



## wwPDB EM Validation Summary Report ⓘ

May 18, 2025 – 12:31 AM JST

PDB ID : 9K41 / pdb\_00009k41  
EMDB ID : EMD-62038  
Title : Cryo-EM structure of Arabidopsis thaliana H2A.W-nucleosome with Arabidopsis native 147bp DNA 15.2.2 (C2 symmetry)  
Authors : Wang, Y.; Dong, A.  
Deposited on : 2024-10-21  
Resolution : 2.81 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](mailto: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 Phenix2.0rc1  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.43.1

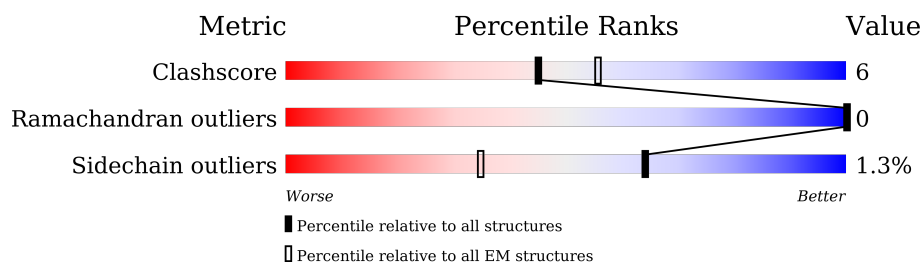
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 2.81 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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  $\geq 3$ , 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	136	65% 5% 29%
1	E	136	63% 7% 29%
2	B	103	71% 7% 22%
2	F	103	73% 5% 22%
3	C	150	58% 11% 31%
3	G	150	61% 9% 31%
4	D	148	55% 7% 37%
4	H	148	57% 6% 37%
5	I	147	63% 36% .

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	J	147	<div><div></div><div>51%</div><div>48%</div><div></div></div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11828 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 Histone H3.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	96	Total	C	N	O	S	0	0
			783	497	149	135	2		
1	E	96	Total	C	N	O	S	0	0
			783	497	149	135	2		

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	80	Total	C	N	O	S	0	0
			644	406	127	110	1		
2	F	80	Total	C	N	O	S	0	0
			644	406	127	110	1		

- Molecule 3 is a protein called Probable histone H2A.7.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	104	Total	C	N	O	S	0	0
			785	501	146	136	2		
3	G	104	Total	C	N	O	S	0	0
			785	501	146	136	2		

- Molecule 4 is a protein called Histone H2B.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	93	Total	C	N	O	S	0	0
			731	468	125	136	2		
4	H	93	Total	C	N	O	S	0	0
			731	468	125	136	2		

- Molecule 5 is a DNA chain called 15.2.2 DNA (147-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
5	I	145	Total	C	N	O	P	0	0
			2950	1413	522	870	145		

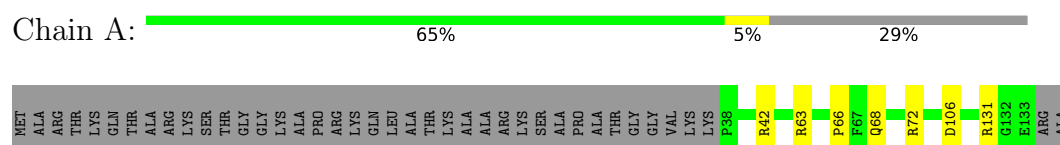
- Molecule 6 is a DNA chain called 15.2.2 DNA (147-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	J	145	Total	C	N	O	P	0	0
			2992	1428	552	868	144		

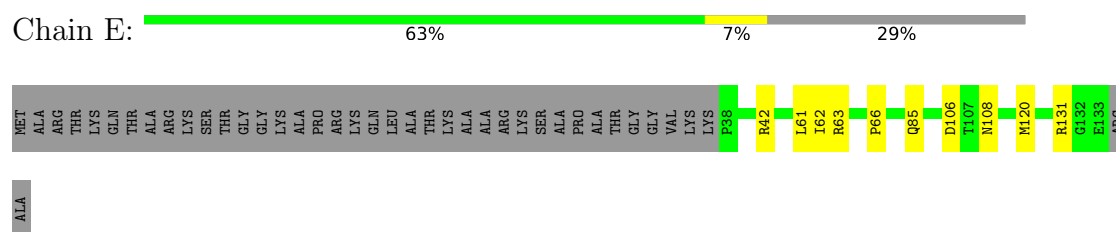
### 3 Residue-property plots [i](#)

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.

