



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 31, 2020 – 08:25 AM BST

PDB ID : 1AOI
Title : COMPLEX BETWEEN NUCLEOSOME CORE PARTICLE
(H3,H4,H2A,H2B) AND 146 BP LONG DNA FRAGMENT
Authors : Luger, K.; Maeder, A.W.; Richmond, R.K.; Sargent, D.F.; Richmond, T.J.
Deposited on : 1997-07-03
Resolution : 2.80 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.13

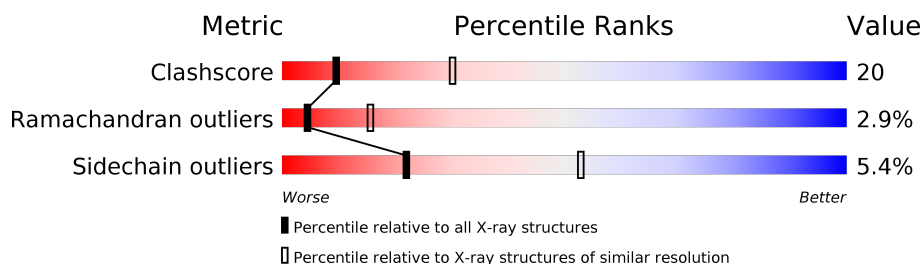
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)

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 on 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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	I	146	<div> <div>27%</div> <div>72%</div> <div>.</div> </div>
1	J	146	<div> <div>29%</div> <div>68%</div> <div>.</div> </div>
2	A	116	<div> <div>61%</div> <div>22%</div> <div>.</div> <div>16%</div> </div>
2	E	116	<div> <div>79%</div> <div>16%</div> <div>..</div> </div>
3	B	87	<div> <div>78%</div> <div>15%</div> <div>..</div> <div>5%</div> </div>
3	F	87	<div> <div>72%</div> <div>24%</div> <div>.</div> </div>
4	C	116	<div> <div>73%</div> <div>22%</div> <div>..</div> </div>
4	G	116	<div> <div>68%</div> <div>22%</div> <div>.</div> <div>7%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	D	99	<div><div></div><div>76%</div><div>21%</div><div></div></div>
5	H	99	<div><div></div><div>68%</div><div>24%</div><div>8%</div></div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12385 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 DNA chain called PALINDROMIC 146 BP DNA REPEAT 8/9 FROM HUMAN X-CHROMOSOME ALPHA SATELLITE DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	146	Total	C	N	O	P	0	0	0
			2990	1431	540	874	145			
1	J	146	Total	C	N	O	P	0	0	0
			2990	1431	540	874	145			

- Molecule 2 is a protein called HISTONE H3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	98	Total	C	N	O	S	0	0	0
			808	509	156	140	3			
2	E	116	Total	C	N	O	S	0	0	0
			930	585	181	161	3			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	ALA	VAL	CONFLICT	UNP P02302
A	26	ARG	LYS	CONFLICT	UNP P02302
A	28	SER	CYS	CONFLICT	UNP P02302
A	86	SER	ARG	CONFLICT	UNP P02302
E	21	ALA	VAL	CONFLICT	UNP P02302
E	26	ARG	LYS	CONFLICT	UNP P02302
E	28	SER	CYS	CONFLICT	UNP P02302
E	86	SER	ARG	CONFLICT	UNP P02302

- Molecule 3 is a protein called HISTONE H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	83	Total	C	N	O	S	0	0	0
			662	418	129	114	1			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	87	Total	C	N	O	S	0	0	0
			703	442	142	118	1			

- Molecule 4 is a protein called HISTONE H2A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	C	115	Total	C	N	O		0	0	0
			880	552	175	153				
4	G	108	Total	C	N	O		0	0	0
			833	525	163	145				

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	99	ARG	GLY	CONFLICT	UNP P06897
G	99	ARG	GLY	CONFLICT	UNP P06897

- Molecule 5 is a protein called HISTONE H2B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	D	99	Total	C	N	O	S	0	0	0
			785	493	146	144	2			
5	H	99	Total	C	N	O	S	0	0	0
			785	493	146	144	2			

There are 2 discrepancies between the modelled and reference sequences:

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

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	3	Total	Mn	0	0
			3	3		
6	I	3	Total	Mn	0	0
			3	3		

- Molecule 7 is water.

