wwPDB X-ray Structure Validation Summary Report (i)

Aug 7, 2023 – 01:10 AM EDT

PDB ID	:	1KX3
Title	:	X-Ray Structure of the Nucleosome Core Particle, NCP146, at 2.0 A Resolution
Authors	:	Davey, C.A.; Sargent, D.F.; Luger, K.; Maeder, A.W.; Richmond, T.J.
Deposited on	:	2002-01-31
Resolution	:	2.00 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 (i) 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 (i)) were used in the production of this report:

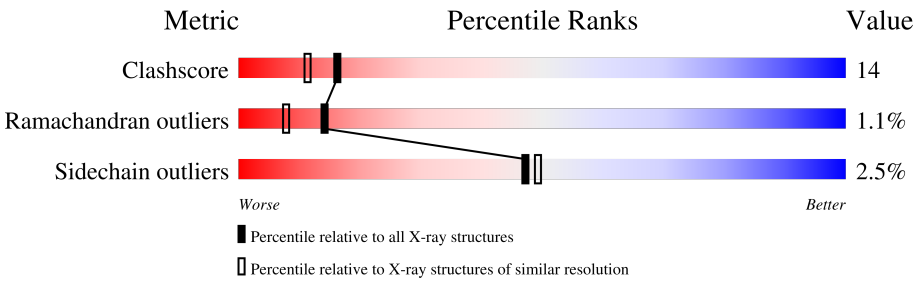
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	I	146	50% 50%
1	J	146	47% 51% •
2	A	135	57% 14% • 27%
2	E	135	55% 16% • 27%
3	B	102	61% 18% •• 20%
3	F	102	66% 17% • 15%
4	C	128	66% 15% •• 16%
4	G	128	66% 16% • 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	D	125	58% 15% • 25%
5	H	125	63% 11% • 25%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 13023 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 DNA (5'(ATCAATATCCACCTGCAGATTCTACCAAAAGTGTATTTGGAAACTGCTCCATCAAAAGGCATGTTCAGCTGAATTCAGCTGAACATGCCTTTTGATGGAGCAGTTTCCAAATACACTTTTGGTAGAATCTGCAGGTGGATATTGAT)3').

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

- Molecule 2 is a protein called histone H3.

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

There are 2 discrepancies between the modelled and reference sequences:

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

- Molecule 3 is a protein called histone H4.

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

- Molecule 4 is a protein called histone H2A.1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	C	107	825	520	161	144	0	0	0
4	G	106	818	516	160	142	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	99	ARG	GLY	variant	UNP P06897
C	123	SER	ALA	conflict	UNP P06897
C	?	-	ALA	deletion	UNP P06897
G	99	ARG	GLY	variant	UNP P06897
G	123	SER	ALA	conflict	UNP P06897
G	?	-	ALA	deletion	UNP P06897

- Molecule 5 is a protein called histone H2B.2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	D	94	736	463	132	139	2	0	0	0
5	H	94	736	463	132	139	2	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

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

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

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

- Molecule 7 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O			
7	I	141	141	141		0	0
			Total	O			
7	J	133	133	133		0	0
			Total	O			
7	A	74	74	74		0	0
			Total	O			
7	B	62	62	62		0	0
			Total	O			
7	C	107	107	107		0	0
			Total	O			
7	D	77	77	77		0	0
			Total	O			
7	E	116	116	116		0	0
			Total	O			
7	F	90	90	90		0	0
			Total	O			
7	G	98	98	98		0	0
			Total	O			
7	H	45	45	45		0	0

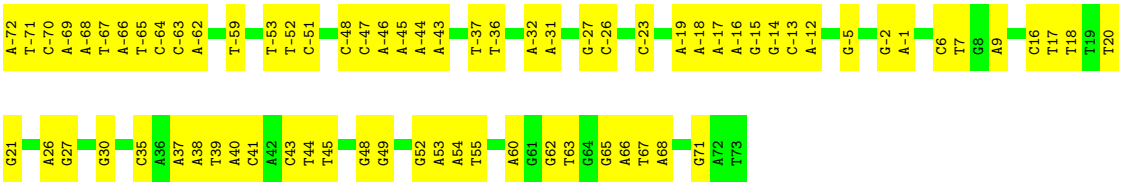
3 Residue-property plots (i)

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.