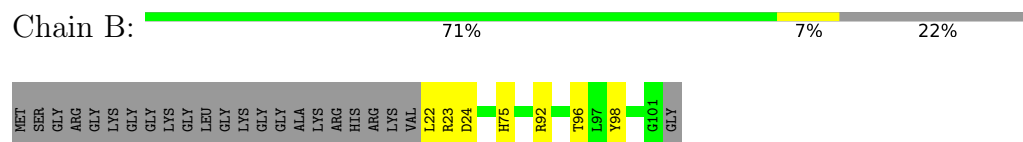
#### • Molecule 1: Histone H3.1



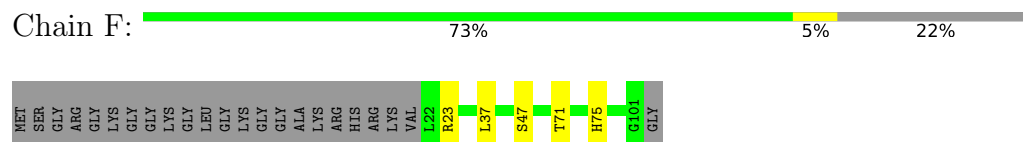
#### • Molecule 1: Histone H3.1



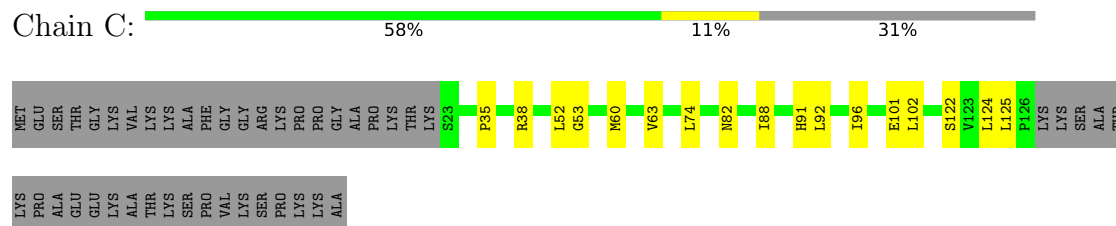
#### • Molecule 2: Histone H4



#### • Molecule 2: Histone H4

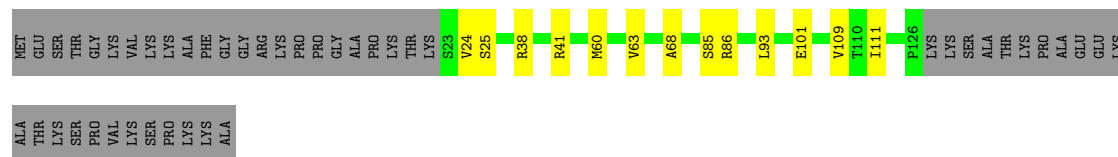


#### • Molecule 3: Probable histone H2A.7



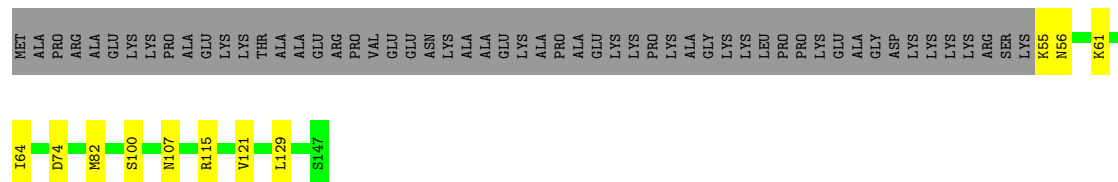
- Molecule 3: Probable histone H2A.7

Chain G: 



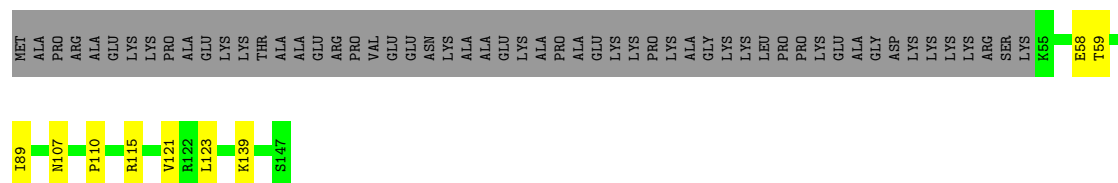
- Molecule 4: Histone H2B.1

Chain D: 