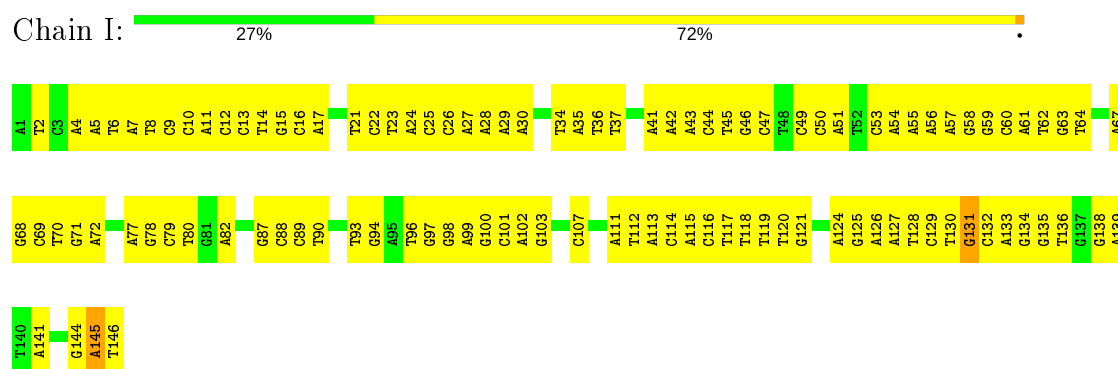
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	I	4	Total 4	O 4	0	0
7	J	5	Total 5	O 5	0	0
7	B	1	Total 1	O 1	0	0
7	E	1	Total 1	O 1	0	0
7	F	2	Total 2	O 2	0	0

3 Residue-property plots [i](#)

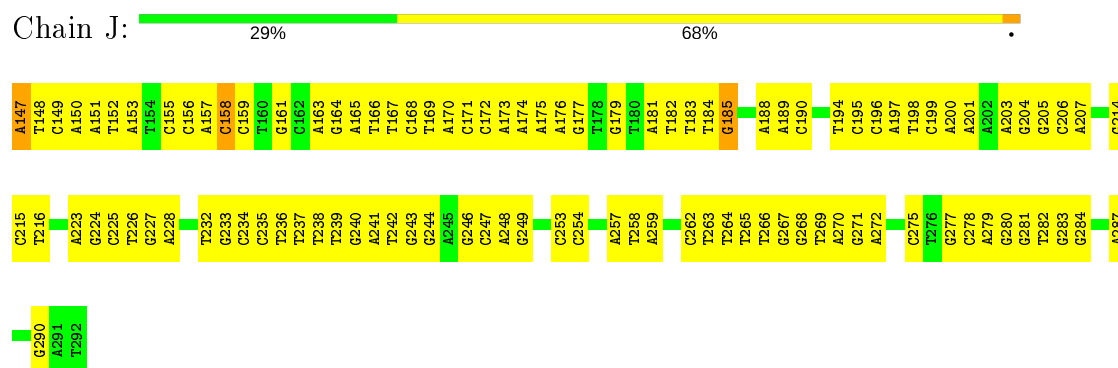
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

Note EDS was not executed.

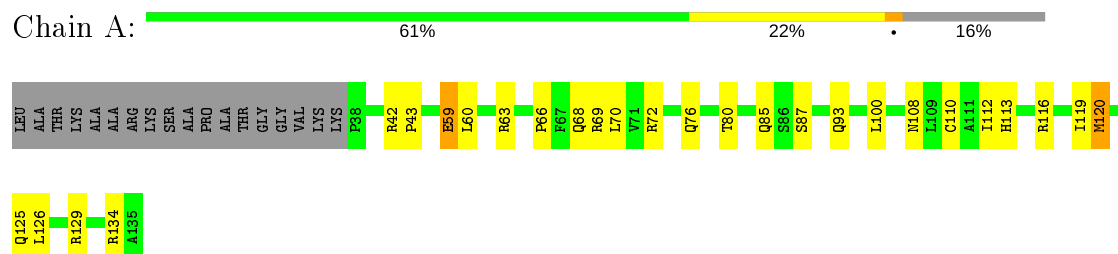
- Molecule 1: PALINDROMIC 146 BP DNA REPEAT 8/9 FROM HUMAN X-CHROMOSOME ALPHA SATELLITE DNA




- Molecule 1: PALINDROMIC 146 BP DNA REPEAT 8/9 FROM HUMAN X-CHROMOSOME ALPHA SATELLITE DNA

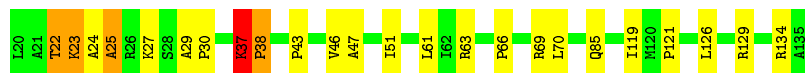


- Molecule 2: HISTONE H3




- Molecule 2: HISTONE H3

Chain E:  79% 16% ..



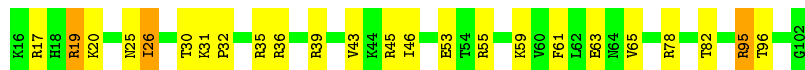
- Molecule 3: HISTONE H4

Chain B:  78% 15% .. 5%



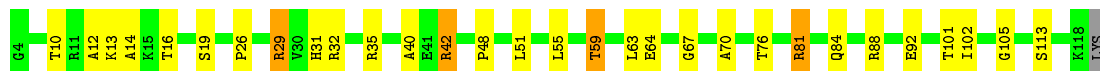
- Molecule 3: HISTONE H4

Chain F:  72% 24% .



