wwPDB EM Validation Summary Report (i)

Nov 29, 2022 – 01:18 AM EST

PDB ID	:	7UBN
EMDB ID	:	EMD-26439
Title	:	Transcription antitermination complex: NusA-containing "engaged" Qlambda-loading complex
Authors	:	Yin, Z.; Ebright, R.H.
Deposited on	:	2022-03-15
Resolution	:	3.36 Å(reported)
Based on initial models	:	1LB2, 5LM9, 5LM7, 4YLN

This is a wwPDB EM 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/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

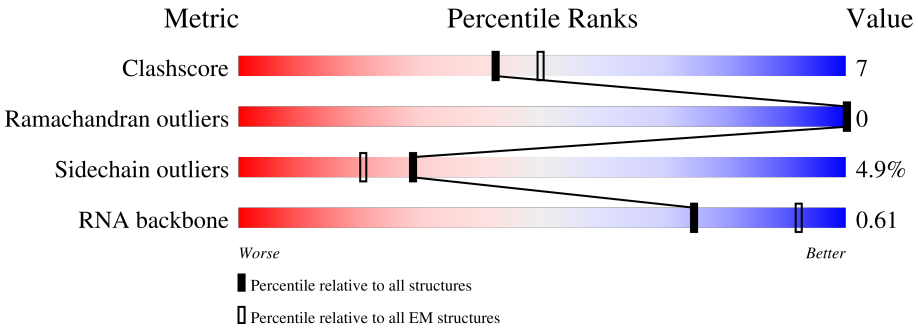
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 3.36 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.



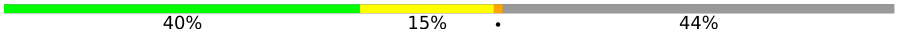
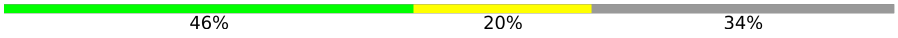
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826
RNA backbone	4643	859

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=5%

Mol	Chain	Length	Quality of chain
1	1	61	56% 31% 13%
2	2	61	64% 21% 15%
3	A	329	59% 11% 30%
3	B	329	66% 22% • 11%
4	C	1342	81% 18% •
5	D	1430	76% 16% 7%
6	E	91	69% 11% • 18%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	F	627	
8	N	515	
9	Q	207	
10	R	11	

2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 35069 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 (53-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
1	1	53	1096	525	195	323	53	0	0

- Molecule 2 is a DNA chain called DNA (52-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	2	52	1052	506	184	310	52	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	A	230	1786	1112	317	351	6	0	0
3	B	294	Total	C	N	O	S	0	0
			2274	1420	400	447	7		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	C	1339	10556	6620	1840	2053	43	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	D	1327	10319	6484	1839	1947	49	0	0

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1408	LEU	-	expression tag	UNP P0A8T7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	1409	GLU	-	expression tag	UNP P0A8T7
D	1410	ARG	-	expression tag	UNP P0A8T7
D	1411	ARG	-	expression tag	UNP P0A8T7
D	1412	ALA	-	expression tag	UNP P0A8T7
D	1413	SER	-	expression tag	UNP P0A8T7
D	1414	GLU	-	expression tag	UNP P0A8T7
D	1415	ASN	-	expression tag	UNP P0A8T7
D	1416	LEU	-	expression tag	UNP P0A8T7
D	1417	TYR	-	expression tag	UNP P0A8T7
D	1418	PHE	-	expression tag	UNP P0A8T7
D	1419	GLN	-	expression tag	UNP P0A8T7
D	1420	GLY	-	expression tag	UNP P0A8T7
D	1421	HIS	-	expression tag	UNP P0A8T7
D	1422	HIS	-	expression tag	UNP P0A8T7
D	1423	HIS	-	expression tag	UNP P0A8T7
D	1424	HIS	-	expression tag	UNP P0A8T7
D	1425	HIS	-	expression tag	UNP P0A8T7
D	1426	HIS	-	expression tag	UNP P0A8T7
D	1427	HIS	-	expression tag	UNP P0A8T7
D	1428	HIS	-	expression tag	UNP P0A8T7
D	1429	HIS	-	expression tag	UNP P0A8T7
D	1430	HIS	-	expression tag	UNP P0A8T7

