



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 07:00 PM UTC

PDB ID : 5AVB / pdb_00005avb
Title : human nucleosome core particle
Authors : Wakamori, M.; Fujii, Y.; Umehara, T.; Yokoyama, S.
Deposited on : 2015-06-12
Resolution : 2.40 Å(reported)

This is a wwPDB X-ray Structure 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/XrayValidationReportHelp>

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:

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

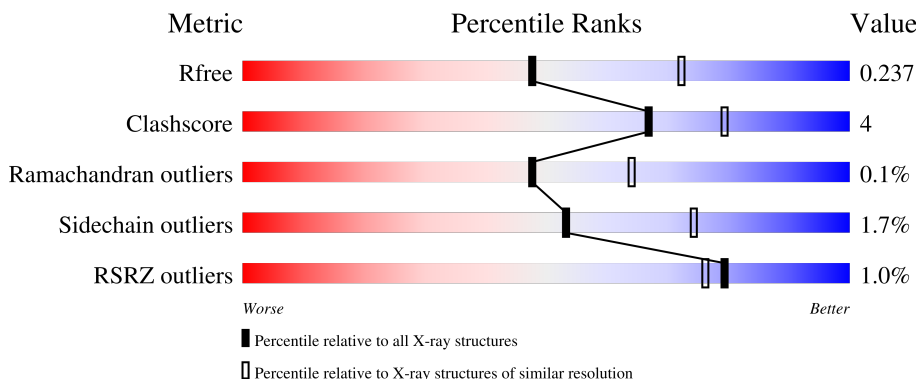
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	139	0.1% 62% 9% • 29%
1	E	139	0.1% 66% • 29%
2	B	104	2% 67% 11% • 21%
2	F	104	3% 75% 9% 16%
3	C	133	74% 5% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	133	 71% 8% 21%
4	D	129	 2% 68% 7% 25%
4	H	129	 65% 8% 27%
5	I	147	 85% 15%
6	J	147	 77% 23%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 12350 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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					ZeroOcc	AltConf	Trace
1	A	99	Total	C	N	O	S	0	0	0
			815	514	158	139	4			
1	E	99	Total	C	N	O	S	0	0	0
			816	514	158	140	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	GLY	-	expression tag	UNP P68431
A	-2	SER	-	expression tag	UNP P68431
A	-1	HIS	-	expression tag	UNP P68431
E	-3	GLY	-	expression tag	UNP P68431
E	-2	SER	-	expression tag	UNP P68431
E	-1	HIS	-	expression tag	UNP P68431

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	82	Total	C	N	O	S	0	0	0
			653	412	127	113	1			
2	F	87	Total	C	N	O	S	0	0	0
			703	442	142	118	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP P62805
F	-1	GLY	-	expression tag	UNP P62805

- Molecule 3 is a protein called Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	106	Total	C	N	O	0	0	0
			819	517	160	142			
3	G	105	Total	C	N	O	0	0	0
			810	511	158	141			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP P04908
C	-2	SER	-	expression tag	UNP P04908
C	-1	HIS	-	expression tag	UNP P04908
G	-3	GLY	-	expression tag	UNP P04908
G	-2	SER	-	expression tag	UNP P04908
G	-1	HIS	-	expression tag	UNP P04908

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	97	Total	C	N	O	S	0	0	0
			766	480	142	142	2			
4	H	94	Total	C	N	O	S	0	0	0
			736	462	134	138	2			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-3	GLY	-	expression tag	UNP P06899
D	-2	SER	-	expression tag	UNP P06899
D	-1	HIS	-	expression tag	UNP P06899
H	-3	GLY	-	expression tag	UNP P06899
H	-2	SER	-	expression tag	UNP P06899
H	-1	HIS	-	expression tag	UNP P06899

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	147	Total	C	N	O	P	0	0	0
			3011	1440	546	879	146			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	J	147	Total	C	N	O	P	0	0	0
			3010	1440	543	881	146			

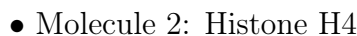
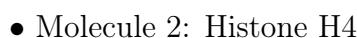
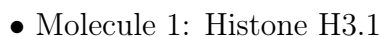
- Molecule 7 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	E	1	Total	Mn	0	0
			1	1		
7	I	7	Total	Mn	0	0
			7	7		
7	J	6	Total	Mn	0	0
			6	6		

- Molecule 8 is water.

