 18 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 19 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
19	A	1	Total	Mg	0
			1	1	

- Molecule 20 is SULFATE ION (three-letter code: SO4) (formula: O₄S).

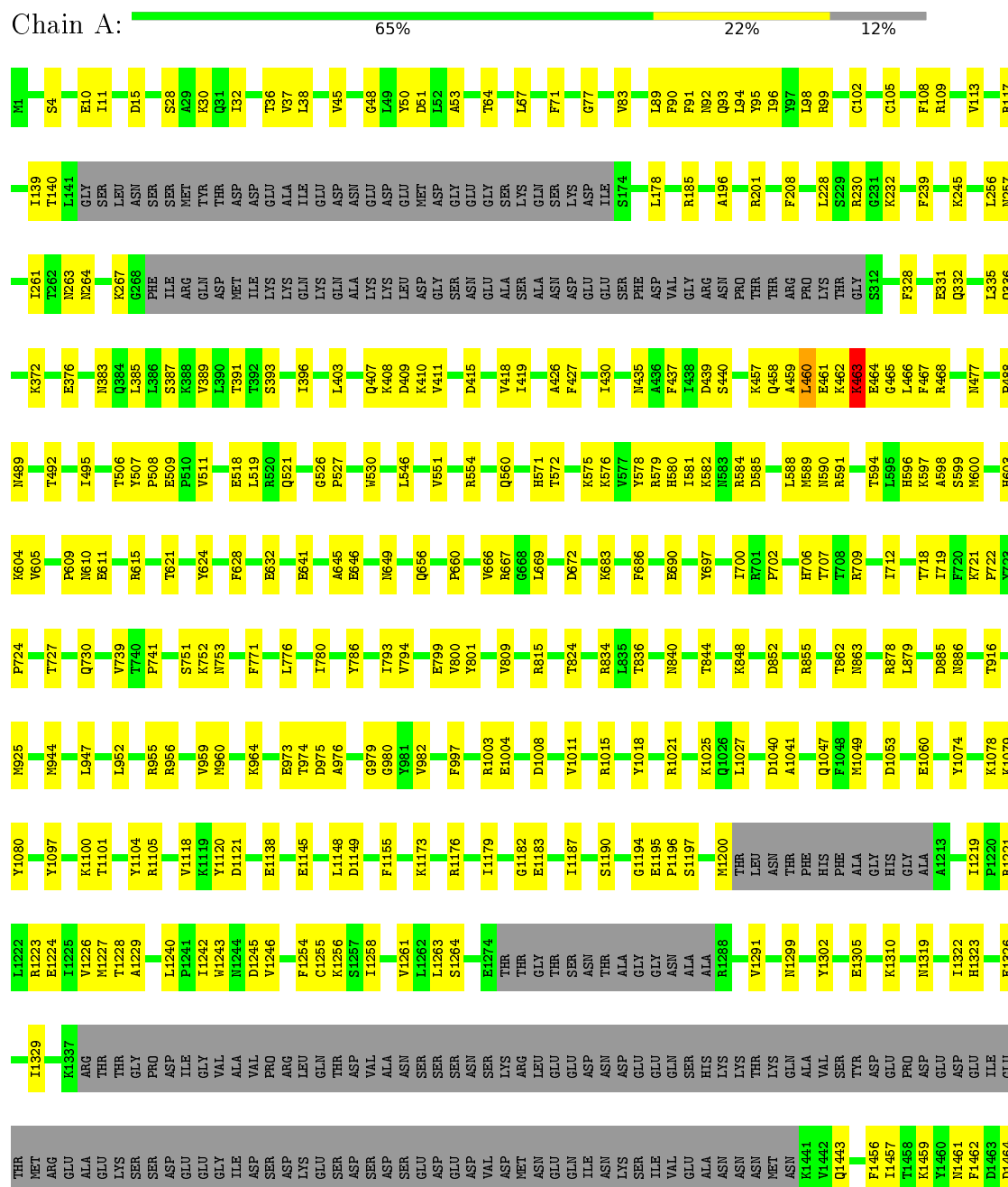


