



wwPDB EM Validation Summary Report ⓘ

May 10, 2026 – 12:19 AM JST

PDB ID : 9K18 / pdb_00009k18
EMDB ID : EMD-61968
Title : A cryo-EM structure of *B. oleracea* RNA polymerase V in complex with 7U scaffold at 3.42 Angstrom
Authors : Xie, G.; Du, X.; Du, J.
Deposited on : 2024-10-16
Resolution : 3.42 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

EMDB validation analysis : **FAILED**
MolProbity : 4-5-2 with Phenix2.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
EM percentile statistics : **NOT EXECUTED**
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

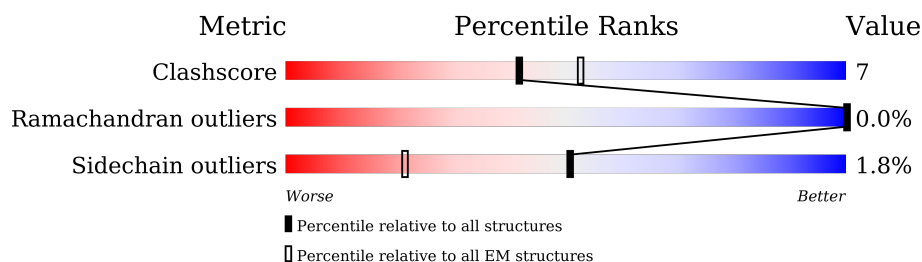
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.42 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.



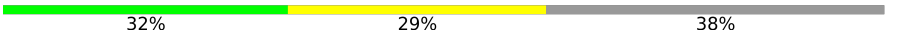
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	229148	23984
Ramachandran outliers	224038	23583
Sidechain outliers	223484	23102

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	2032	
2	B	1169	
3	C	319	
4	E	230	
5	F	144	
6	H	146	
7	I	114	
8	J	71	
9	K	116	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	L	51	 71% 20% 10%
11	T	34	 12% 26% 62%
12	N	34	 32% 29% 38%

2 Entry composition

There are 14 unique types of molecules in this entry. The entry contains 23692 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 V largest subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	854	Total	C	N	O	S	0	0
			6672	4210	1149	1268	45		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1006	Total	C	N	O	S	0	0
			7995	5030	1415	1506	44		

- Molecule 3 is a protein called DNA-directed RNA polymerase RpoA/D/Rpb3-type domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	300	Total	C	N	O	S	0	0
			2340	1463	393	471	13		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	THR	SER	conflict	UNP A0A0D3D418

- Molecule 4 is a protein called RNA polymerase subunit H/Rpb5 C-terminal domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	E	208	Total	C	N	O	S	0	0
			1698	1079	302	315	2		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	101	GLY	SER	conflict	UNP A0A0D3DTU3
E	182	GLN	HIS	conflict	UNP A0A0D3DTU3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	210	ILE	VAL	conflict	UNP A0A0D3DTU3

- Molecule 5 is a protein called DNA-directed RNA polymerase RpoA/D/Rpb3-type domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	F	80	Total	C	N	O	S	0	0
			660	420	114	122	4		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	51	GLU	ASP	conflict	UNP A0A0D3BZZ8

- Molecule 6 is a protein called DNA-directed RNA polymerase II, IV and V subunit 8.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	H	142	Total	C	N	O	S	0	0
			1135	732	184	210	9		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
7	I	91	Total	C	N	O	S	0	0
			713	436	130	136	11		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	20	ARG	LYS	conflict	UNP A0A0D3A7P5
I	40	ASP	ASN	conflict	UNP A0A0D3A7P5

- Molecule 8 is a protein called DNA-directed RNA polymerase II, IV and V subunit 10.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	J	64	Total	C	N	O	S	0	0
			514	328	86	93	7		

- Molecule 9 is a protein called DNA-directed RNA polymerase RBP11-like dimerisation domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	K	108	Total	C	N	O	S	0	0
			887	563	154	168	2		

- Molecule 10 is a protein called DNA-directed RNA polymerase II, IV and V subunit 12.

Mol	Chain	Residues	Atoms					AltConf	Trace
10	L	46	Total	C	N	O	S	0	0
			375	229	71	71	4		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	18	GLU	LYS	conflict	UNP A0A0D2ZPP3
L	32	CYS	ARG	conflict	UNP A0A0D2ZPP3

- Molecule 11 is a DNA chain called DNA (34-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
11	T	13	Total	C	N	O	P	0	0
			264	127	47	77	13		

- Molecule 12 is a DNA chain called DNA (34-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
12	N	21	Total	C	N	O	P	0	0
			431	205	80	125	21		