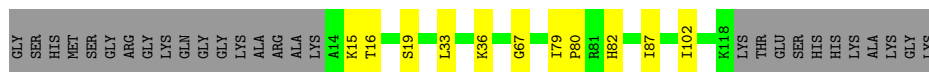
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	25	Total	O	0	0
			25	25		
8	B	20	Total	O	0	0
			20	20		
8	C	28	Total	O	0	0
			28	28		
8	D	14	Total	O	0	0
			14	14		
8	E	33	Total	O	0	0
			33	33		
8	F	29	Total	O	0	0
			29	29		
8	G	16	Total	O	0	0
			16	16		
8	H	6	Total	O	0	0
			6	6		
8	I	11	Total	O	0	0
			11	11		
8	J	15	Total	O	0	0
			15	15		

- Molecule 1: Histone H3.1



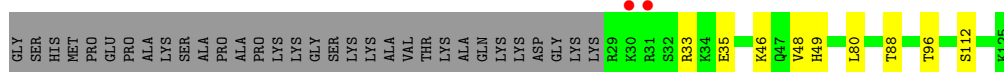
- Molecule 3: Histone H2A type 1-B/E

Chain G:  71% 8% 21%



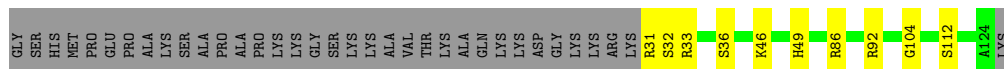
- Molecule 4: Histone H2B type 1-J

Chain D:  2% 68% 7% 25%



- Molecule 4: Histone H2B type 1-J

Chain H:  65% 8% 27%



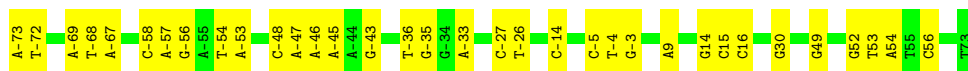
- Molecule 5: DNA (147-MER)

Chain I:  85% 15%



- Molecule 6: DNA (147-MER)

Chain J:  77% 23%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	105.95Å 110.09Å 181.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.40 20.00 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.00-2.40) 99.7 (20.00-2.40)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	0.17	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.46 (at 2.41Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1690)	Depositor
R, R_{free}	0.197 , 0.236 0.201 , 0.237	Depositor DCC
R_{free} test set	4168 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	50.0	Xtriage
Anisotropy	0.272	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 46.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.017 for k,h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	12350	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MN

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	0.29	0/827	0.71	1/1109 (0.1%)
1	E	0.32	0/828	0.72	1/1109 (0.1%)
2	B	0.31	0/660	0.74	0/883
2	F	0.32	0/711	0.73	0/948
3	C	0.31	0/829	0.69	0/1118
3	G	0.30	0/820	0.71	0/1107
4	D	0.31	0/777	0.71	0/1040
4	H	0.28	0/747	0.71	0/1004
5	I	0.19	0/3378	0.40	0/5212
6	J	0.19	0/3376	0.41	0/5209
All	All	0.25	0/12953	0.56	2/18739 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	117	VAL	N-CA-C	-5.51	107.42	113.43
1	A	117	VAL	N-CA-C	-5.16	108.10	113.10

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	815	0	856	8	0
1	E	816	0	856	7	0
2	B	653	0	696	8	0
2	F	703	0	755	4	0
3	C	819	0	879	5	0
3	G	810	0	866	7	0
4	D	766	0	797	5	0
4	H	736	0	758	8	0
5	I	3011	0	1662	19	0
6	J	3010	0	1663	30	0
7	E	1	0	0	0	0
7	I	7	0	0	0	0
7	J	6	0	0	0	0
8	A	25	0	0	0	0
8	B	20	0	0	0	0
8	C	28	0	0	0	0
8	D	14	0	0	0	0
8	E	33	0	0	0	0
8	F	29	0	0	0	0
8	G	16	0	0	0	0
8	H	6	0	0	0	0
8	I	11	0	0	0	0
8	J	15	0	0	0	0
All	All	12350	0	9788	79	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 4.

The worst 5 of 79 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:J:-69:DA:H2''	6:J:-68:DT:H5''	1.69	0.73
1:A:60:LEU:HD12	1:A:64:LYS:HE2	1.76	0.68
1:E:69:ARG:NH2	5:I:17:DT:OP2	2.29	0.66
1:A:73:GLU:OE1	2:B:25:ASN:ND2	2.28	0.65
6:J:-68:DT:H2''	6:J:-67:DA:C8	2.37	0.60

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 X-ray entries followed by that with respect to entries of similar resolution.

