



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:17 PM BST

PDB ID : 5FLM
EMDB ID: : EMD-3218
Title : Structure of transcribing mammalian RNA polymerase II
Authors : Bernecky, C.; Herzog, F.; Baumeister, W.; Plitzko, J.M.; Cramer, P.
Deposited on : 2015-10-26
Resolution : 3.40 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
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) : trunk27241

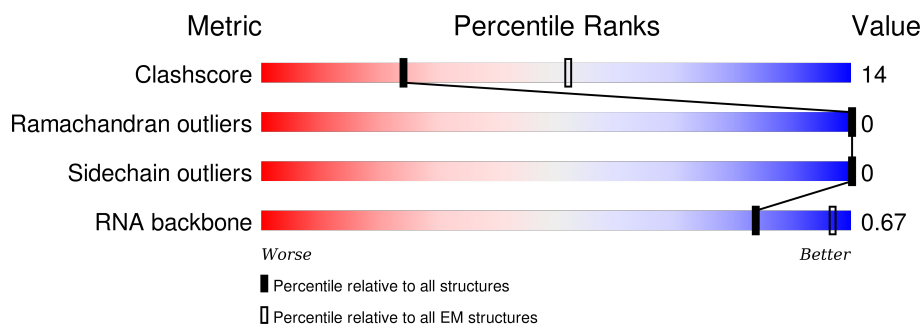
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.40 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.







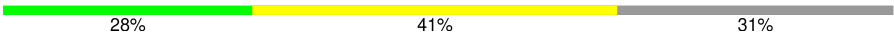

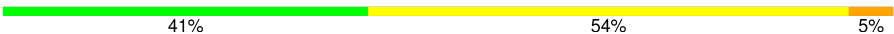
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686
RNA backbone	3027	244

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	1970	50% 22% 28%
2	B	1174	63% 33% .
3	C	275	64% 29% 7%
4	D	142	47% 43% 10%
5	E	210	73% 27%
6	F	127	49% 16% 35%
7	G	172	53% 46% .
8	H	150	67% 31% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
9	I	125	
10	J	67	
11	K	117	
12	L	58	
13	N	39	
14	P	20	
15	T	39	

2 Entry composition

There are 17 unique types of molecules in this entry. The entry contains 32712 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.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1427	Total	C	N	O	S	0	0
			11291	7102	2021	2096	72		

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1134	Total	C	N	O	S	0	0
			9062	5732	1595	1671	64		

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	257	Total	C	N	O	S	0	0
			2059	1294	351	408	6		

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	128	Total	C	N	O	S	0	0
			1005	632	172	197	4		

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	209	Total	C	N	O	S	0	0
			1711	1084	300	319	8		

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC2.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	82	Total	C	N	O	S	0	0
			658	419	113	121	5		

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	171	Total	C	N	O	S	0	0
			1334	867	216	243	8		

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC3.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	148	Total	C	N	O	S	0	0
			1186	750	194	237	5		

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	114	Total	C	N	O	S	0	0
			927	571	166	179	11		

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC5.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	J	67	Total	C	N	O	S	0	0
			533	345	90	92	6		

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	115	Total	C	N	O	S	0	0
			920	593	152	173	2		

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	44	Total	C	N	O	S	0	0
			372	231	72	63	6		

- Molecule 13 is a DNA chain called DNA, DNA-RNA ELONGATION SCAFFOLD.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	N	27	Total	C	N	O	P	0	0
			552	266	103	158	25		

- Molecule 14 is a RNA chain called RNA, DNA-RNA ELONGATION SCAFFOLD.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	P	14	Total	C	N	O	P	0	0
			301	135	59	93	14		

- Molecule 15 is a DNA chain called DNA, DNA-RNA ELONGATION SCAFFOLD.