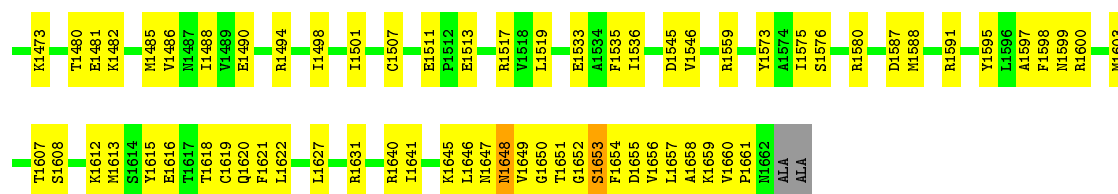
Mol	Chain	Residues	Atoms			AltConf
			Total	O	S	
20	B	1	5	4	1	0

3 Residue-property plots

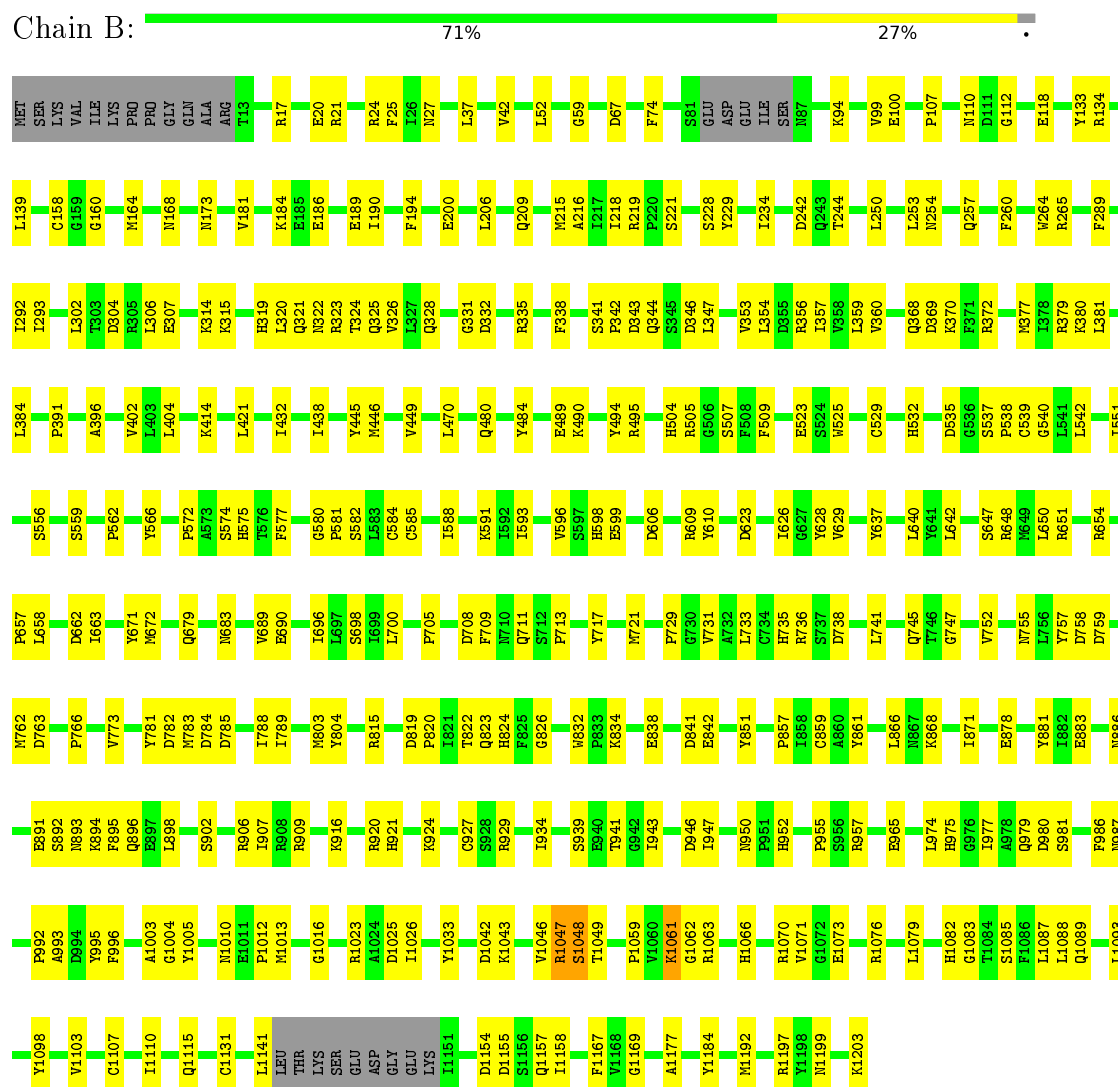
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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