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	97/139 (70%)	97 (100%)	0	0	100	100
1	E	97/139 (70%)	96 (99%)	1 (1%)	0	100	100
2	B	80/104 (77%)	79 (99%)	1 (1%)	0	100	100
2	F	85/104 (82%)	83 (98%)	2 (2%)	0	100	100
3	C	104/133 (78%)	103 (99%)	1 (1%)	0	100	100
3	G	103/133 (77%)	102 (99%)	1 (1%)	0	100	100
4	D	95/129 (74%)	94 (99%)	1 (1%)	0	100	100
4	H	92/129 (71%)	91 (99%)	0	1 (1%)	11	18
All	All	753/1010 (75%)	745 (99%)	7 (1%)	1 (0%)	48	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	H	104	GLY

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 X-ray entries followed by that with respect to entries of similar resolution.

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	86/113 (76%)	84 (98%)	2 (2%)	44	66
1	E	86/113 (76%)	86 (100%)	0	100	100
2	B	67/79 (85%)	66 (98%)	1 (2%)	57	77
2	F	72/79 (91%)	69 (96%)	3 (4%)	26	45

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	C	84/102 (82%)	84 (100%)	0	100	100
3	G	83/102 (81%)	83 (100%)	0	100	100
4	D	83/107 (78%)	80 (96%)	3 (4%)	31	52
4	H	80/107 (75%)	78 (98%)	2 (2%)	42	64
All	All	641/802 (80%)	630 (98%)	11 (2%)	53	74

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	21	VAL
2	F	47	SER
4	H	112	SER
4	H	36	SER
4	D	88	THR

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

Mol	Chain	Res	Type
3	C	24	GLN
4	D	109	HIS
1	E	68	GLN
4	H	49	HIS
4	H	95	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 14 ligands modelled in this entry, 14 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.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	99/139 (71%)	-0.32	1 (1%) 79 76	34, 48, 71, 88	0
1	E	99/139 (71%)	-0.54	2 (2%) 65 60	29, 39, 63, 101	0
2	B	82/104 (78%)	-0.32	2 (2%) 59 55	36, 47, 73, 134	0
2	F	87/104 (83%)	-0.43	3 (3%) 48 44	30, 40, 68, 113	0
3	C	106/133 (79%)	-0.48	0 100 100	32, 44, 69, 101	0
3	G	105/133 (78%)	-0.29	0 100 100	37, 52, 78, 98	0
4	D	97/129 (75%)	-0.24	2 (2%) 63 59	34, 46, 93, 111	0
4	H	94/129 (72%)	-0.19	0 100 100	36, 52, 87, 114	0
5	I	147/147 (100%)	0.29	1 (0%) 84 81	50, 92, 126, 144	0
6	J	147/147 (100%)	0.31	0 100 100	52, 95, 126, 134	0
All	All	1063/1304 (81%)	-0.17	11 (1%) 79 76	29, 53, 113, 144	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	135	ALA	4.5
2	F	101	GLY	3.9
1	A	135	ALA	3.6
4	D	30	LYS	3.0
2	B	21	VAL	2.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
7	MN	J	106	1/1	0.71	0.21	99,99,99,99	1
7	MN	I	104	1/1	0.75	0.19	99,99,99,99	1
7	MN	I	102	1/1	0.84	0.12	107,107,107,107	0
7	MN	J	105	1/1	0.89	0.07	106,106,106,106	0
7	MN	I	105	1/1	0.89	0.12	76,76,76,76	1
7	MN	I	107	1/1	0.91	0.09	99,99,99,99	0
7	MN	J	104	1/1	0.91	0.10	104,104,104,104	0
7	MN	I	106	1/1	0.93	0.16	71,71,71,71	1
7	MN	I	101	1/1	0.96	0.14	74,74,74,74	1
7	MN	J	101	1/1	0.96	0.14	82,82,82,82	0
7	MN	J	102	1/1	0.96	0.04	78,78,78,78	0
7	MN	J	103	1/1	0.97	0.08	77,77,77,77	0
7	MN	I	103	1/1	0.98	0.08	76,76,76,76	0
7	MN	E	201	1/1	1.00	0.03	40,40,40,40	0

6.5 Other polymers [i](#)

There are no such residues in this entry.