



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jan 14, 2018 – 10:21 PM EST

PDB ID : 6ESI  
EMDB ID: : EMD-3950  
Title : Nucleosome breathing : Class 4  
Authors : Bilokapic, S.; Halic, M.  
Deposited on : 2017-10-20  
Resolution : 6.30 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

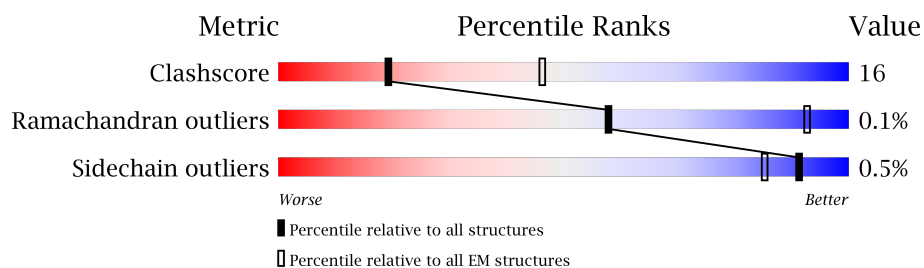
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 6.30 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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  $\geq 3$ , 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	135	
1	E	135	
2	B	102	
2	F	102	
3	C	129	
3	G	129	
4	D	122	
5	H	122	
6	I	147	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
7	J	147	 A horizontal bar chart showing the quality of chain J. The bar is divided into three segments: green (48%), yellow (39%), and orange (10%). The segments are labeled with their respective percentages: 48%, 39%, and 10%.

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 11173 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 H3.2.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	98	Total	C	N	O	S	0	0
			811	512	157	139	3		
1	E	87	Total	C	N	O	S	0	0
			721	455	137	126	3		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	ALA	GLY	variant	UNP P84233
E	102	ALA	GLY	variant	UNP P84233

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	78	Total	C	N	O	S	0	0
			619	391	120	107	1		
2	F	82	Total	C	N	O	S	0	0
			657	416	128	112	1		

- Molecule 3 is a protein called Histone H2A.

Mol	Chain	Residues	Atoms				AltConf	Trace
3	C	87	Total	C	N	O	0	0
			679	426	135	118		
3	G	102	Total	C	N	O	0	0
			786	495	153	138		

- Molecule 4 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	92	Total	C	N	O	S	0	0
			719	453	129	135	2		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	29	THR	SER	variant	UNP P02281
D	122	ARG	LYS	variant	UNP P02281

- Molecule 5 is a protein called Histone H2B 1.1.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	H	93	Total	C	N	O	S	0	0
			728	459	131	136	2		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	29	THR	SER	variant	UNP P02281

- Molecule 6 is a DNA chain called DNA (133-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	I	133	Total	C	N	O	P	0	0
			2740	1295	514	798	133		

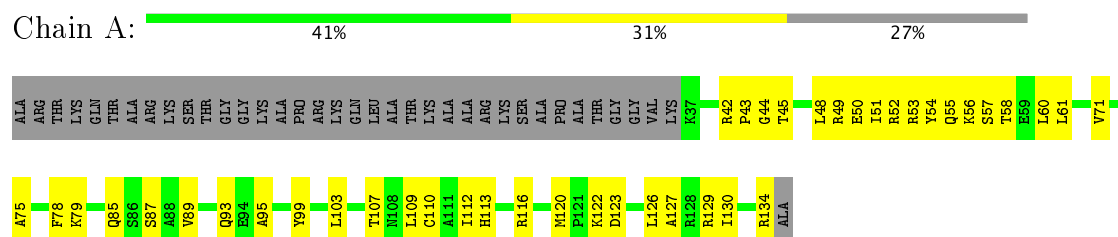
- Molecule 7 is a DNA chain called DNA (133-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	J	133	Total	C	N	O	P	0	0
			2713	1286	496	798	133		

