



wwPDB/EMDatabank EM Map/Model Validation Summary Report ⓘ

Dec 10, 2019 – 02:17 PM EST

PDB ID : 6C6T
EMDB ID: : EMD-7350
Title : CryoEM structure of E.coli RNA polymerase elongation complex bound with RfaH
Authors : Kang, J.Y.; Artsimovitch, I.; Landick, R.; Darst, S.A.
Deposited on : 2018-01-19
Resolution : 3.50 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

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

with specific help available everywhere you see the ⓘ symbol.

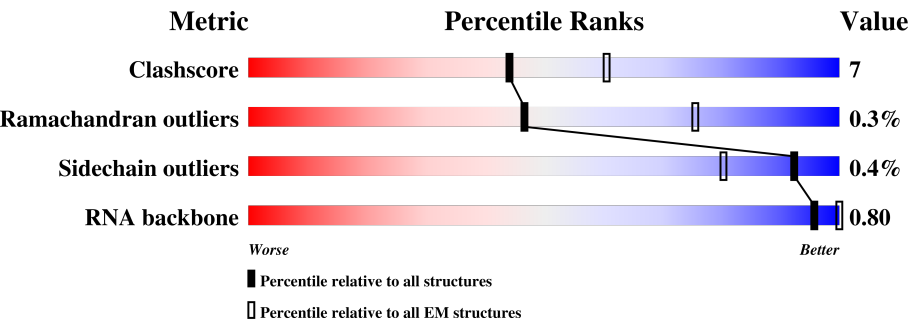
MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : 2.4

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.50 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.



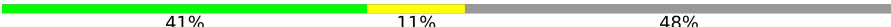
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531
RNA backbone	3747	458

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	29	<div><div>62%</div><div>34%</div><div>.</div></div>
2	B	29	<div><div>83%</div><div>17%</div></div>
3	R	20	<div><div>30%</div><div>20%</div><div>50%</div></div>
4	G	239	<div><div>79%</div><div>13%</div><div>8%</div></div>
4	H	239	<div><div>66%</div><div>26%</div><div>8%</div></div>
5	I	1342	<div><div>80%</div><div>18%</div><div>.</div></div>
6	J	1407	<div><div>74%</div><div>20%</div><div>5%</div></div>
7	K	91	<div><div>74%</div><div>18%</div><div>9%</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	D	162	 A horizontal bar chart showing the quality of chain D. The bar is divided into three segments: a green segment on the left labeled '41%', a yellow segment in the middle labeled '11%', and a grey segment on the right labeled '48%'. The segments are separated by thin white lines.

2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 26915 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 DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	29	Total	C	N	O	P	0	0
			598	282	117	171	28		

- Molecule 2 is a DNA chain called DNA (29-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	29	Total	C	N	O	P	0	0
			585	279	102	176	28		

- Molecule 3 is a RNA chain called RNA (5'-R(*GP*CP*AP*UP*UP*CP*AP*AP*AP*GP*CP*CP*GP*AP*GP*AP*GP*GP*UP*A)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
3	R	10	Total	C	N	O	P	0	0
			218	97	43	68	10		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	G	221	Total	C	N	O	S	0	0
			1694	1061	296	331	6		
4	H	219	Total	C	N	O	S	0	0
			1687	1053	298	330	6		

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	235	GLU	-	expression tag	UNP P0A7Z4
G	236	VAL	-	expression tag	UNP P0A7Z4
G	237	LEU	-	expression tag	UNP P0A7Z4
G	238	PHE	-	expression tag	UNP P0A7Z4
G	239	GLN	-	expression tag	UNP P0A7Z4
H	235	GLU	-	expression tag	UNP P0A7Z4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	236	VAL	-	expression tag	UNP P0A7Z4
H	237	LEU	-	expression tag	UNP P0A7Z4
H	238	PHE	-	expression tag	UNP P0A7Z4
H	239	GLN	-	expression tag	UNP P0A7Z4

- Molecule 5 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	1319	Total	C	N	O	S	0	0
			10407	6530	1814	2020	43		

- Molecule 6 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	1335	Total	C	N	O	S	0	0
			10388	6526	1854	1958	50		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
7	K	83	Total	C	N	O	S	0	0
			655	399	123	132	1		