Mol	Chain	Residues	Atoms					AltConf	Trace
15	T	39	Total	C	N	O	P	0	0
			792	380	142	232	38		

- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

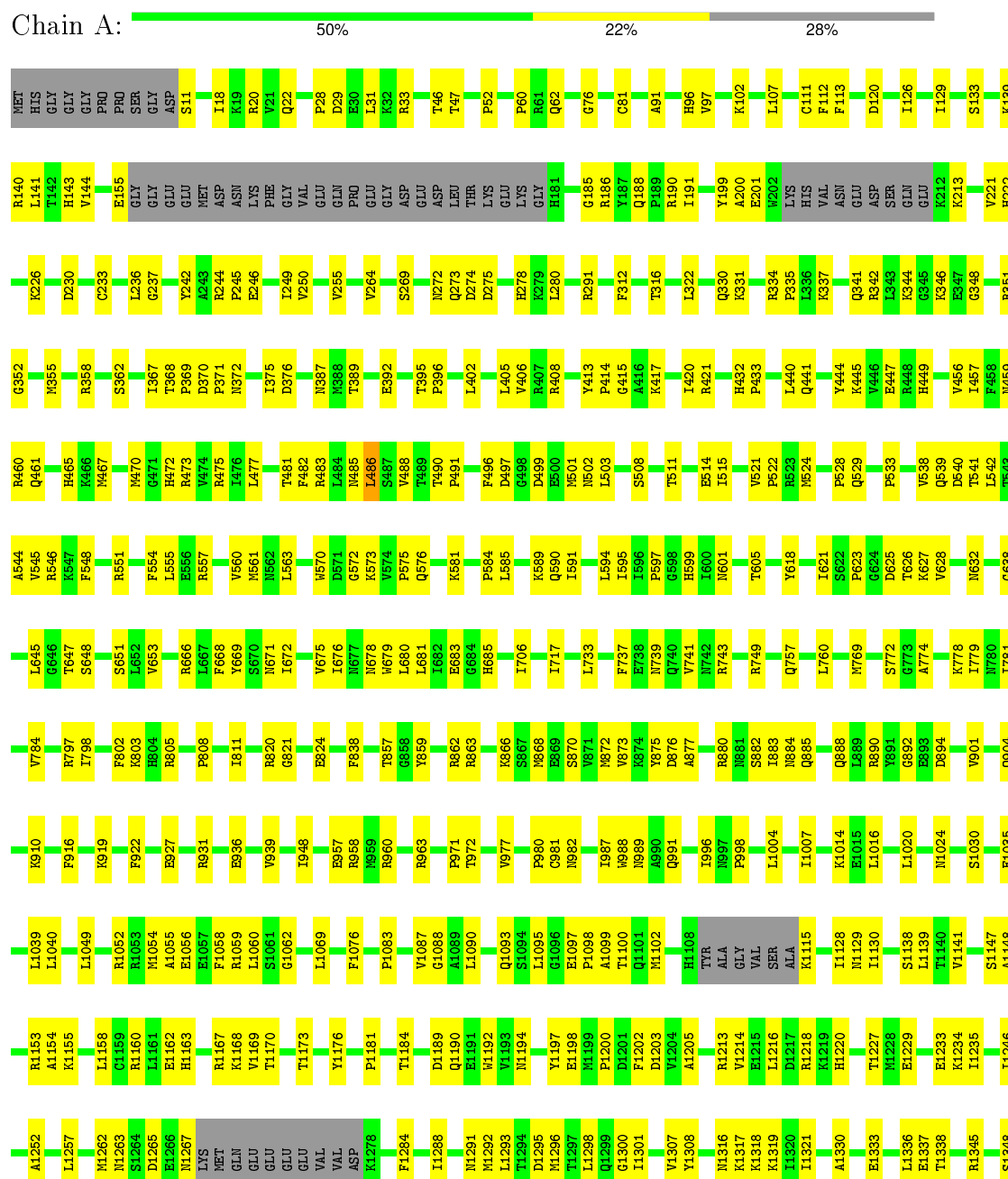
- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