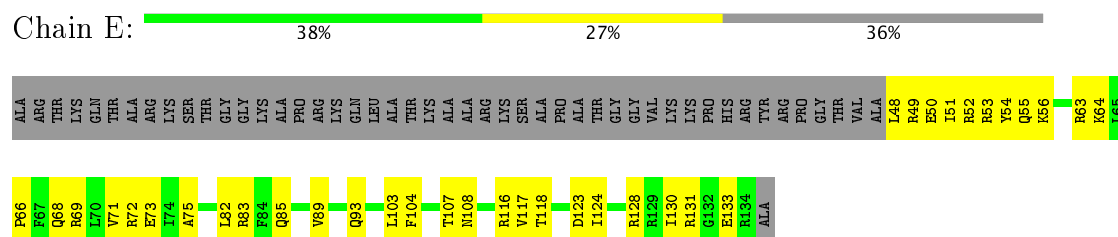
### 3 Residue-property plots

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.

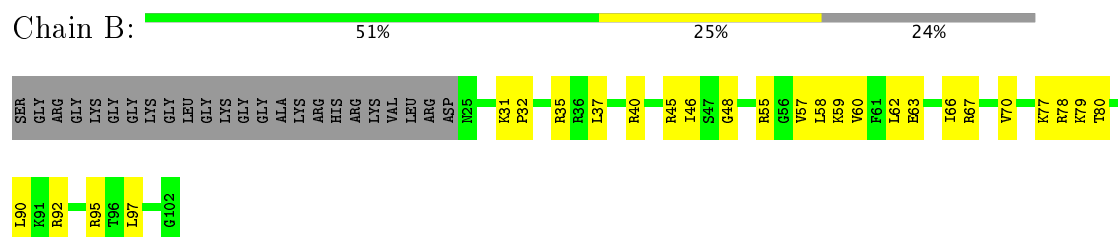
#### • Molecule 1: Histone H3.2



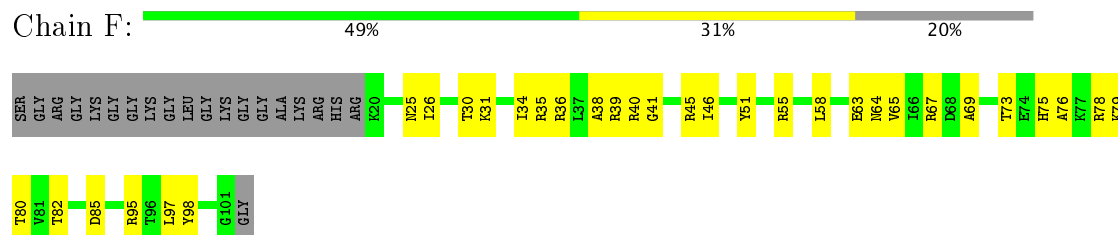
#### • Molecule 1: Histone H3.2



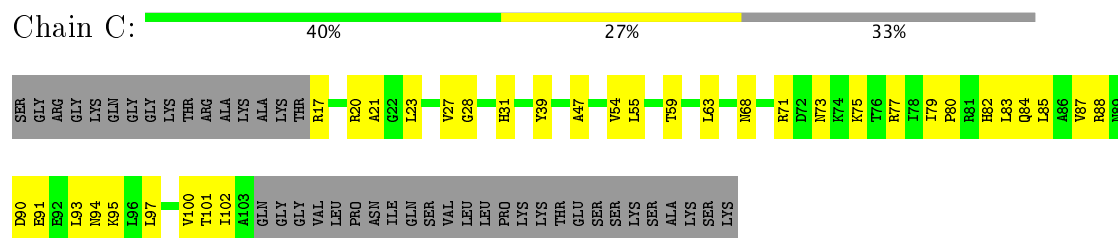
#### • Molecule 2: Histone H4



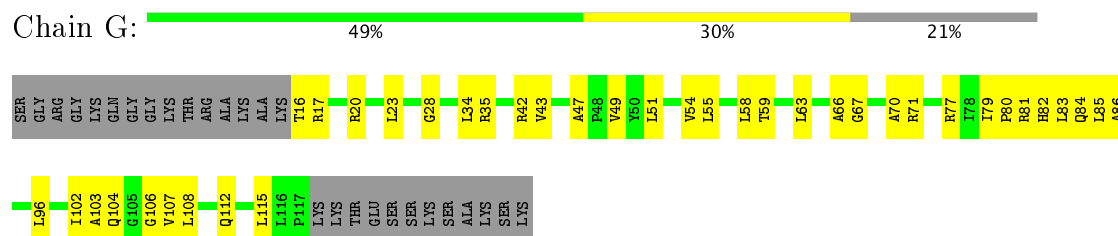