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

Mol	Chain	Residues	Atoms					AltConf	Trace
6	E	75	Total	C	N	O	S	0	0
			600	365	114	120	1		

- Molecule 7 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	F	353	Total	C	N	O	S	0	0
			2900	1823	521	539	17		

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-13	MET	-	expression tag	UNP Q0P6L9
F	-12	GLY	-	expression tag	UNP Q0P6L9
F	-11	SER	-	expression tag	UNP Q0P6L9
F	-10	SER	-	expression tag	UNP Q0P6L9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	-9	HIS	-	expression tag	UNP Q0P6L9
F	-8	HIS	-	expression tag	UNP Q0P6L9
F	-7	HIS	-	expression tag	UNP Q0P6L9
F	-6	HIS	-	expression tag	UNP Q0P6L9
F	-5	HIS	-	expression tag	UNP Q0P6L9
F	-4	HIS	-	expression tag	UNP Q0P6L9
F	-3	SER	-	expression tag	UNP Q0P6L9
F	-2	SER	-	expression tag	UNP Q0P6L9
F	-1	GLY	-	expression tag	UNP Q0P6L9
F	0	HIS	-	expression tag	UNP Q0P6L9

- Molecule 8 is a protein called Transcription termination/antitermination protein NusA.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	N	340	Total	C	N	O	S	1	0
			2674	1662	481	522	9		

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	-19	MET	-	expression tag	UNP C3SSN7
N	-18	GLY	-	expression tag	UNP C3SSN7
N	-17	SER	-	expression tag	UNP C3SSN7
N	-16	SER	-	expression tag	UNP C3SSN7
N	-15	HIS	-	expression tag	UNP C3SSN7
N	-14	HIS	-	expression tag	UNP C3SSN7
N	-13	HIS	-	expression tag	UNP C3SSN7
N	-12	HIS	-	expression tag	UNP C3SSN7
N	-11	HIS	-	expression tag	UNP C3SSN7
N	-10	HIS	-	expression tag	UNP C3SSN7
N	-9	SER	-	expression tag	UNP C3SSN7
N	-8	SER	-	expression tag	UNP C3SSN7
N	-7	GLY	-	expression tag	UNP C3SSN7
N	-6	LEU	-	expression tag	UNP C3SSN7
N	-5	VAL	-	expression tag	UNP C3SSN7
N	-4	PRO	-	expression tag	UNP C3SSN7
N	-3	ARG	-	expression tag	UNP C3SSN7
N	-2	GLY	-	expression tag	UNP C3SSN7
N	-1	SER	-	expression tag	UNP C3SSN7
N	0	HIS	-	expression tag	UNP C3SSN7

- Molecule 9 is a protein called Antitermination protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	Q	207	Total	C	N	O	S	0	0
			1567	972	287	291	17		

- Molecule 10 is a RNA chain called RNA (5'-R(*UP*GP*GP*GP*AP*GP*AP*GP*GP*UP *A)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
10	R	11	Total	C	N	O	P	0	0
			241	108	49	74	10		

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
11	D	2	Total	Zn	0
			2	2	
11	Q	1	Total	Zn	0
			1	1	

- Molecule 12 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

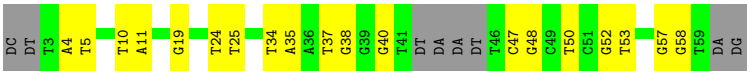
Mol	Chain	Residues	Atoms		AltConf
12	D	1	Total	Mg	0
			1	1	

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

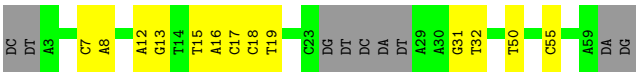
• Molecule 1: DNA (53-MER)

Chain 1:



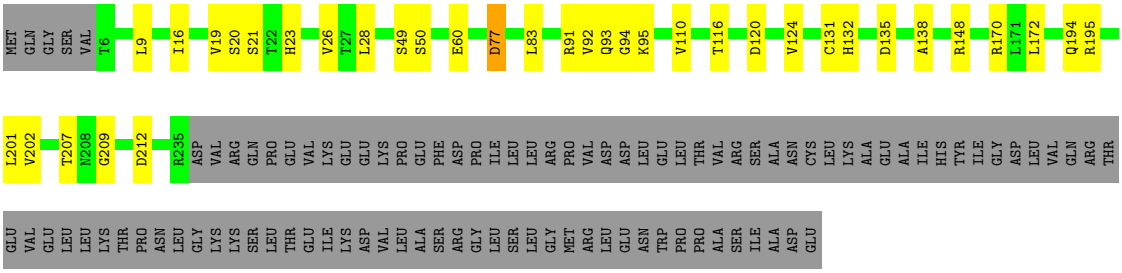
• Molecule 2: DNA (52-MER)

Chain 2:



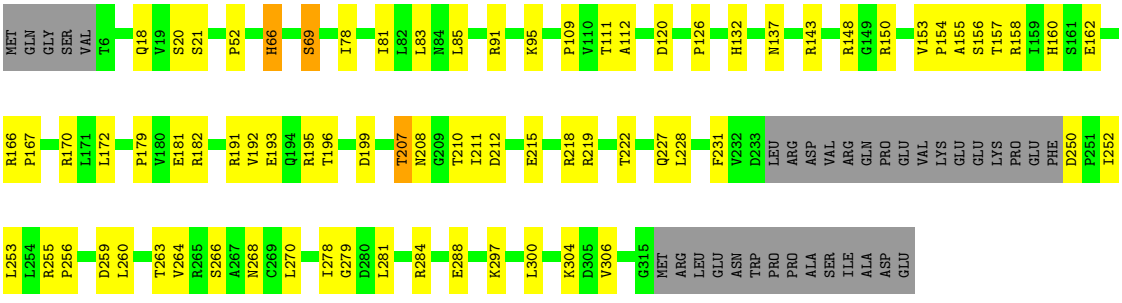
• Molecule 3: DNA-directed RNA polymerase subunit alpha

Chain A:



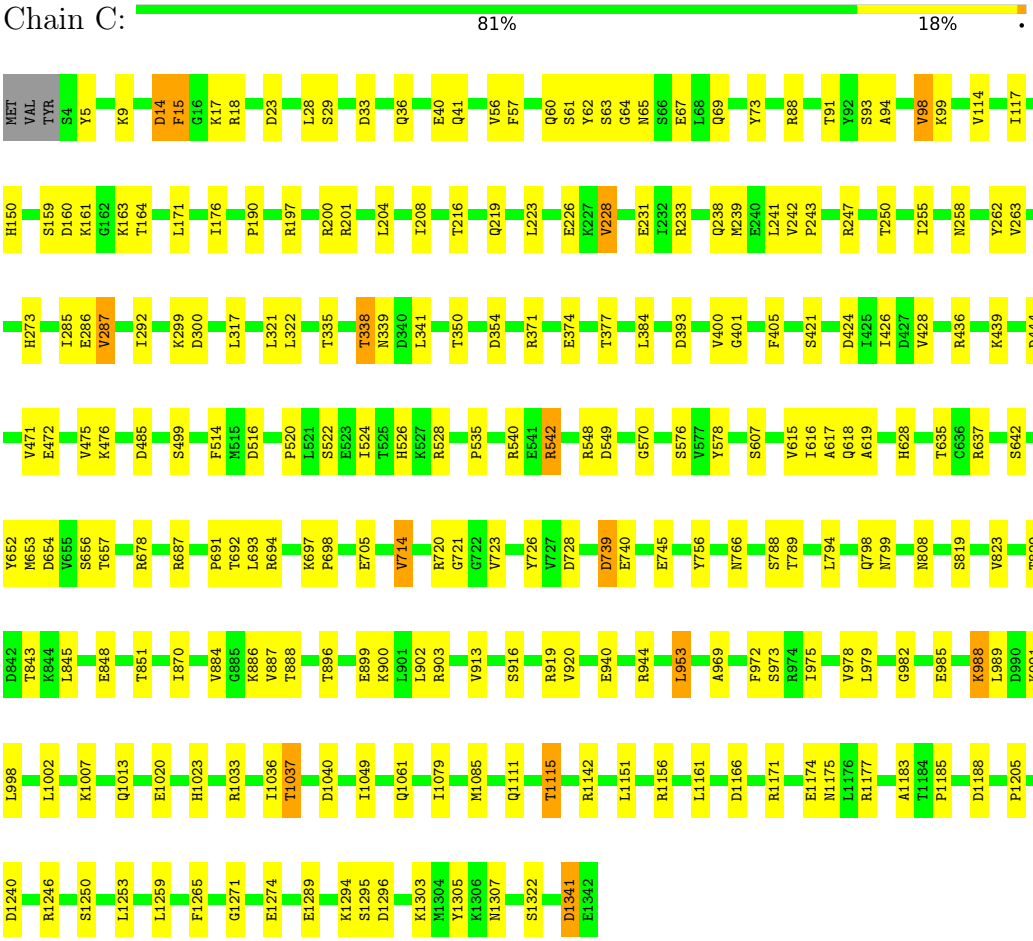
• Molecule 3: DNA-directed RNA polymerase subunit alpha