- Molecule 8 is a protein called Transcription antitermination protein RfaH.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	D	85	Total	C	N	O	S	0	0
			680	438	119	120	3		

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
9	R	1	Total	Mg	0
			1	1	

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
10	J	2	Total	Zn	0
			2	2	

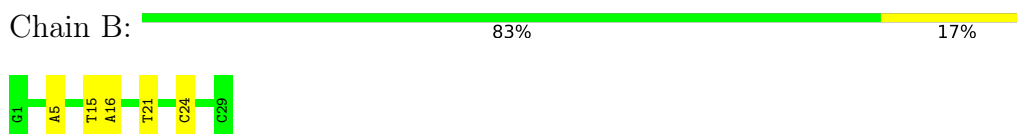
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

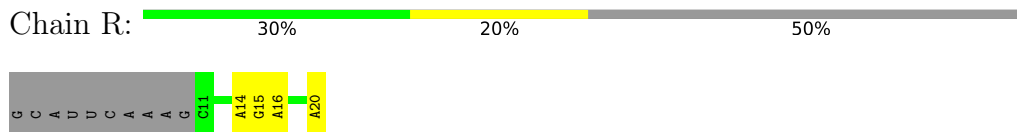
- Molecule 1: DNA (29-MER)



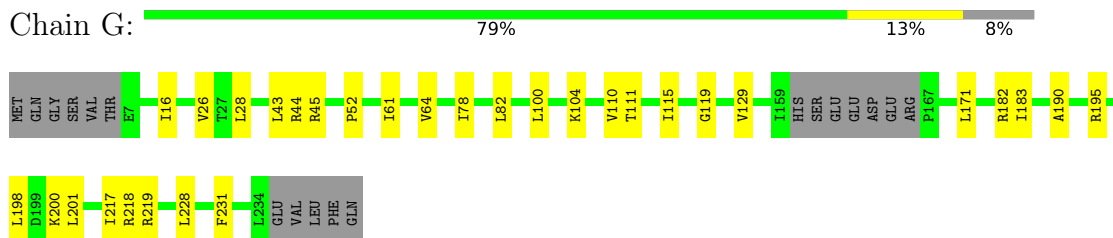
- Molecule 2: DNA (29-MER)



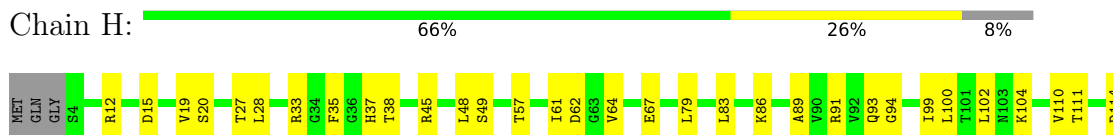
- Molecule 3: RNA (5'-R(*GP*CP*AP*UP*UP*CP*AP*AP*AP*GP*CP*CP*GP*AP*GP*AP*GP*GP*UP*A)-3')