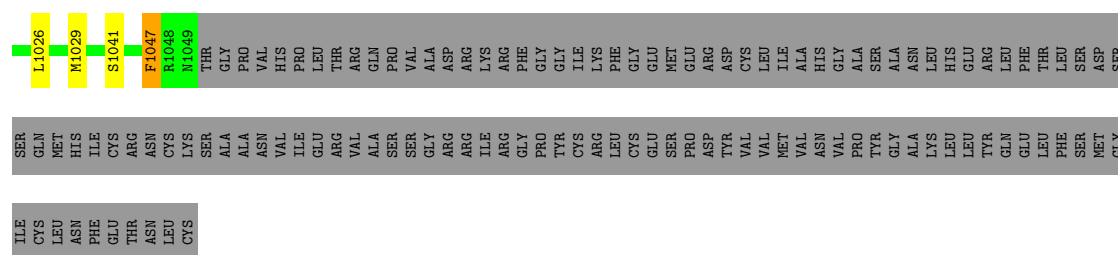
- Molecule 13 is ZINC ION (CCD ID: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
13	A	1	Total	Zn	0
			1	1	
13	C	1	Total	Zn	0
			1	1	
13	H	1	Total	Zn	0
			1	1	
13	I	2	Total	Zn	0
			2	2	
13	J	1	Total	Zn	0
			1	1	
13	L	1	Total	Zn	0
			1	1	

- Molecule 14 is CALCIUM ION (CCD ID: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

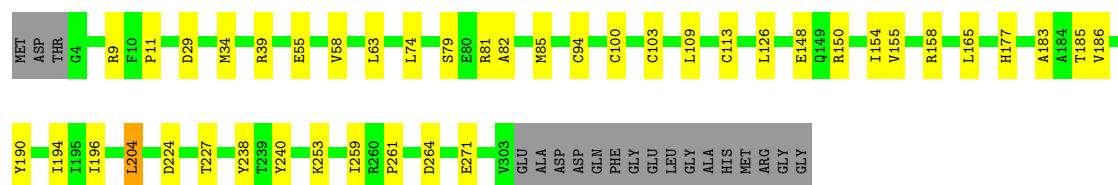
Mol	Chain	Residues	Atoms		AltConf
14	A	1	Total 1	Ca 1	0





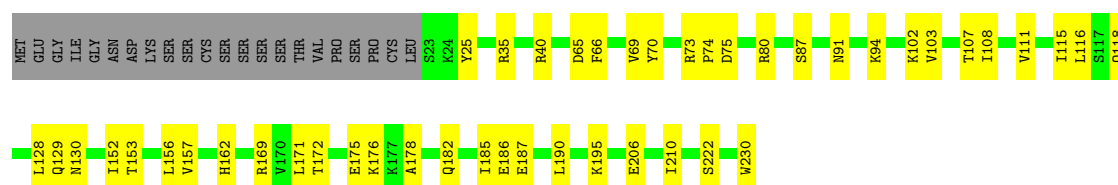
- Molecule 3: DNA-directed RNA polymerase RpoA/D/Rpb3-type domain-containing protein

Chain C: 81% 13% 6%



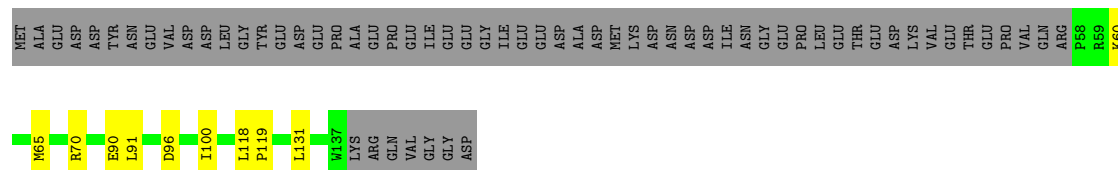
- Molecule 4: RNA polymerase subunit H/Rpb5 C-terminal domain-containing protein

Chain E: 70% 20% 10%



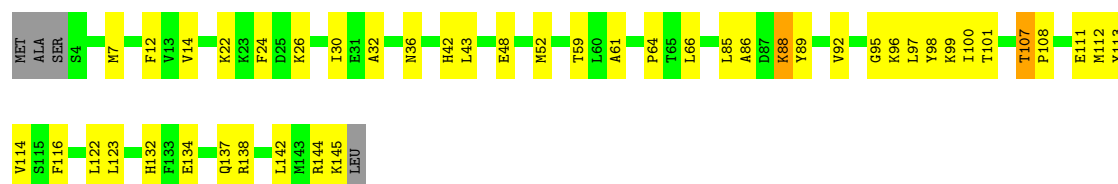
- Molecule 5: DNA-directed RNA polymerase RpoA/D/Rpb3-type domain-containing protein

Chain F: 49% 7% 44%



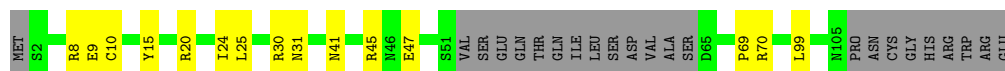
- Molecule 6: DNA-directed RNA polymerase II, IV and V subunit 8