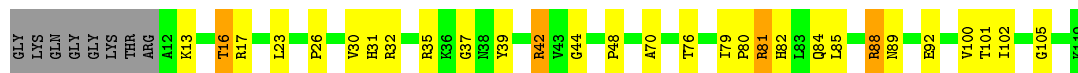
- Molecule 4: HISTONE H2A

Chain C:  73% 22% ..



- Molecule 4: HISTONE H2A

Chain G:  68% 22% . 7%



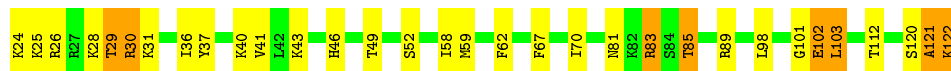
- Molecule 5: HISTONE H2B

Chain D:  76% 21% .



- Molecule 5: HISTONE H2B

Chain H:  68% 24% 8%



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	106.04Å 181.78Å 110.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.80	Depositor
% Data completeness (in resolution range)	(Not available) (25.00-2.80)	Depositor
R_{merge}	0.06	Depositor
R_{sym}	3.60	Depositor
Refinement program	X-PLOR 3.843	Depositor
R, R_{free}	0.224 , 0.302	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12385	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

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	I	0.54	0/3354	0.77	0/5175
1	J	0.53	0/3354	0.79	0/5175
2	A	0.54	0/820	0.73	0/1099
2	E	0.59	0/943	0.79	0/1264
3	B	0.52	0/669	0.75	0/894
3	F	0.59	0/711	0.81	0/948
4	C	0.53	0/890	0.76	0/1197
4	G	0.49	0/843	0.75	0/1135
5	D	0.52	0/796	0.70	0/1065
5	H	0.53	0/796	0.71	0/1065
All	All	0.54	0/13176	0.77	0/19017

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	I	0	3
1	J	0	4
All	All	0	7

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	131	DG	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	I	145	DA	Sidechain
1	I	96	DT	Sidechain
1	J	147	DA	Sidechain
1	J	158	DC	Sidechain

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	I	2990	0	1652	162	0
1	J	2990	0	1652	170	0
2	A	808	0	846	22	0
2	E	930	0	987	28	0
3	B	662	0	709	11	0
3	F	703	0	755	18	0
4	C	880	0	945	26	0
4	G	833	0	895	31	0
5	D	785	0	825	21	1
5	H	785	0	825	24	0
6	I	3	0	0	0	0
6	J	3	0	0	0	0
7	B	1	0	0	0	0
7	E	1	0	0	0	1
7	F	2	0	0	0	0
7	I	4	0	0	0	0
7	J	5	0	0	1	0
All	All	12385	0	10091	434	1

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 20.

The worst 5 of 434 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:J:157:DA:H2''	1:J:158:DC:H5''	1.21	1.13
1:J:238:DT:H2''	1:J:239:DT:C5'	1.80	1.11

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:J:194:DT:H2''	1:J:195:DC:H5'	1.33	1.10
1:I:99:DA:H2''	1:I:100:DG:H5'	1.10	1.08
1:I:61:DA:H2''	1:I:62:DT:H5'	1.33	1.07

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:D:45:VAL:O	7:E:6:HOH:O[2_565]	2.17	0.03

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	
2	A	96/116 (83%)	92 (96%)	4 (4%)	0	100	100
2	E	114/116 (98%)	103 (90%)	4 (4%)	7 (6%)	1	4
3	B	81/87 (93%)	77 (95%)	3 (4%)	1 (1%)	13	39
3	F	85/87 (98%)	81 (95%)	1 (1%)	3 (4%)	3	12
4	C	113/116 (97%)	100 (88%)	9 (8%)	4 (4%)	3	12
4	G	106/116 (91%)	100 (94%)	5 (5%)	1 (1%)	17	46
5	D	97/99 (98%)	91 (94%)	3 (3%)	3 (3%)	4	14
5	H	97/99 (98%)	86 (89%)	7 (7%)	4 (4%)	3	9
All	All	789/836 (94%)	730 (92%)	36 (5%)	23 (3%)	4	15

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	21	VAL
4	C	14	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	D	101	GLY
2	E	22	THR
2	E	23	LYS

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	
2	A	85/96 (88%)	82 (96%)	3 (4%)	36	70
2	E	96/96 (100%)	94 (98%)	2 (2%)	53	84
3	B	68/72 (94%)	64 (94%)	4 (6%)	19	49
3	F	72/72 (100%)	68 (94%)	4 (6%)	21	51
4	C	89/90 (99%)	82 (92%)	7 (8%)	12	34
4	G	85/90 (94%)	80 (94%)	5 (6%)	19	49
5	D	85/85 (100%)	82 (96%)	3 (4%)	36	70
5	H	85/85 (100%)	77 (91%)	8 (9%)	8	26
All	All	665/686 (97%)	629 (95%)	36 (5%)	22	53

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	D	83	ARG
3	F	20	LYS
5	H	103	LEU
2	E	129	ARG
3	F	26	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
4	C	84	GLN
4	G	31	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	D	92	GLN
3	B	75	HIS
5	D	46	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 6 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 [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.