



wwPDB/EMDatabank EM Map/Model Validation Summary Report

Sep 4, 2019 – 12:26 PM EDT


PDB ID : 6SEG
EMDB ID: : EMD-10155
Title : Class1: CENP-A nucleosome in complex with CENP-C central region
Authors : Ali-Ahmad, A.; Bilokapic, S.; Schafer, I.B.; Halic, M.; Sekulic, N.
Deposited on : 2019-07-30
Resolution : 3.10 Å(reported)

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB 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.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : 2.4

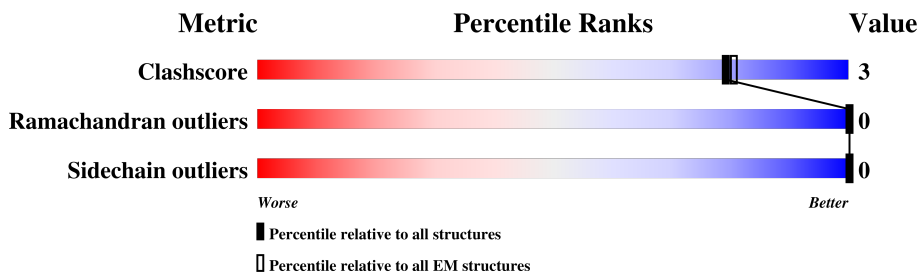
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.10 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	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	140	59% 9% 32%
1	E	140	61% 6% 32%
2	B	103	68% 8% 24%
2	F	103	69% 7% 24%
3	C	130	70% . 26%
3	G	130	73% 7% 20%
4	D	126	67% 6% 27%
4	H	126	65% 6% 29%
5	I	145	83% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	J	145	 85% 14%

2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 11689 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-like centromeric protein A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	95	779	506	143	129	1	0	0
1	E	95	779	506	143	129	1	0	0

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	78	618	391	120	106	1	0	0
2	F	78	618	391	120	106	1	0	0

- Molecule 3 is a protein called Histone H2A type 2-A.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	96	739	464	144	130	1	0	0
3	G	104	799	501	158	139	1	0	0

- Molecule 4 is a protein called Histone H2B type 1-C/E/F/G/I.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	92	719	452	129	136	2	0	0
4	H	90	699	440	123	134	2	0	0

- Molecule 5 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
5	I	145	2952	1404	537	867	144	0	0

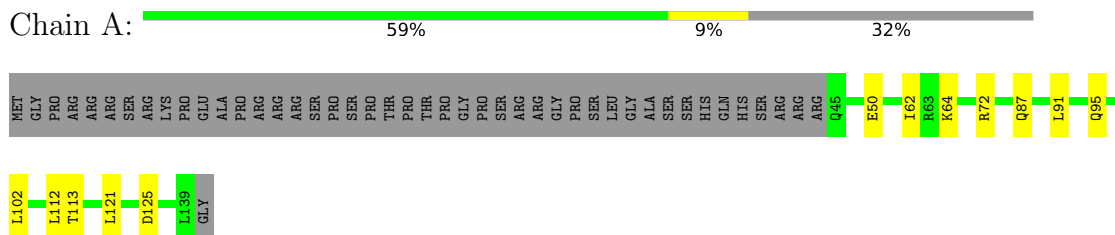
- Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
6	J	145	2987	1416	558	869	144	0	0

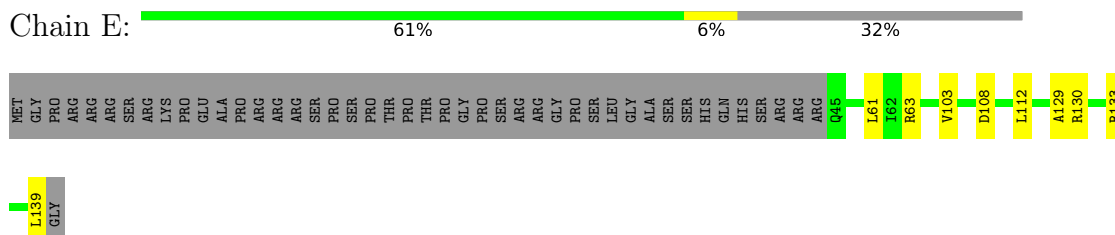
3 Residue-property plots i

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

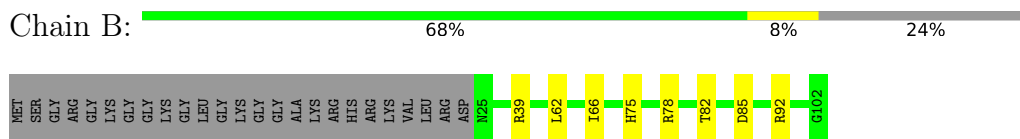
- Molecule 1: Histone H3-like centromeric protein A



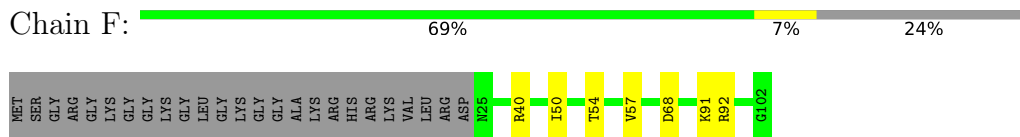
- Molecule 1: Histone H3-like centromeric protein A