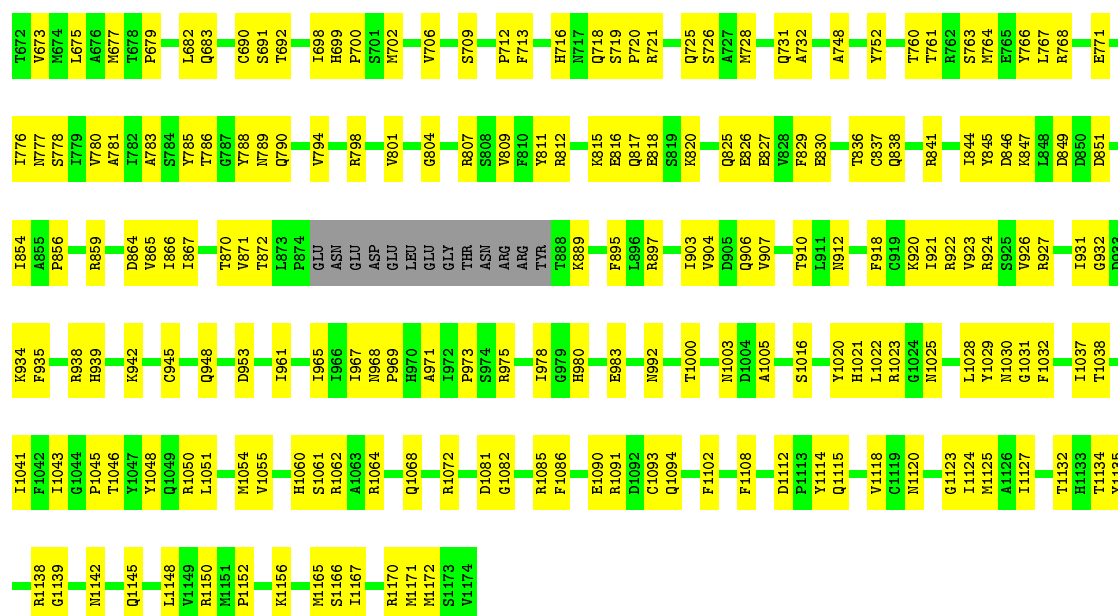
Mol	Chain	Residues	Atoms		AltConf
17	A	1	Total	Mg	0
			1	1	

3 Residue-property plots

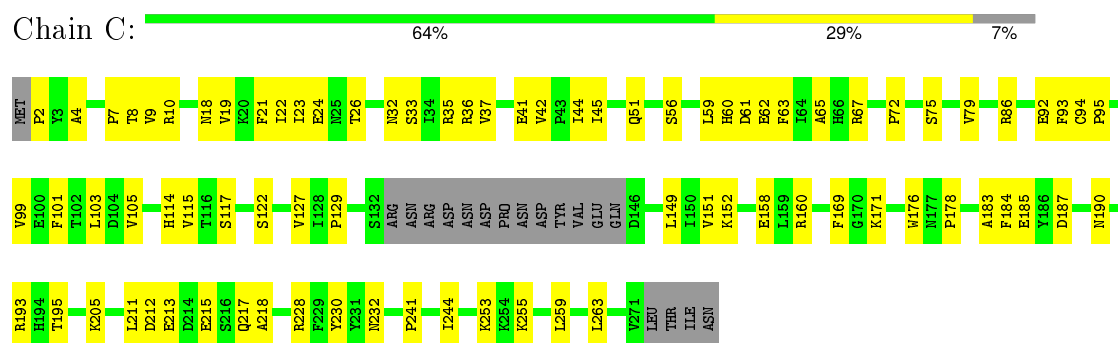
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric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