Chain B:



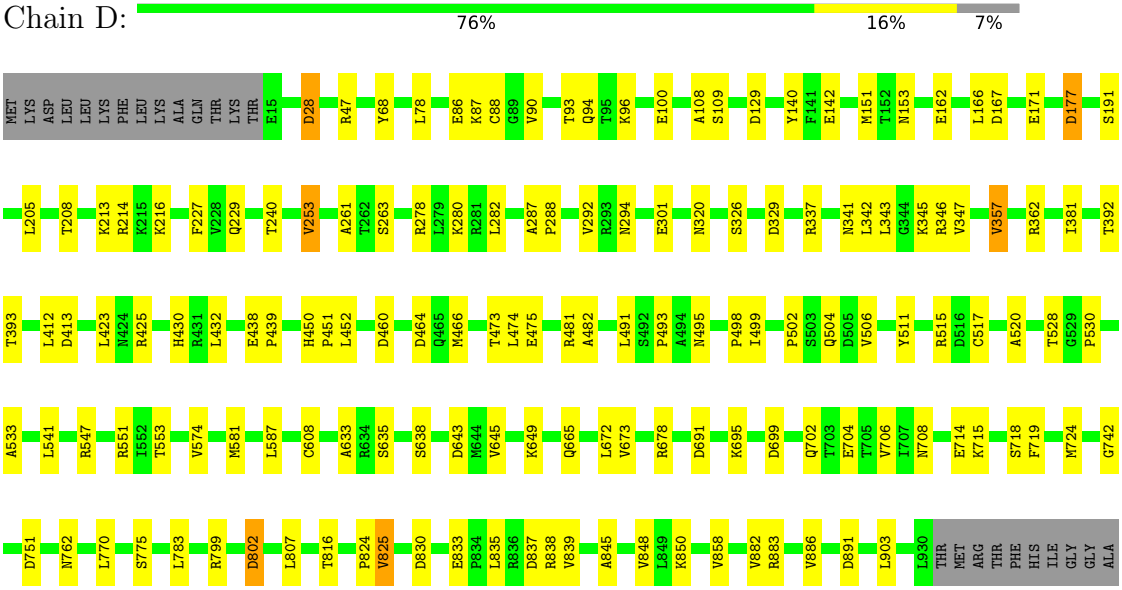
• Molecule 4: DNA-directed RNA polymerase subunit beta