Chain H: 66% 29% 5%



- Molecule 7: DNA-directed RNA polymerase subunit

Chain I:  67% 13% 20%



- Molecule 8: DNA-directed RNA polymerase II, IV and V subunit 10

Chain J:  79% 11% 10%



- Molecule 9: DNA-directed RNA polymerase RBP11-like dimerisation domain-containing protein

Chain K:  83% 10% 7%



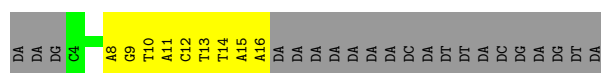
- Molecule 10: DNA-directed RNA polymerase II, IV and V subunit 12

Chain L:  71% 20% 10%

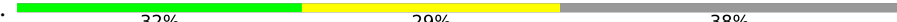


- Molecule 11: DNA (34-MER)

Chain T:  12% 26% 62%



- Molecule 12: DNA (34-MER)

Chain N:  32% 29% 38%



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	62139	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)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	1.5625	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k 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, CA

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	A	0.15	0/6788	0.39	0/9165
2	B	0.18	0/8156	0.40	1/10992 (0.0%)
3	C	0.17	0/2371	0.37	0/3204
4	E	0.13	0/1724	0.35	0/2321
5	F	0.19	0/673	0.38	0/905
6	H	0.20	0/1161	0.47	0/1565
7	I	0.14	0/726	0.39	0/979
8	J	0.21	0/522	0.42	0/706
9	K	0.17	0/905	0.40	0/1221
10	L	0.21	0/379	0.53	0/506
11	T	0.21	0/295	0.47	0/452
12	N	0.33	0/483	0.45	0/743
All	All	0.17	0/24183	0.40	1/32759 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	332	GLU	N-CA-C	-5.12	108.30	114.75

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	A	6672	0	6729	102	0
2	B	7995	0	7910	94	0
3	C	2340	0	2337	27	0
4	E	1698	0	1748	29	0
5	F	660	0	669	7	0
6	H	1135	0	1133	34	0
7	I	713	0	655	21	0
8	J	514	0	521	7	0
9	K	887	0	876	7	0
10	L	375	0	373	8	0
11	T	264	0	148	12	0
12	N	431	0	237	15	0
13	A	1	0	0	0	0
13	C	1	0	0	0	0
13	H	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
14	A	1	0	0	0	0
All	All	23692	0	23336	321	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 321 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:1018:GLU:CG	7:I:20:ARG:HH22	1.34	1.37
1:A:1018:GLU:CB	7:I:20:ARG:HH22	1.47	1.26
1:A:1018:GLU:HB3	7:I:20:ARG:NH2	1.65	1.12
1:A:1018:GLU:OE2	7:I:20:ARG:NH2	1.83	1.11
1:A:1018:GLU:CG	7:I:20:ARG:NH2	2.14	1.11

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	852/2032 (42%)	764 (90%)	87 (10%)	1 (0%)	48	78
2	B	1000/1169 (86%)	891 (89%)	109 (11%)	0	100	100
3	C	298/319 (93%)	279 (94%)	19 (6%)	0	100	100
4	E	206/230 (90%)	189 (92%)	17 (8%)	0	100	100
5	F	78/144 (54%)	73 (94%)	5 (6%)	0	100	100
6	H	140/146 (96%)	119 (85%)	21 (15%)	0	100	100
7	I	87/114 (76%)	80 (92%)	7 (8%)	0	100	100
8	J	62/71 (87%)	55 (89%)	7 (11%)	0	100	100
9	K	106/116 (91%)	96 (91%)	10 (9%)	0	100	100
10	L	44/51 (86%)	33 (75%)	11 (25%)	0	100	100
All	All	2873/4392 (65%)	2579 (90%)	293 (10%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	419	HIS

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	756/1709 (44%)	745 (98%)	11 (2%)	57	69
2	B	884/1026 (86%)	868 (98%)	16 (2%)	51	66
3	C	263/276 (95%)	260 (99%)	3 (1%)	65	73
4	E	189/209 (90%)	185 (98%)	4 (2%)	47	64
5	F	72/128 (56%)	70 (97%)	2 (3%)	38	59
6	H	124/127 (98%)	121 (98%)	3 (2%)	43	62
7	I	77/101 (76%)	76 (99%)	1 (1%)	61	71

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
8	J	57/63 (90%)	55 (96%)	2 (4%)	32	55
9	K	98/105 (93%)	96 (98%)	2 (2%)	48	65
10	L	41/46 (89%)	39 (95%)	2 (5%)	22	48
All	All	2561/3790 (68%)	2515 (98%)	46 (2%)	51	66

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	196	ILE
5	F	90	GLU
3	C	204	LEU
4	E	222	SER
6	H	88	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	867	HIS
2	B	892	ASN
10	L	28	GLN
7	I	41	ASN
9	K	85	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 oligosaccharides in this entry.

5.6 Ligand geometry

Of 8 ligands modelled in this entry, 8 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.