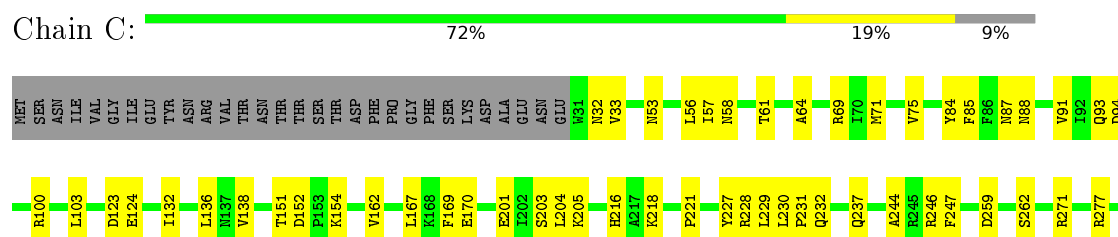




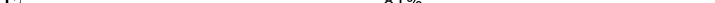
• Molecule 2: DNA-directed RNA polymerase I subunit RPA135

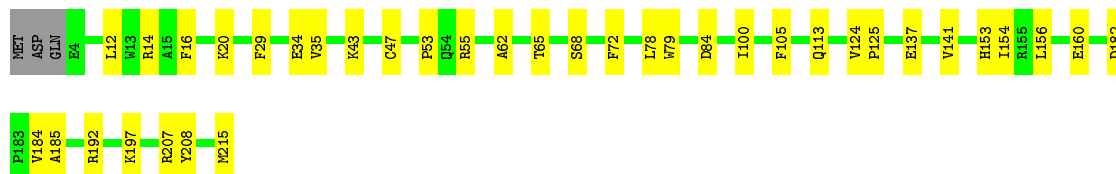


• Molecule 3: DNA-directed RNA polymerases I and III subunit RPAC1



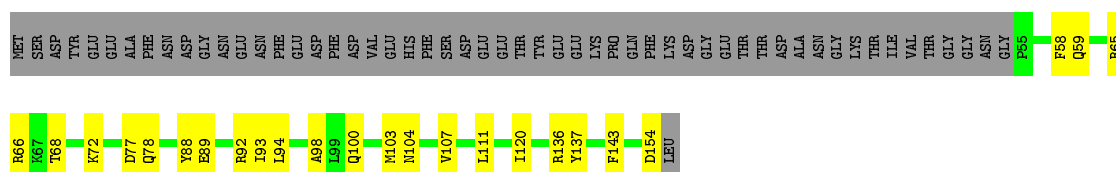
- Molecule 4: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:  81% 17% .