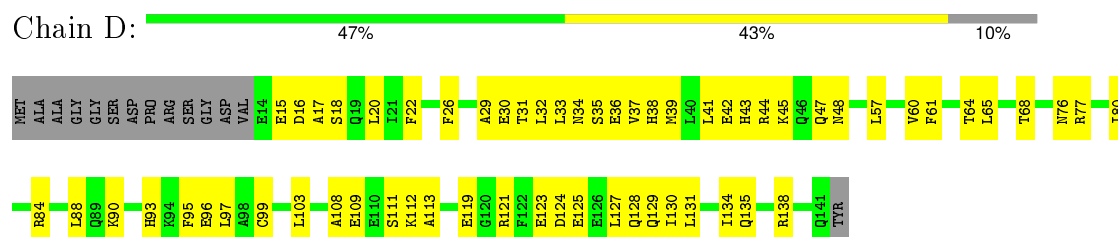




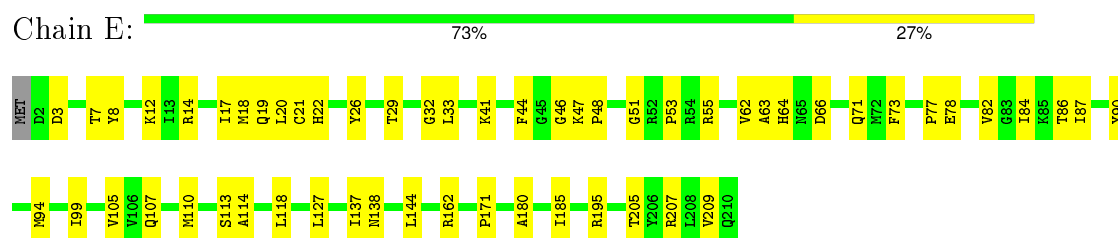
• Molecule 3: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3



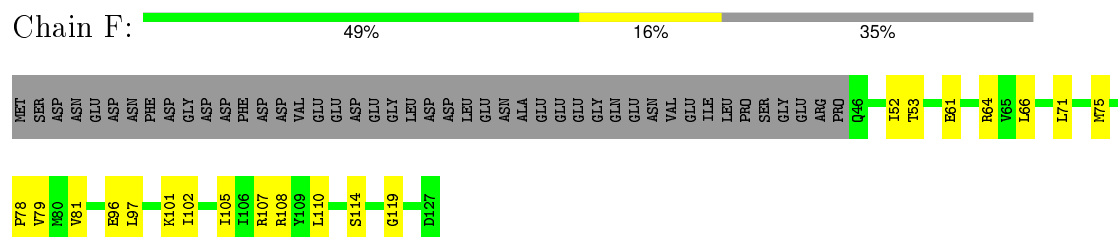
• Molecule 4: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4



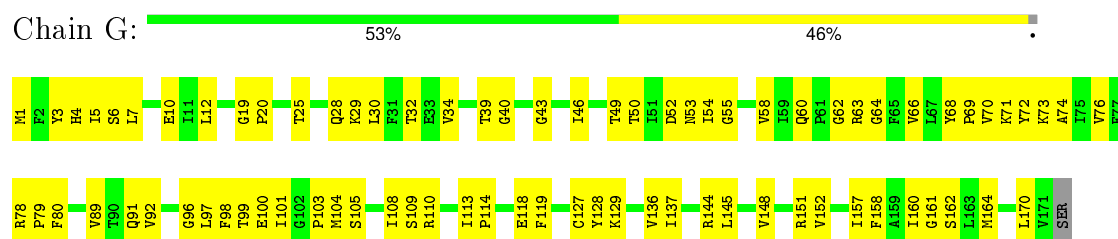
• Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC1



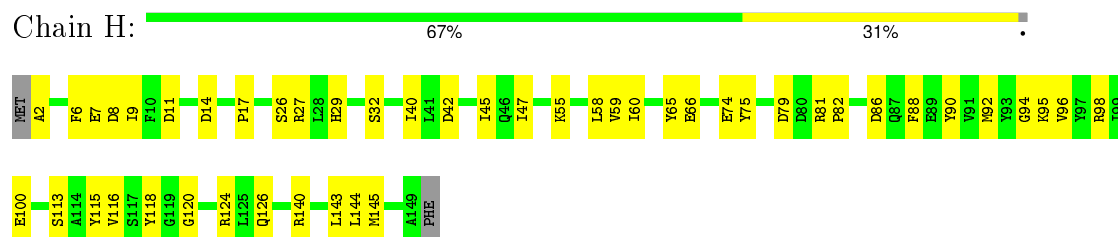
- Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC2