#### • Molecule 2: Histone H4



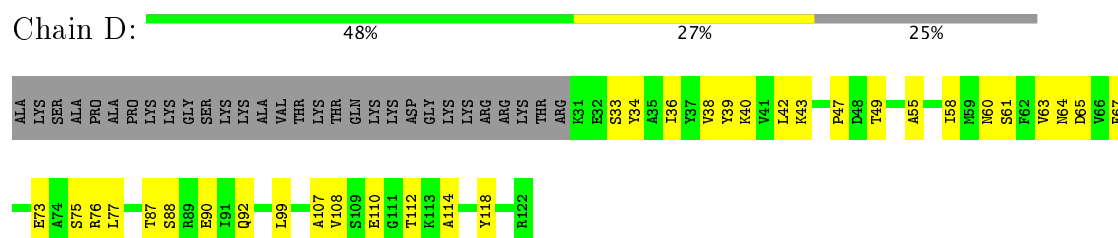
#### • Molecule 3: Histone H2A



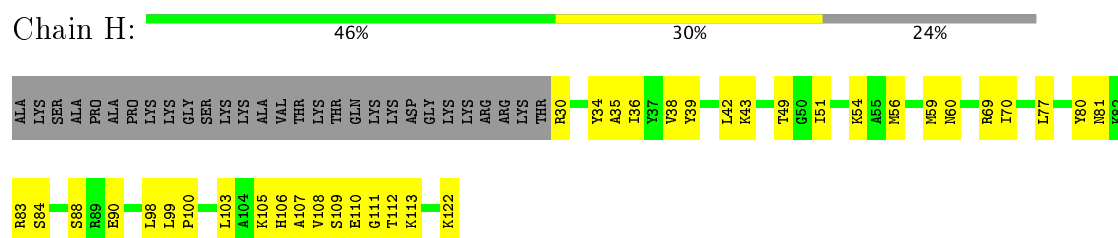
- Molecule 3: Histone H2A



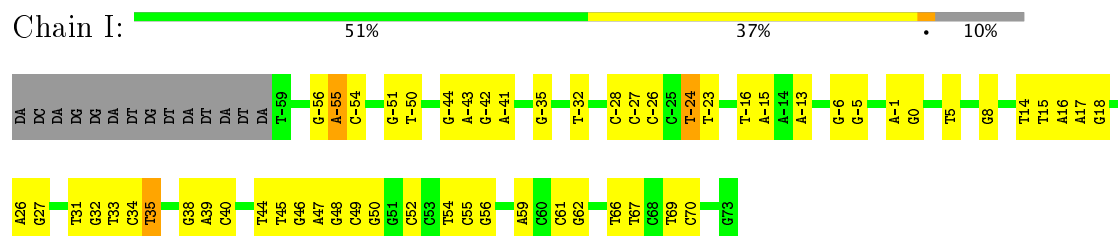
- Molecule 4: Histone H2B 1.1



- Molecule 5: Histone H2B 1.1

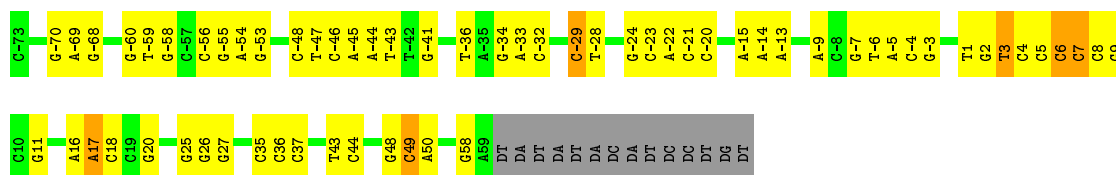


- Molecule 6: DNA (133-MER)