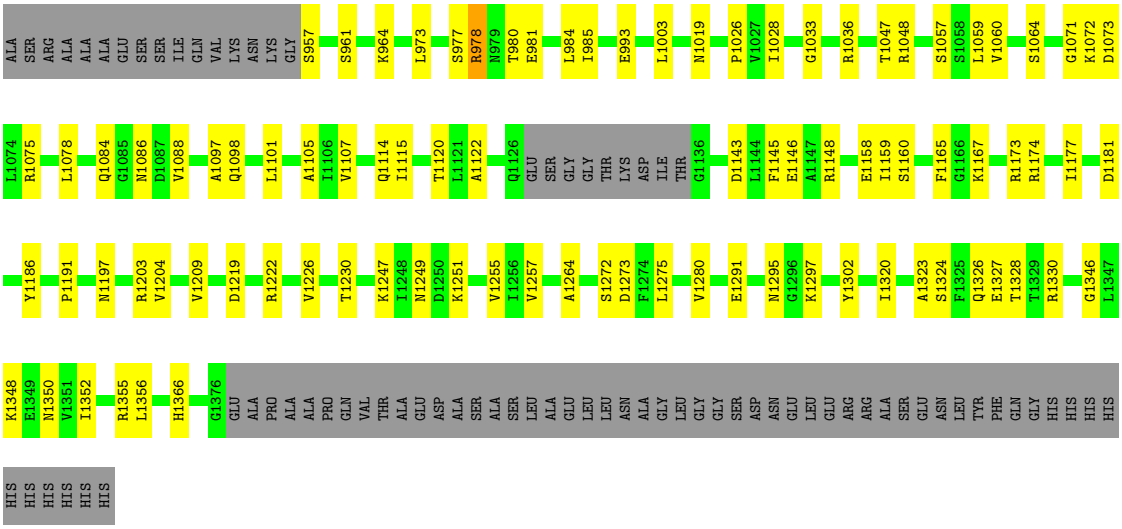
Chain C:



• Molecule 5: DNA-directed RNA polymerase subunit beta'

Chain D:



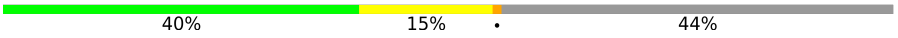


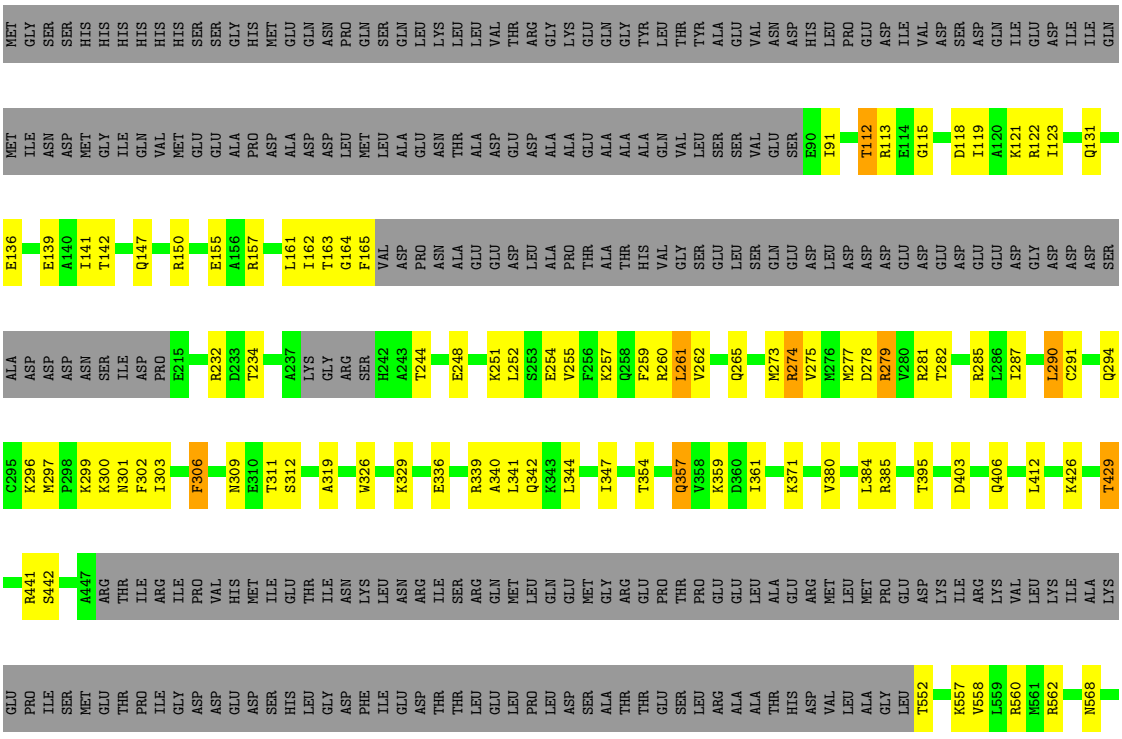
• Molecule 6: DNA-directed RNA polymerase subunit omega