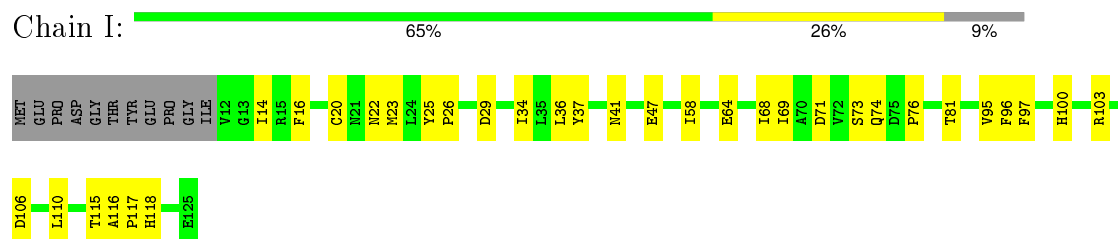
- Molecule 7: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7



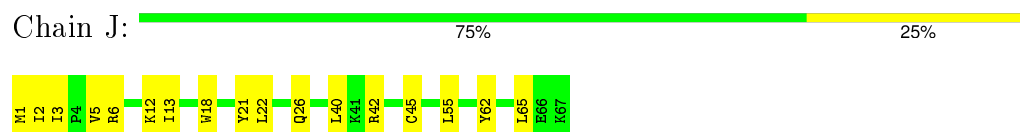
- Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC3



- Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9

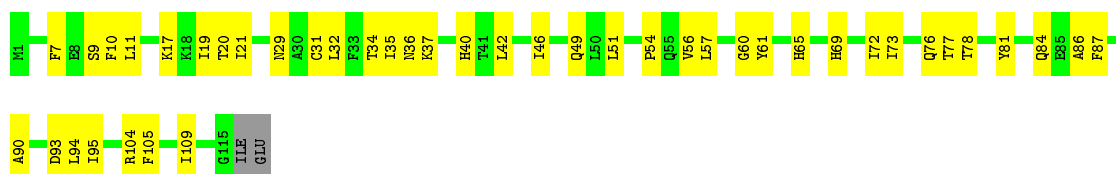


- Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC5



- Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11





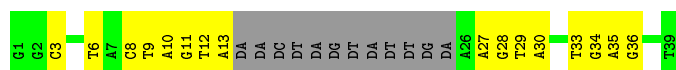
- Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC4

Chain L: 57% 19% 24%



- Molecule 13: DNA, DNA-RNA ELONGATION SCAFFOLD

Chain N: 28% 41% 31%



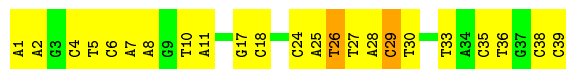
- Molecule 14: RNA, DNA-RNA ELONGATION SCAFFOLD

Chain P: 25% 25% 20% 30%



- Molecule 15: DNA, DNA-RNA ELONGATION SCAFFOLD

Chain T: 41% 54% 5%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	INDIVIDUAL PARTICLES, Not provided	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	43	Depositor
Minimum defocus (nm)	600	Depositor
Maximum defocus (nm)	3100	Depositor
Magnification	37000	Depositor
Image detector	GATAN K2 SUMMIT	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, 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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.25	0/11497	0.47	1/15521 (0.0%)
10	J	0.25	0/542	0.42	0/730
11	K	0.26	0/939	0.46	0/1271
12	L	0.27	0/377	0.52	0/500
13	N	0.78	4/619 (0.6%)	1.13	0/953
14	P	0.19	0/337	0.73	0/523
15	T	0.93	8/887 (0.9%)	1.13	2/1366 (0.1%)
2	B	0.26	0/9243	0.47	0/12475
3	C	0.26	0/2102	0.47	0/2857
4	D	0.28	0/1019	0.42	0/1374
5	E	0.26	0/1742	0.46	0/2353
6	F	0.53	1/668 (0.1%)	0.51	0/903
7	G	0.35	0/1365	0.50	0/1853
8	H	0.26	0/1207	0.50	0/1628
9	I	0.25	0/948	0.45	0/1284
All	All	0.32	13/33492 (0.0%)	0.53	3/45591 (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
1	A	0	1
11	K	0	1
2	B	0	1
All	All	0	3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	T	26	DT	O3'-P	-15.93	1.42	1.61
6	F	102	ILE	C-N	11.88	1.56	1.34
13	N	8	DC	C1'-N1	5.23	1.56	1.49
13	N	6	DT	C1'-N1	5.23	1.56	1.49
13	N	9	DT	C1'-N1	5.22	1.56	1.49