- Molecule 7: DNA (133-MER)







## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	28000	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	80, 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), FEI FALCON II (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	0.33	0/823	0.66	0/1104
1	E	0.35	0/728	0.65	0/973
2	B	0.35	0/626	0.64	0/837
2	F	0.31	0/664	0.63	0/889
3	C	0.32	0/687	0.60	0/926
3	G	0.33	0/795	0.62	0/1074
4	D	0.33	0/730	0.61	0/983
5	H	0.32	0/739	0.60	0/994
6	I	0.95	0/3076	1.09	5/4749 (0.1%)
7	J	0.95	4/3040 (0.1%)	1.10	15/4686 (0.3%)
All	All	0.72	4/11908 (0.0%)	0.91	20/17215 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	J	-7	DG	C3'-O3'	-6.43	1.35	1.44
7	J	3	DT	C3'-O3'	-5.57	1.36	1.44
7	J	-22	DA	N9-C4	-5.41	1.34	1.37
7	J	-14	DA	C3'-O3'	-5.26	1.37	1.44

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	I	26	DA	O4'-C4'-C3'	-7.67	101.40	106.00
6	I	52	DC	O4'-C1'-N1	7.61	113.33	108.00
7	J	49	DC	O5'-P-OP2	-7.50	98.95	105.70
7	J	-29	DC	O4'-C1'-N1	7.01	112.90	108.00
7	J	7	DC	O4'-C1'-N1	6.54	112.58	108.00

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	811	0	853	38	0
1	E	721	0	759	39	0
2	B	619	0	659	23	0
2	F	657	0	706	31	0
3	C	679	0	717	31	0
3	G	786	0	832	34	0
4	D	719	0	740	30	0
5	H	728	0	753	32	0
6	I	2740	0	1491	61	0
7	J	2713	0	1491	57	0
All	All	11173	0	9001	289	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 16.

The worst 5 of 289 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:51:ILE:O	1:A:55:GLN:HB3	1.60	1.00
4:D:36:ILE:O	4:D:40:LYS:HB2	1.64	0.98
1:E:51:ILE:O	1:E:55:GLN:HB3	1.70	0.91
2:B:58:LEU:O	2:B:62:LEU:HB2	1.72	0.90
2:F:35:ARG:O	2:F:39:ARG:HB2	1.74	0.87

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	96/135 (71%)	92 (96%)	4 (4%)	0	100	100
1	E	83/135 (62%)	78 (94%)	5 (6%)	0	100	100
2	B	76/102 (74%)	72 (95%)	4 (5%)	0	100	100
2	F	80/102 (78%)	76 (95%)	4 (5%)	0	100	100
3	C	85/129 (66%)	79 (93%)	6 (7%)	0	100	100
3	G	98/129 (76%)	89 (91%)	9 (9%)	0	100	100
4	D	90/122 (74%)	85 (94%)	5 (6%)	0	100	100
5	H	91/122 (75%)	86 (94%)	4 (4%)	1 (1%)	17	59
All	All	699/976 (72%)	657 (94%)	41 (6%)	1 (0%)	58	88

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	H	100	PRO

### 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	86/110 (78%)	86 (100%)	0	100	100
1	E	77/110 (70%)	76 (99%)	1 (1%)	73	87
2	B	63/78 (81%)	62 (98%)	1 (2%)	68	85
2	F	68/78 (87%)	68 (100%)	0	100	100
3	C	68/101 (67%)	68 (100%)	0	100	100
3	G	81/101 (80%)	81 (100%)	0	100	100
4	D	78/102 (76%)	77 (99%)	1 (1%)	73	87
5	H	79/102 (78%)	79 (100%)	0	100	100
All	All	600/782 (77%)	597 (100%)	3 (0%)	91	95

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	37	LEU
4	D	99	LEU
1	E	93	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	93	GLN
1	A	108	ASN
1	E	93	GLN
2	F	75	HIS
3	G	112	GLN

### 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	G	1
1	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	G	40:ALA	C	41:GLU	N	8.41
1	E	80:THR	C	81:ASP	N	3.80