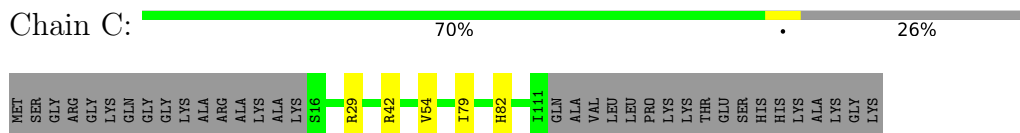
- Molecule 2: Histone H4



- Molecule 2: Histone H4

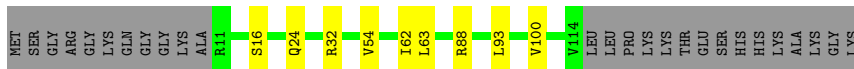


- Molecule 3: Histone H2A type 2-A



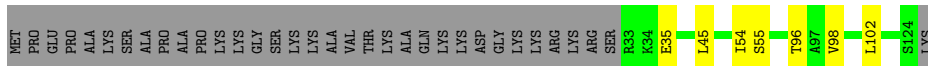
- Molecule 3: Histone H2A type 2-A

Chain G:  73% 7% 20%



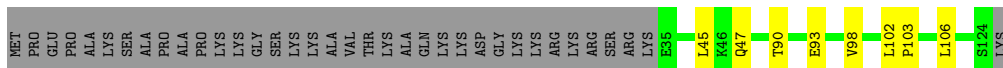
- Molecule 4: Histone H2B type 1-C/E/F/G/I

Chain D:  67% 6% 27%




- Molecule 4: Histone H2B type 1-C/E/F/G/I

Chain H:  65% 6% 29%




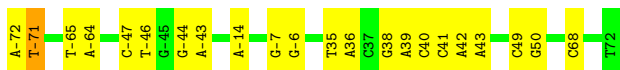
- Molecule 5: DNA (145-MER)

Chain I:  83% 17%



- Molecule 6: DNA (145-MER)

Chain J:  85% 14%



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	320000	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 ($e^-/\text{\AA}^2$)	100	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k 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 >2	RMSZ	# Z >2
1	A	0.34	0/794	0.58	1/1069 (0.1%)
1	E	0.33	0/794	0.52	0/1069
2	B	0.33	0/625	0.53	0/837
2	F	0.33	0/625	0.53	0/837
3	C	0.34	0/748	0.53	0/1008
3	G	0.33	0/808	0.54	0/1087
4	D	0.35	0/730	0.52	0/982
4	H	0.35	0/710	0.51	0/957
5	I	0.72	0/3308	1.04	4/5099 (0.1%)
6	J	0.71	0/3354	1.01	2/5180 (0.0%)
All	All	0.57	0/12496	0.85	7/18125 (0.0%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	121	LEU	CA-CB-CG	6.40	130.03	115.30
5	I	70	DG	O4'-C4'-C3'	-5.62	102.25	104.50
5	I	-35	DA	O4'-C1'-N9	-5.44	104.19	108.00
5	I	-68	DG	O4'-C1'-N9	5.39	111.77	108.00
6	J	-71	DT	O4'-C1'-N1	5.17	111.62	108.00

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	779	0	815	9	0
1	E	779	0	815	9	0
2	B	618	0	659	6	0
2	F	618	0	659	7	0
3	C	739	0	780	6	0
3	G	799	0	851	8	0
4	D	719	0	738	7	0
4	H	699	0	712	6	0
5	I	2952	0	1629	11	0
6	J	2987	0	1630	13	0
All	All	11689	0	9288	63	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 3.

The worst 5 of 63 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:E:129:ALA:O	1:E:133:ARG:HB2	1.79	0.83
3:C:29:ARG:HH21	4:D:35:GLU:HB2	1.55	0.69
3:G:16:SER:HA	6:J:-43:DA:H5''	1.78	0.65
1:A:112:LEU:HD21	1:E:112:LEU:HD22	1.78	0.64
1:A:72:ARG:HH22	1:A:91:LEU:HD11	1.66	0.60

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	93/140 (66%)	93 (100%)	0	0	100 100
1	E	93/140 (66%)	93 (100%)	0	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	76/103 (74%)	75 (99%)	1 (1%)	0	100	100
2	F	76/103 (74%)	74 (97%)	2 (3%)	0	100	100
3	C	94/130 (72%)	93 (99%)	1 (1%)	0	100	100
3	G	102/130 (78%)	98 (96%)	4 (4%)	0	100	100
4	D	90/126 (71%)	89 (99%)	1 (1%)	0	100	100
4	H	88/126 (70%)	88 (100%)	0	0	100	100
All	All	712/998 (71%)	703 (99%)	9 (1%)	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	A	80/118 (68%)	80 (100%)	0	100	100
1	E	80/118 (68%)	80 (100%)	0	100	100
2	B	63/79 (80%)	63 (100%)	0	100	100
2	F	63/79 (80%)	63 (100%)	0	100	100
3	C	75/99 (76%)	75 (100%)	0	100	100
3	G	80/99 (81%)	80 (100%)	0	100	100
4	D	79/106 (74%)	79 (100%)	0	100	100
4	H	77/106 (73%)	77 (100%)	0	100	100
All	All	597/804 (74%)	597 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
1	A	89	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	G	110	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 carbohydrates 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.