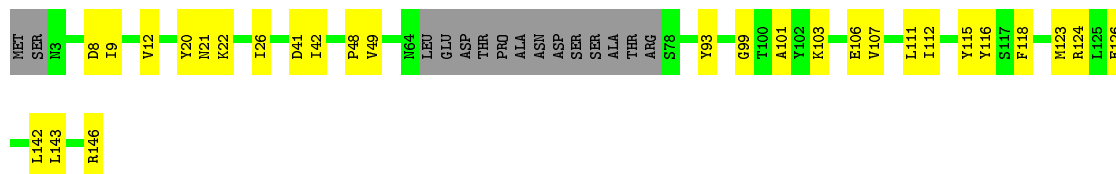
- Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F:  49% 15% 35%

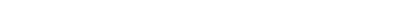


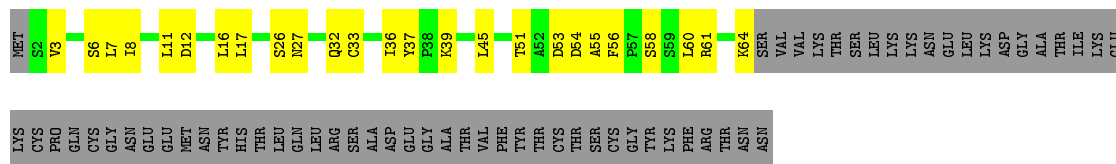
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H: 71% 19% 10%



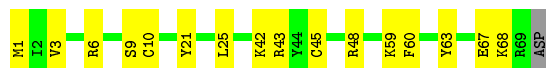
- Molecule 7: DNA-directed RNA polymerase I subunit RPA12

Chain I:  30% 20% 50%

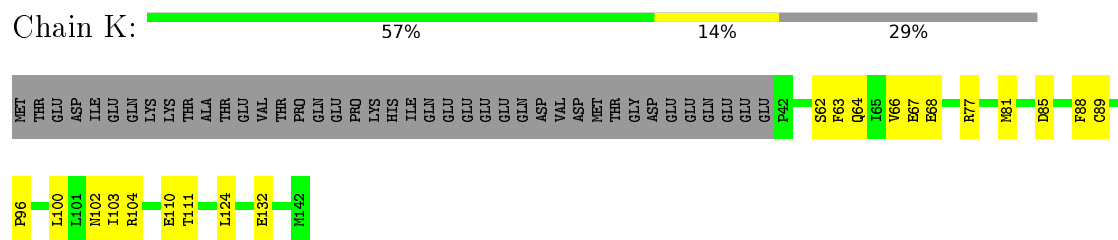


- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC5

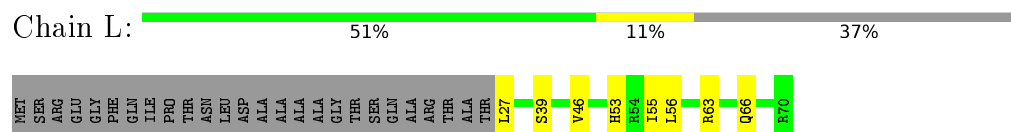
Chain J:  76% 23%



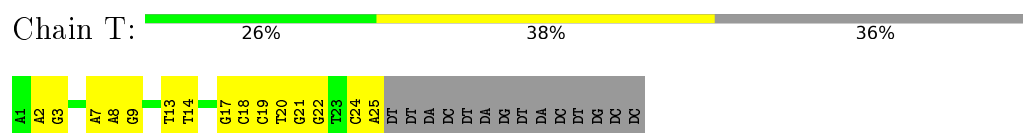
- Molecule 9: DNA-directed RNA polymerases I and III subunit RPAC2



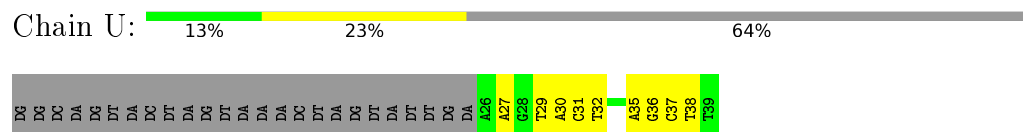
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC4



