wwPDB EM Validation Summary Report (i)

May 4, 2024 – 06:41 pm BST

PDB ID	:	6GML
EMDB ID	:	EMD-0038
Title	:	Structure of paused transcription complex Pol II-DSIF-NELF
Authors	:	Vos, S.M.; Farnung, L.; Urlaub, H.; Cramer, P.
Deposited on	:	2018-05-27
Resolution	:	3.20 Å(reported)

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.dev92
MolProbity	:	FAILED
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)	:	2.36.2

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 3.20 Å.

There are no overall percentile quality scores available for this entry.

2 Entry composition (i)

There are 23 unique types of molecules in this entry. The entry contains 44827 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 subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	1407	11142	7014	1997	2063	68	0	0

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

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1116	8928	5652	1568	1644	64	0	0

- Molecule 3 is a protein called RNA polymerase II subunit C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	258	2072	1301	353	412	6	0	0

- Molecule 4 is a protein called RNA polymerase II subunit E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	E	209	1721	1089	300	324	8	0	0

- Molecule 5 is a protein called RNA polymerase II subunit F.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	F	82	658	418	113	122	5	0	0

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

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

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
7	I	117	950	587	169	183	11	0	0

- Molecule 8 is a protein called Uncharacterized protein.

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

- Molecule 9 is a protein called Uncharacterized protein.

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

- Molecule 10 is a protein called Uncharacterized protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
10	L	46	389	241	75	67	6	0	0

- Molecule 11 is a DNA chain called Non-template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
11	N	36	730	349	131	214	36	0	0

- Molecule 12 is a RNA chain called TAR RNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
12	P	17	361	162	66	116	17	0	0

- Molecule 13 is a DNA chain called Template DNA.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
13	T	45	933	443	172	273	45	0	0

- Molecule 14 is a protein called Negative elongation factor A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
14	U	183	1410	895	239	269	7	0	0

- Molecule 15 is a protein called Negative elongation factor B,Negative elongation factor B.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
15	V	483	1932	966	483	483	0	0

- Molecule 16 is a protein called Negative elongation factor C/D.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
16	W	538	3858	2440	670	728	20	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	7	SER	-	expression tag	UNP Q8IXH7
W	8	ASN	-	expression tag	UNP Q8IXH7
W	9	ALA	-	expression tag	UNP Q8IXH7
W	51	UNK	PHE	conflict	UNP Q8IXH7
W	52	UNK	SER	conflict	UNP Q8IXH7
W	53	UNK	THR	conflict	UNP Q8IXH7
W	54	UNK	ARG	conflict	UNP Q8IXH7
W	55	UNK	ASP	conflict	UNP Q8IXH7
W	56	UNK	TYR	conflict	UNP Q8IXH7
W	57	UNK	ILE	conflict	UNP Q8IXH7
W	58	UNK	MET	conflict	UNP Q8IXH7
W	59	UNK	GLU	conflict	UNP Q8IXH7
W	60	UNK	PRO	conflict	UNP Q8IXH7
W	61	UNK	SER	conflict	UNP Q8IXH7
W	62	UNK	ILE	conflict	UNP Q8IXH7
W	63	UNK	PHE	conflict	UNP Q8IXH7
W	64	UNK	ASN	conflict	UNP Q8IXH7
W	65	UNK	THR	conflict	UNP Q8IXH7
W	66	UNK	LEU	conflict	UNP Q8IXH7
W	67	UNK	LYS	conflict	UNP Q8IXH7
W	68	UNK	ARG	conflict	UNP Q8IXH7
W	69	UNK	TYR	conflict	UNP Q8IXH7
W	70	UNK	PHE	conflict	UNP Q8IXH7
W	71	UNK	GLN	conflict	UNP Q8IXH7
W	72	UNK	ALA	conflict	UNP Q8IXH7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
W	73	UNK	GLY	conflict	UNP Q8IXH7
W	74	UNK	GLY	conflict	UNP Q8IXH7
W	75	UNK	SER	conflict	UNP Q8IXH7
W	76	UNK	PRO	conflict	UNP Q8IXH7
W	77	UNK	GLU	conflict	UNP Q8IXH7
W	78	UNK	ASN	conflict	UNP Q8IXH7
W	79	UNK	VAL	conflict	UNP Q8IXH7
W	80	UNK	ILE	conflict	UNP Q8IXH7
W	81	UNK	GLN	conflict	UNP Q8IXH7
W	82	UNK	LEU	conflict	UNP Q8IXH7
W	83	UNK	LEU	conflict	UNP Q8IXH7
W	84	UNK	SER	conflict	UNP Q8IXH7
W	85	UNK	GLU	conflict	UNP Q8IXH7
W	86	UNK	ASN	conflict	UNP Q8IXH7
W	87	UNK	TYR	conflict	UNP Q8IXH7
W	88	UNK	THR	conflict	UNP Q8IXH7
W	89	UNK	ALA	conflict	UNP Q8IXH7
W	90	UNK	VAL	conflict	UNP Q8IXH7
W	91	UNK	ALA	conflict	UNP Q8IXH7

- Molecule 17 is a protein called Negative elongation factor E,Negative elongation factor E.

Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
17	X	22	110	66	22	22	0	0

- Molecule 18 is a protein called Transcription elongation factor SPT4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
18	Y	116	911	570	159	173	9	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Y	-3	GLY	-	expression tag	UNP P63272
Y	-2	PRO	-	expression tag	UNP P63272
Y	-1	GLY	-	expression tag	UNP P63272
Y	0	SER	-	expression tag	UNP P63272

- Molecule 19 is a protein called Transcription elongation factor SPT5.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
19	Z	470	3770	2400	663	690	17	0	0

- Molecule 20 is a protein called RNA polymerase II subunit D.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
20	D	129	998	629	170	195	4	0	1

- Molecule 21 is a protein called RNA polymerase II subunit G.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
21	G	171	1305	852	205	240	8	0	0

- Molecule 22 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 23 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

SEQUENCE-PLOTS INFOmissingINFO

3 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	162269	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 ($e^-/Å^2$)	46	Depositor
Minimum defocus (nm)	Not provided	
Maximum defocus (nm)	Not provided	
Magnification	165000	Depositor
Image detector	GATAN K2 QUANTUM (4k x 4k)	Depositor

4 Model quality (i)

4.1 Standard geometry (i)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts (i)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA (i)

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 10 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.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
16	W	2
15	V	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	V	419:ARG	C	420:MET	N	7.63
1	W	185:GLN	C	186:GLY	N	4.08
1	W	588:MET	C	589:VAL	N	3.94

5 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-0038. 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.

5.1 Orthogonal projections (i)

This section was not generated.

5.2 Central slices (i)

This section was not generated.

5.3 Largest variance slices (i)

This section was not generated.

5.4 Orthogonal standard-deviation projections (False-color) (i)

This section was not generated.

5.5 Orthogonal surface views (i)

This section was not generated.

5.6 Mask visualisation (i)

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

6 Map analysis (i)

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

6.1 Map-value distribution (i)

This section was not generated.

6.2 Volume estimate versus contour level (i)

This section was not generated.

6.3 Rotationally averaged power spectrum (i)

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

7 Fourier-Shell correlation (i)

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

8 Map-model fit (i)

This section was not generated.