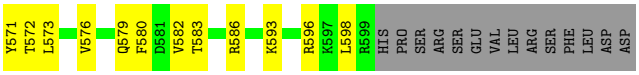
Chain E:

• Molecule 7: RNA polymerase sigma factor RpoD

Chain F:

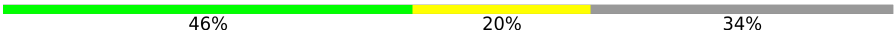


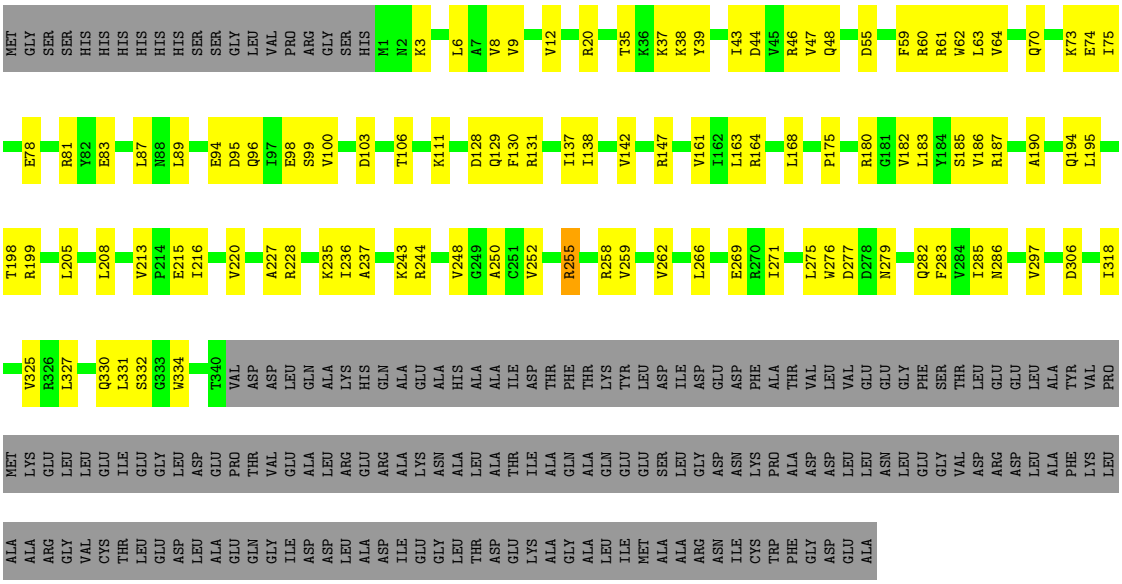




• Molecule 8: Transcription termination/antitermination protein NusA

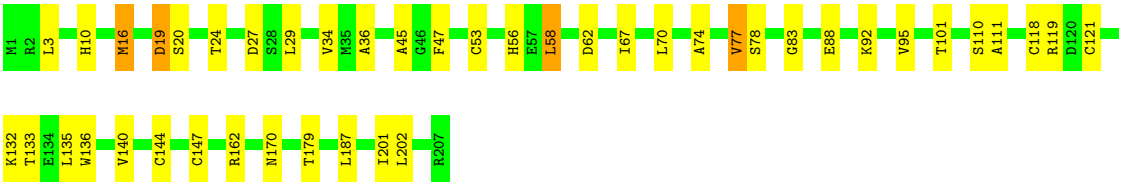
Chain N:





• Molecule 9: Antitermination protein

Chain Q:



• Molecule 10: RNA (5'-R(*UP*GP*GP*GP*AP*GP*AP*GP*GP*UP*A)-3')

Chain R:

4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	12391	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TALOS ARCTICA	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	1.1	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
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: MG, 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	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	1	0.64	0/1228	1.06	0/1894
2	2	0.74	0/1176	1.02	0/1806
3	A	0.32	0/1808	0.48	0/2450
3	B	0.30	0/2300	0.57	0/3116
4	C	0.37	0/10724	0.52	0/14468
5	D	0.36	0/10476	0.53	0/14146
6	E	0.29	0/602	0.49	0/810
7	F	0.29	0/2936	0.54	0/3935
8	N	0.31	0/2708	0.62	0/3659
9	Q	0.32	0/1595	0.55	0/2147
10	R	0.81	0/271	1.07	0/423
All	All	0.38	0/35824	0.60	0/48854

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1	1096	0	605	15	0
2	2	1052	0	590	9	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	1786	0	1813	18	0
3	B	2274	0	2329	48	0
4	C	10556	0	10566	141	0
5	D	10319	0	10530	124	0
6	E	600	0	607	10	0
7	F	2900	0	2956	65	0
8	N	2674	0	2705	64	0
9	Q	1567	0	1572	38	0
10	R	241	0	120	3	0
11	D	2	0	0	0	0
11	Q	1	0	0	0	0
12	D	1	0	0	0	0
All	All	35069	0	34393	496	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 496 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)
9:Q:53:CYS:HA	9:Q:58:LEU:HD11	1.44	0.98
7:F:290:LEU:HD21	7:F:336:GLU:OE1	1.67	0.95
5:D:78:LEU:HD21	9:Q:56:HIS:NE2	1.87	0.90
9:Q:58:LEU:HD12	9:Q:58:LEU:O	1.82	0.79
9:Q:53:CYS:HA	9:Q:58:LEU:CD1	2.15	0.77

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	
3	A	228/329 (69%)	223 (98%)	5 (2%)	0	100	100
3	B	290/329 (88%)	273 (94%)	17 (6%)	0	100	100
4	C	1337/1342 (100%)	1269 (95%)	68 (5%)	0	100	100
5	D	1321/1430 (92%)	1253 (95%)	68 (5%)	0	100	100
6	E	73/91 (80%)	71 (97%)	2 (3%)	0	100	100
7	F	345/627 (55%)	325 (94%)	20 (6%)	0	100	100
8	N	339/515 (66%)	318 (94%)	21 (6%)	0	100	100
9	Q	205/207 (99%)	191 (93%)	14 (7%)	0	100	100
All	All	4138/4870 (85%)	3923 (95%)	215 (5%)	0	100	100

There are no Ramachandran outliers to report.

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 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	
3	A	198/286 (69%)	191 (96%)	7 (4%)	36	66
3	B	254/286 (89%)	242 (95%)	12 (5%)	26	58
4	C	1153/1157 (100%)	1109 (96%)	44 (4%)	33	63
5	D	1110/1189 (93%)	1049 (94%)	61 (6%)	21	53
6	E	65/75 (87%)	61 (94%)	4 (6%)	18	49
7	F	310/552 (56%)	284 (92%)	26 (8%)	11	37
8	N	288/426 (68%)	280 (97%)	8 (3%)	43	71
9	Q	164/164 (100%)	154 (94%)	10 (6%)	18	50
All	All	3542/4135 (86%)	3370 (95%)	172 (5%)	29	57

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	D	1145	PHE
7	F	357	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	D	1226	VAL
7	F	147	GLN
7	F	552	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
5	D	1259	GLN
7	F	409	ASN
7	F	265	GLN
7	F	342	GLN
8	N	112	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
10	R	10/11 (90%)	1 (10%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
10	R	2	G

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 4 ligands modelled in this entry, 4 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 Map visualisation

This section contains visualisations of the EMDB entry EMD-26439. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal surface views

This section was not generated.

6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.