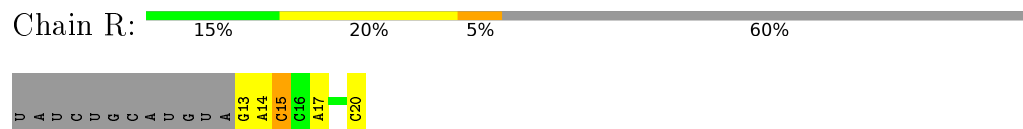
- Molecule 11: template DNA



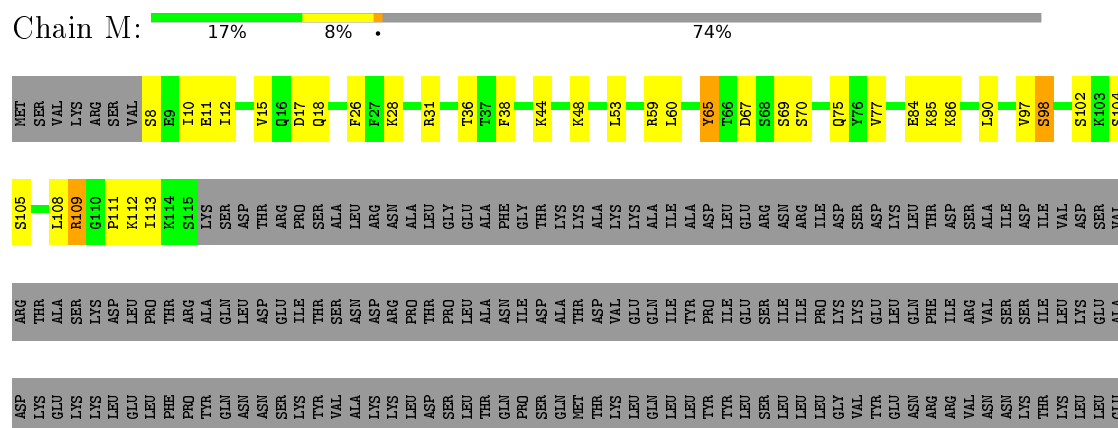
- Molecule 12: non-template DNA



- Molecule 13: RNA



- Molecule 14: DNA-directed RNA polymerase I subunit RPA49



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	94000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

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: MG, ZN, SO4

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	0.35	0/11738	0.50	0/15851
10	L	0.33	0/354	0.50	0/468
11	T	0.60	0/571	0.87	0/879
12	U	0.58	0/319	0.97	0/491
13	R	0.48	0/193	1.08	0/299
14	M	0.41	0/872	0.55	0/1170
15	N	0.42	0/1172	0.57	1/1580 (0.1%)
16	D	0.38	0/436	0.51	0/591
17	G	0.37	0/1564	0.66	3/2127 (0.1%)
2	B	0.37	0/9557	0.52	0/12918
3	C	0.35	0/2475	0.49	0/3354
4	E	0.31	0/1771	0.46	0/2383
5	F	0.32	0/838	0.46	0/1129
6	H	0.34	0/1070	0.49	0/1449
7	I	0.26	0/472	0.48	0/639
8	J	0.37	0/578	0.47	0/775
9	K	0.35	0/804	0.51	0/1083
All	All	0.37	0/34784	0.54	4/47186 (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
2	B	0	3

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
17	G	241	ARG	NE-CZ-NH1	11.89	126.25	120.30
17	G	241	ARG	NE-CZ-NH2	-11.70	114.45	120.30
17	G	241	ARG	CD-NE-CZ	6.04	132.05	123.60
15	N	145	ILE	C-N-CD	5.04	138.99	128.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	1048	SER	Peptide
2	B	1061	LYS	Peptide
2	B	319	HIS	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	11526	0	11614	326	0
2	B	9350	0	9240	362	0
3	C	2423	0	2412	68	0
4	E	1735	0	1764	25	0
5	F	823	0	840	46	0
6	H	1052	0	1021	20	0
7	I	466	0	466	29	0
8	J	569	0	585	21	0
9	K	793	0	790	17	0
10	L	352	0	376	7	0
11	T	509	0	283	17	0
12	U	285	0	161	9	0
13	R	173	0	89	7	0
14	M	856	0	852	97	0
15	N	1151	0	1169	99	0
16	D	431	0	428	11	0
17	G	1526	0	1540	90	0
18	A	2	0	0	0	0
18	B	1	0	0	0	0
18	I	1	0	0	0	0
18	J	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
18	L	1	0	0	0	0
19	A	1	0	0	0	0
20	B	5	0	0	2	0
All	All	34032	0	33630	919	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 14.

