



wwPDB EM Validation Summary Report ⓘ

Nov 17, 2024 – 12:10 AM JST

PDB ID : 8WG5
EMDB ID : EMD-37503
Title : Cryo-EM structure of USP16 bound to H2AK119Ub nucleosome
Authors : Ai, H.S.; He, Z.Z.; Deng, Z.H.; Liu, L.
Deposited on : 2023-09-20
Resolution : 3.05 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:

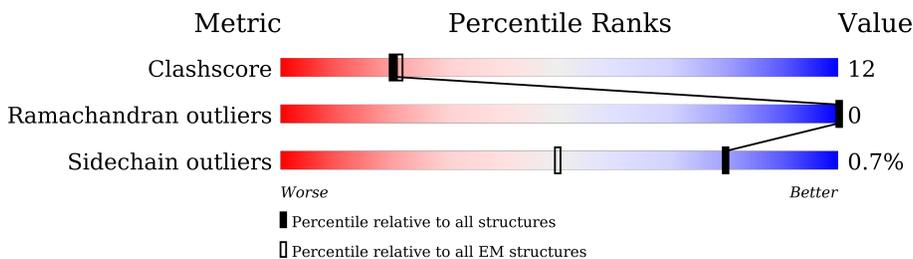
EMDB validation analysis : **FAILED**
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

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 3.05 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 ≥ 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	M	823	23% 12% 64%
2	A	98	70% 30%
2	E	98	79% 20% .
3	B	79	73% 27%
3	F	79	65% 35%
4	C	111	73% 27%
4	G	111	76% 23% .
5	D	94	81% 19%
5	H	94	72% 28%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain	
6	I	147	 46%	54%
7	J	147	 37%	63%
8	U	76	 58%	42%

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 14965 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 Ubiquitin carboxyl-terminal hydrolase 16.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	M	298	2319	1487	395	424	13	0	0

- Molecule 2 is a protein called Histone H3.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	A	98	810	511	157	138	4	0	0
2	E	97	801	505	155	137	4	0	0

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	B	79	627	396	121	109	1	0	0
3	F	79	627	396	121	109	1	0	0

- Molecule 4 is a protein called Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	C	111	849	533	167	148	1	0	0
4	G	109	836	526	165	145		0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	119	CYS	-	expression tag	UNP P04908
C	120	THR	-	expression tag	UNP P04908
G	119	CYS	-	expression tag	UNP P04908

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	120	THR	-	expression tag	UNP P04908

- Molecule 5 is a protein called Histone H2B type 1-K.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	D	94	Total	C	N	O	S	0	0
			735	461	134	138	2		
5	H	94	Total	C	N	O	S	0	0
			735	461	134	138	2		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
6	I	147	Total	C	N	O	P	0	0
			3031	1434	570	880	147		

- Molecule 7 is a DNA chain called DNA (147-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	J	147	Total	C	N	O	P	0	0
			2996	1423	542	884	147		

- Molecule 8 is a protein called Ubiquitin B.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	U	76	Total	C	N	O	S	0	0
			599	376	104	117	2		

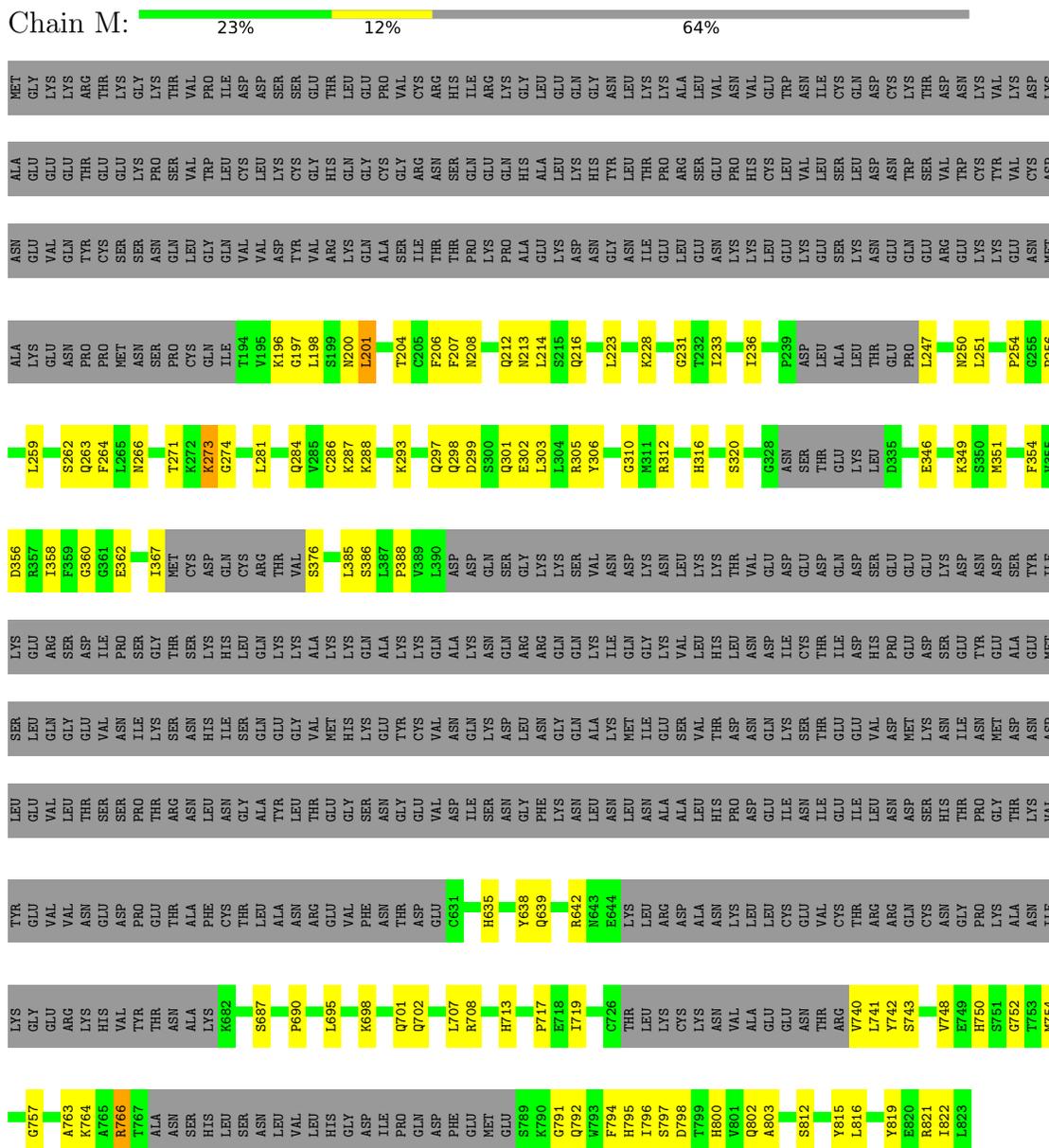
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
U	76	CYS	-	expression tag	UNP J3QS39