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
15	T	26	DT	P-O3'-C3'	11.37	133.35	119.70
15	T	26	DT	O3'-P-O5'	8.16	119.51	104.00
1	A	486	LEU	CA-CB-CG	-5.03	103.72	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	910	LYS	Peptide
2	B	497	LYS	Peptide
11	K	7	PHE	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	A	11291	0	11403	330	0
2	B	9062	0	9107	293	0
3	C	2059	0	2007	66	0
4	D	1005	0	964	53	0
5	E	1711	0	1733	38	0
6	F	658	0	686	16	0
7	G	1334	0	1333	65	0
8	H	1186	0	1147	40	0
9	I	927	0	859	30	0
10	J	533	0	554	15	0
11	K	920	0	942	32	0
12	L	372	0	378	13	0
13	N	552	0	309	16	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
14	P	301	0	153	9	0
15	T	792	0	442	15	0
16	A	2	0	0	0	0
16	B	1	0	0	0	0
16	C	1	0	0	0	0
16	I	2	0	0	0	0
16	J	1	0	0	0	0
16	L	1	0	0	0	0
17	A	1	0	0	0	0
All	All	32712	0	32017	915	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 915 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)
1:A:1481:LYS:HG2	7:G:20:PRO:HA	1.30	1.10
13:N:13:DA:H61	15:T:27:DT:H3	1.05	0.98
13:N:13:DA:N1	15:T:27:DT:O2	2.03	0.91
1:A:948:ILE:HG12	1:A:1007:ILE:HD11	1.52	0.89
1:A:392:GLU:O	1:A:445:LYS:HA	1.74	0.86

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	1417/1970 (72%)	1347 (95%)	70 (5%)	0	100	100
2	B	1128/1174 (96%)	1072 (95%)	56 (5%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	253/275 (92%)	238 (94%)	15 (6%)	0	100	100
4	D	126/142 (89%)	121 (96%)	5 (4%)	0	100	100
5	E	207/210 (99%)	198 (96%)	9 (4%)	0	100	100
6	F	80/127 (63%)	76 (95%)	4 (5%)	0	100	100
7	G	169/172 (98%)	164 (97%)	5 (3%)	0	100	100
8	H	146/150 (97%)	138 (94%)	8 (6%)	0	100	100
9	I	112/125 (90%)	103 (92%)	9 (8%)	0	100	100
10	J	65/67 (97%)	59 (91%)	6 (9%)	0	100	100
11	K	113/117 (97%)	110 (97%)	3 (3%)	0	100	100
12	L	42/58 (72%)	36 (86%)	6 (14%)	0	100	100
All	All	3858/4587 (84%)	3662 (95%)	196 (5%)	0	100	100

There are no Ramachandran outliers to report.

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	1252/1749 (72%)	1252 (100%)	0	100	100
2	B	993/1027 (97%)	993 (100%)	0	100	100
3	C	234/252 (93%)	234 (100%)	0	100	100
4	D	106/127 (84%)	106 (100%)	0	100	100
5	E	189/190 (100%)	189 (100%)	0	100	100
6	F	71/111 (64%)	71 (100%)	0	100	100
7	G	147/153 (96%)	147 (100%)	0	100	100
8	H	129/131 (98%)	129 (100%)	0	100	100
9	I	103/112 (92%)	103 (100%)	0	100	100
10	J	56/56 (100%)	56 (100%)	0	100	100
11	K	104/106 (98%)	104 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
12	L	41/55 (74%)	41 (100%)	0	100	100
All	All	3425/4069 (84%)	3425 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1462	GLN
2	B	370	HIS
9	I	118	HIS
2	B	245	GLN
2	B	525	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	13/20 (65%)	3 (23%)	1 (7%)

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
14	P	8	A
14	P	11	A
14	P	13	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
14	P	12	A

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 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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.