The worst 5 of 919 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:328:GLN:NE2	14:M:112:LYS:HA	1.22	1.49
1:A:1661:PRO:CD	17:G:55:GLU:CA	1.86	1.44
2:B:987:ASN:ND2	15:N:157:ARG:CD	1.85	1.39
2:B:328:GLN:CD	14:M:112:LYS:HA	1.39	1.39
3:C:301:ASN:ND2	15:N:173:THR:HB	1.37	1.37

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all 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	1447/1664 (87%)	1325 (92%)	119 (8%)	3 (0%)	52	86
2	B	1171/1203 (97%)	1109 (95%)	60 (5%)	2 (0%)	52	86
3	C	303/335 (90%)	286 (94%)	17 (6%)	0	100	100
4	E	210/215 (98%)	200 (95%)	10 (5%)	0	100	100
5	F	98/155 (63%)	96 (98%)	2 (2%)	0	100	100
6	H	127/146 (87%)	124 (98%)	3 (2%)	0	100	100
7	I	61/125 (49%)	53 (87%)	8 (13%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
8	J	67/70 (96%)	63 (94%)	4 (6%)	0	100	100
9	K	99/142 (70%)	93 (94%)	6 (6%)	0	100	100
10	L	42/70 (60%)	37 (88%)	5 (12%)	0	100	100
14	M	106/415 (26%)	96 (91%)	8 (8%)	2 (2%)	10	54
15	N	139/233 (60%)	122 (88%)	14 (10%)	3 (2%)	8	52
16	D	50/137 (36%)	48 (96%)	2 (4%)	0	100	100
17	G	189/326 (58%)	174 (92%)	13 (7%)	2 (1%)	17	65
All	All	4109/5236 (78%)	3826 (93%)	271 (7%)	12 (0%)	50	83

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	463	LYS
1	A	1653	SER
1	A	91	PHE
2	B	1047	ARG
14	M	85	LYS

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	1289/1465 (88%)	1284 (100%)	5 (0%)	93	97
2	B	1030/1053 (98%)	1030 (100%)	0	100	100
3	C	269/296 (91%)	269 (100%)	0	100	100
4	E	194/197 (98%)	194 (100%)	0	100	100
5	F	90/137 (66%)	90 (100%)	0	100	100
6	H	115/128 (90%)	115 (100%)	0	100	100
7	I	55/110 (50%)	55 (100%)	0	100	100
8	J	64/65 (98%)	64 (100%)	0	100	100
9	K	91/130 (70%)	91 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
10	L	39/57 (68%)	39 (100%)	0	100	100
14	M	98/371 (26%)	88 (90%)	10 (10%)	9	43
15	N	135/220 (61%)	129 (96%)	6 (4%)	35	73
16	D	52/116 (45%)	47 (90%)	5 (10%)	10	46
17	G	171/291 (59%)	159 (93%)	12 (7%)	19	60
All	All	3692/4636 (80%)	3654 (99%)	38 (1%)	83	92

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
15	N	135	LYS
16	D	15	THR
17	G	230	ARG
15	N	167	LYS
16	D	29	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	715	ASN
2	B	1082	HIS
10	L	53	HIS
2	B	975	HIS
2	B	1089	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
13	R	7/20 (35%)	2 (28%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
13	R	15	C
13	R	17	A

There are no RNA pucker outliers to report.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 7 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	$\# Z > 2$
20	SO4	B	1301	-	4,4,4	0.36	0	6,6,6	0.12	0

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
20	SO4	B	1301	-	-	0/0/0/0	0/0/0/0

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.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
20	B	1301	SO4	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.