3 Residue-property plots

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: Ubiquitin carboxyl-terminal hydrolase 16



- Molecule 2: Histone H3.1

Chain A:  70% 30%



- Molecule 2: Histone H3.1

Chain E:  79% 20%



- Molecule 3: Histone H4

Chain B:  73% 27%



- Molecule 3: Histone H4

Chain F:  65% 35%



- Molecule 4: Histone H2A type 1-B/E

Chain C:  73% 27%



- Molecule 4: Histone H2A type 1-B/E

Chain G:  76% 23%



- Molecule 5: Histone H2B type 1-K

Chain D:  81% 19%



- Molecule 5: Histone H2B type 1-K

Chain H:  72% 28%



- Molecule 6: DNA (147-MER)



- Molecule 7: DNA (147-MER)



- Molecule 8: Ubiquitin B



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	144457	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1300	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	M	0.25	0/2360	0.45	0/3181
2	A	0.27	0/822	0.57	0/1102
2	E	0.27	0/813	0.59	0/1090
3	B	0.26	0/634	0.57	0/850
3	F	0.29	0/634	0.64	0/850
4	C	0.27	0/859	0.54	0/1160
4	G	0.27	0/846	0.52	0/1142
5	D	0.26	0/746	0.49	0/1003
5	H	0.29	0/746	0.54	0/1003
6	I	0.50	0/3404	0.90	0/5256
7	J	0.49	0/3356	0.89	0/5173
8	U	0.25	0/605	0.57	0/815
All	All	0.38	0/15825	0.72	0/22625

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	M	2319	0	2307	69	0
2	A	810	0	851	26	0
2	E	801	0	839	15	0
3	B	627	0	662	20	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	F	627	0	662	21	0
4	C	849	0	903	22	0
4	G	836	0	891	23	0
5	D	735	0	756	13	0
5	H	735	0	756	25	0
6	I	3031	0	1649	65	0
7	J	2996	0	1651	63	0
8	U	599	0	620	29	0
All	All	14965	0	12547	333	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 12.

The worst 5 of 333 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:I:-10:DG:N2	7:J:10:DC:O2	2.12	0.81
1:M:288:LYS:NZ	1:M:306:TYR:O	2.14	0.80
6:I:-10:DG:N1	7:J:10:DC:N3	2.28	0.76
1:M:719:ILE:HD11	1:M:741:LEU:HB3	1.69	0.75
5:H:80:LEU:HA	5:H:83:TYR:HD2	1.52	0.74

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	M	282/823 (34%)	273 (97%)	9 (3%)	0	100 100
2	A	96/98 (98%)	92 (96%)	4 (4%)	0	100 100
2	E	95/98 (97%)	93 (98%)	2 (2%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	B	77/79 (98%)	77 (100%)	0	0	100	100
3	F	77/79 (98%)	74 (96%)	3 (4%)	0	100	100
4	C	109/111 (98%)	107 (98%)	2 (2%)	0	100	100
4	G	107/111 (96%)	105 (98%)	2 (2%)	0	100	100
5	D	92/94 (98%)	91 (99%)	1 (1%)	0	100	100
5	H	92/94 (98%)	88 (96%)	4 (4%)	0	100	100
8	U	74/76 (97%)	72 (97%)	2 (3%)	0	100	100
All	All	1101/1663 (66%)	1072 (97%)	29 (3%)	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	
1	M	253/749 (34%)	249 (98%)	4 (2%)	58	76
2	A	86/86 (100%)	86 (100%)	0	100	100
2	E	85/86 (99%)	85 (100%)	0	100	100
3	B	64/65 (98%)	64 (100%)	0	100	100
3	F	64/65 (98%)	64 (100%)	0	100	100
4	C	86/87 (99%)	85 (99%)	1 (1%)	67	82
4	G	84/87 (97%)	83 (99%)	1 (1%)	67	82
5	D	80/80 (100%)	79 (99%)	1 (1%)	65	80
5	H	80/80 (100%)	80 (100%)	0	100	100
8	U	68/69 (99%)	68 (100%)	0	100	100
All	All	950/1454 (65%)	943 (99%)	7 (1%)	80	89

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	M	766	ARG
4	C	11	ARG
4	G	11	ARG
5	D	108	LYS
1	M	708	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	93	GLN
4	G	38	ASN

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.