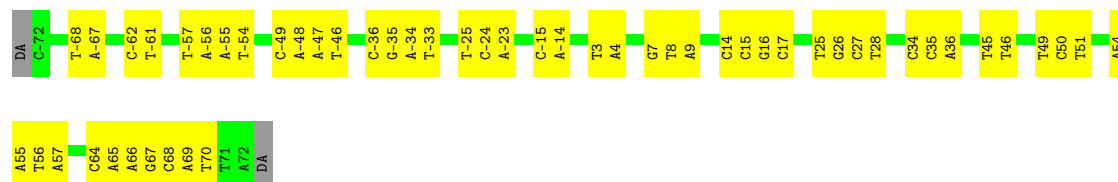
- Molecule 4: Histone H2B.1

Chain H: 



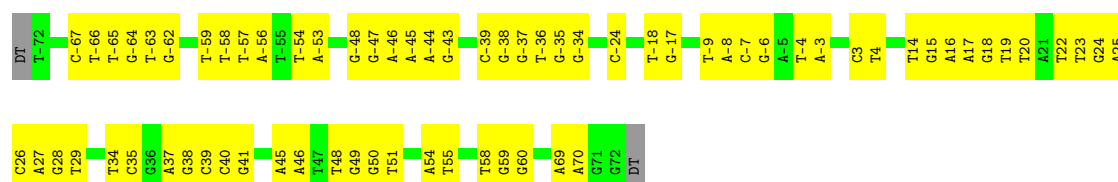
- Molecule 5: 15.2.2 DNA (147-MER)

Chain I: 



- Molecule 6: 15.2.2 DNA (147-MER)

Chain J: 



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	119108	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	50	Depositor
Minimum defocus (nm)	900	Depositor
Maximum defocus (nm)	1800	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor



## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.16	0/795	0.21	0/1067
1	E	0.16	0/795	0.22	0/1067
2	B	0.18	0/651	0.33	0/872
2	F	0.19	0/651	0.37	0/872
3	C	0.15	0/796	0.24	0/1072
3	G	0.15	0/796	0.25	0/1072
4	D	0.15	0/742	0.26	0/995
4	H	0.14	0/742	0.23	0/995
5	I	0.20	0/3304	0.43	0/5091
6	J	0.19	0/3361	0.38	0/5192
All	All	0.18	0/12633	0.35	0/18295

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 [i](#)

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	783	0	822	6	0
1	E	783	0	822	7	0
2	B	644	0	686	6	0
2	F	644	0	686	7	0
3	C	785	0	841	13	0
3	G	785	0	841	13	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	731	0	771	9	0
4	H	731	0	771	10	0
5	I	2950	0	1639	38	0
6	J	2992	0	1640	51	0
All	All	11828	0	9519	129	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 6.

The worst 5 of 129 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)
4:D:55:LYS:NZ	5:I:-46:DT:OP1	2.21	0.71
3:G:41:ARG:NH2	4:H:58:GLU:OE1	2.27	0.67
1:E:106:ASP:OD2	1:E:131:ARG:NH1	2.31	0.64
1:A:106:ASP:OD2	1:A:131:ARG:NH1	2.31	0.64
5:I:-36:DC:H2''	5:I:-35:DG:C8	2.33	0.64

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	
1	A	94/136 (69%)	92 (98%)	2 (2%)	0	100	100
1	E	94/136 (69%)	92 (98%)	2 (2%)	0	100	100
2	B	78/103 (76%)	77 (99%)	1 (1%)	0	100	100
2	F	78/103 (76%)	76 (97%)	2 (3%)	0	100	100
3	C	102/150 (68%)	101 (99%)	1 (1%)	0	100	100
3	G	102/150 (68%)	100 (98%)	2 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	91/148 (62%)	90 (99%)	1 (1%)	0	100	100
4	H	91/148 (62%)	90 (99%)	1 (1%)	0	100	100
All	All	730/1074 (68%)	718 (98%)	12 (2%)	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	82/109 (75%)	82 (100%)	0	100	100
1	E	82/109 (75%)	81 (99%)	1 (1%)	67	89
2	B	66/79 (84%)	63 (96%)	3 (4%)	23	54
2	F	66/79 (84%)	65 (98%)	1 (2%)	60	85
3	C	81/117 (69%)	81 (100%)	0	100	100
3	G	81/117 (69%)	80 (99%)	1 (1%)	67	89
4	D	82/124 (66%)	80 (98%)	2 (2%)	44	76
4	H	82/124 (66%)	82 (100%)	0	100	100
All	All	622/858 (72%)	614 (99%)	8 (1%)	64	88

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	24	VAL
2	F	23	ARG
4	D	100	SER
4	D	74	ASP
1	E	120	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	82	ASN
3	G	77	ASN
4	D	107	ASN
4	H	107	ASN
2	F	75	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

### 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 oligosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 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.