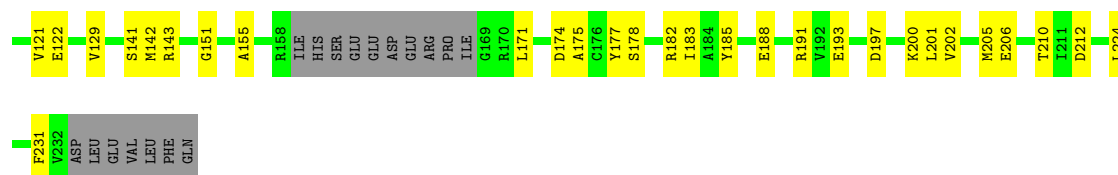


- Molecule 4: DNA-directed RNA polymerase subunit alpha



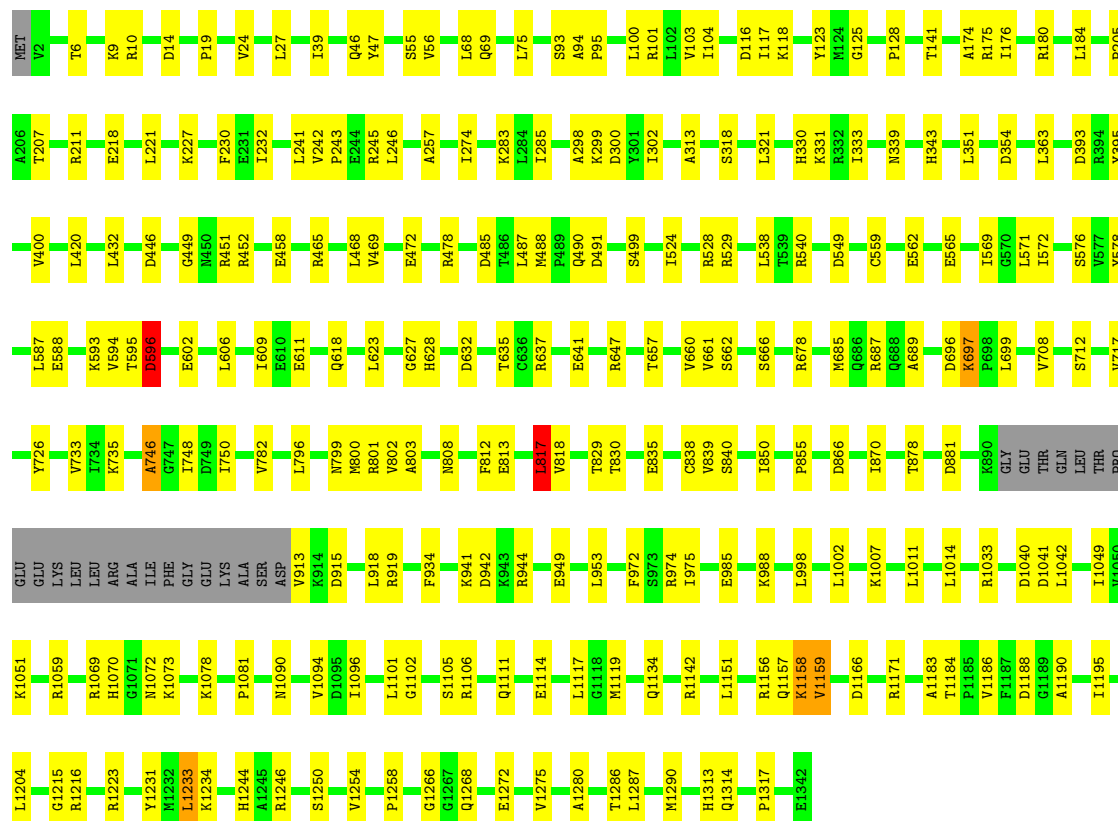
- Molecule 4: DNA-directed RNA polymerase subunit alpha





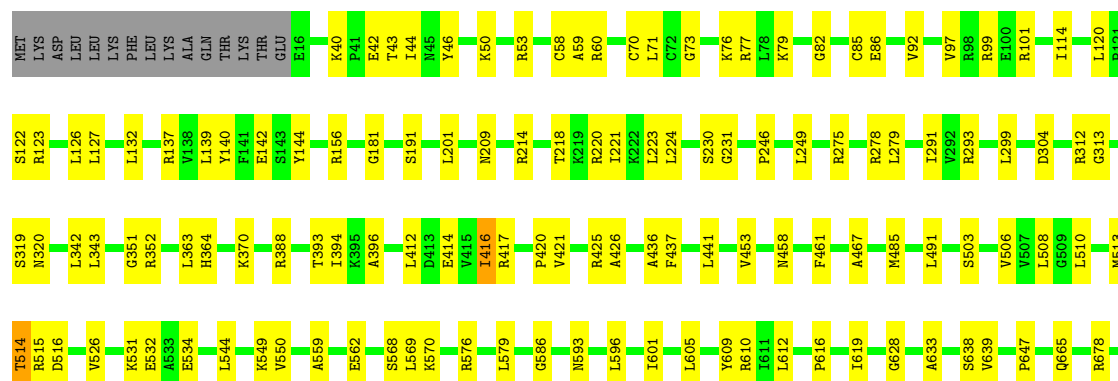
• Molecule 5: DNA-directed RNA polymerase subunit beta

Chain I: 80% 18%



• Molecule 6: DNA-directed RNA polymerase subunit beta'

Chain J: 74% 20% 5%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	174600	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 ($e^-/\text{\AA}^2$)	71.0	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	38462	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, 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	0.92	0/672	1.00	2/1037 (0.2%)
2	B	1.03	0/653	1.03	0/1005
3	R	0.90	0/244	0.98	0/379
4	G	0.48	0/1714	0.67	1/2323 (0.0%)
4	H	0.41	0/1706	0.68	0/2311
5	I	0.50	0/10573	0.64	2/14265 (0.0%)
6	J	0.48	0/10545	0.65	2/14236 (0.0%)
7	K	0.41	0/657	0.66	0/886
8	D	0.49	0/698	0.61	0/951
All	All	0.52	0/27462	0.68	7/37393 (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
5	I	0	9
6	J	0	5
7	K	0	1
All	All	0	15

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	J	903	LEU	C-N-CA	5.86	136.36	121.70
4	G	61	ILE	CG1-CB-CG2	-5.65	98.96	111.40
6	J	363	LEU	CA-CB-CG	5.59	128.16	115.30
5	I	1233	LEU	CA-CB-CG	5.46	127.85	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2	DG	O4'-C4'-C3'	-5.38	102.35	104.50

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	I	205	PRO	Peptide
5	I	594	VAL	Peptide
5	I	595	THR	Peptide
5	I	596	ASP	Mainchain
5	I	696	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	A	598	0	325	8	0
2	B	585	0	328	5	0
3	R	218	0	109	3	0
4	G	1694	0	1722	21	0
4	H	1687	0	1722	40	0
5	I	10407	0	10420	157	0
6	J	10388	0	10611	182	0
7	K	655	0	663	15	0
8	D	680	0	671	10	0
9	R	1	0	0	0	0
10	J	2	0	0	0	0
All	All	26915	0	26571	396	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 7.

The worst 5 of 396 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)
6:J:425:ARG:NH1	6:J:458:ASN:O	2.08	0.86

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:K:3:ARG:NH1	7:K:55:GLU:OE2	2.09	0.84
6:J:665:GLN:HE22	6:J:678:ARG:HH11	1.25	0.84
5:I:1314:GLN:HB2	7:K:28:ARG:HH12	1.43	0.83
6:J:40:LYS:NZ	6:J:42:GLU:OE1	2.15	0.80

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	
4	G	217/239 (91%)	197 (91%)	19 (9%)	1 (0%)	31	71
4	H	215/239 (90%)	202 (94%)	13 (6%)	0	100	100
5	I	1315/1342 (98%)	1204 (92%)	106 (8%)	5 (0%)	36	75
6	J	1329/1407 (94%)	1216 (92%)	111 (8%)	2 (0%)	49	83
7	K	81/91 (89%)	74 (91%)	7 (9%)	0	100	100
8	D	81/162 (50%)	74 (91%)	5 (6%)	2 (2%)	6	37
All	All	3238/3480 (93%)	2967 (92%)	261 (8%)	10 (0%)	47	78

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	I	596	ASP
5	I	1158	LYS
6	J	805	GLN
8	D	53	ASN
6	J	904	ALA

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	
4	G	186/206 (90%)	186 (100%)	0	100	100
4	H	187/206 (91%)	186 (100%)	1 (0%)	90	95
5	I	1138/1157 (98%)	1135 (100%)	3 (0%)	93	97
6	J	1120/1168 (96%)	1113 (99%)	7 (1%)	87	95
7	K	70/75 (93%)	70 (100%)	0	100	100
8	D	75/142 (53%)	75 (100%)	0	100	100
All	All	2776/2954 (94%)	2765 (100%)	11 (0%)	92	96

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	J	53	ARG
6	J	60	ARG
6	J	836	ARG
5	I	1159	VAL
6	J	709	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
6	J	450	HIS
6	J	777	HIS
8	D	20	HIS
6	J	665	GLN
6	J	805	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	9/20 (45%)	0	0

There are no RNA backbone outliers to report.

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 3 ligands modelled in this entry, 3 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.