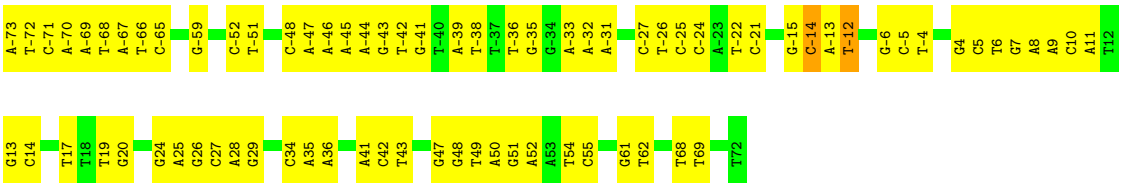
• Molecule 1: DNA (5'(ATCAATATCCACCTGCAGATTCTACCAAAAGTGTATTTGGAAACTGCTCCATCAAAAGGCATGTTCAGCTGAATTCAGCTGAACATGCCTTTTGATGGAGCAGTTTCCAAATACACTTTTGGTAGAATCTGCAGGTGGATATTGAT)3')

Chain I: 50% 50%



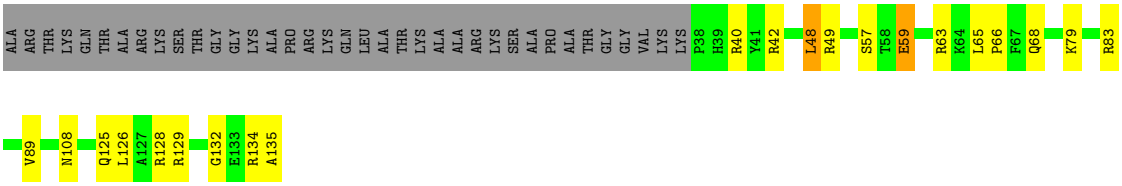
• Molecule 1: DNA (5'(ATCAATATCCACCTGCAGATTCTACCAAAAGTGTATTTGGAAACTGCTCCATCAAAAGGCATGTTCAGCTGAATTCAGCTGAACATGCCTTTTGATGGAGCAGTTTCCAAATACACTTTTGGTAGAATCTGCAGGTGGATATTGAT)3')

Chain J: 47% 51%



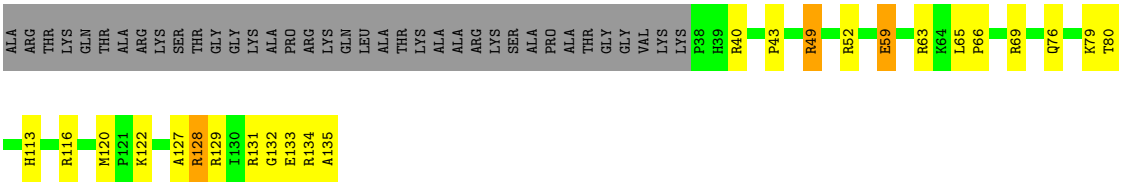
• Molecule 2: histone H3

Chain A: 57% 14% • 27%



• Molecule 2: histone H3

Chain E: 55% 16% • 27%

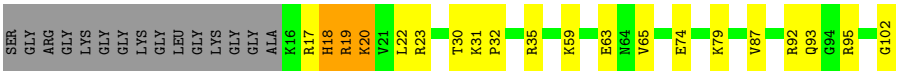


- Molecule 3: histone H4

Chain B:

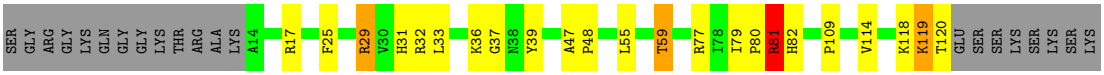
- Molecule 3: histone H4

Chain F:



- Molecule 4: histone H2A.1

Chain C:

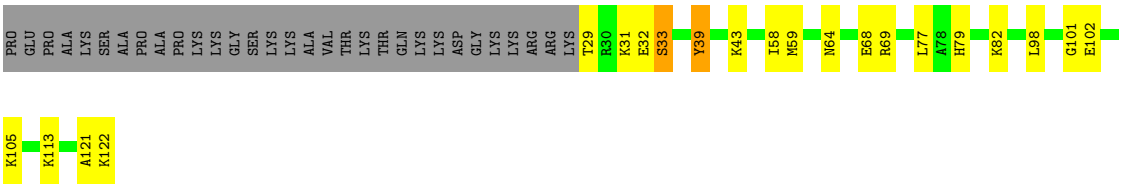


- Molecule 4: histone H2A.1

Chain G:

- Molecule 5: histone H2B.2

Chain D:



- Molecule 5: histone H2B.2

Chain H:

PRO
GLU
PRO
ALA
LYS
SER
ALA
PRO
ALA
PRO
LYS
LYS
GLY
SER
LYS
LYS
ALA
VAL
THR
LYS
THR
GLN
LYS
LYS
ASP
GLY
LYS
LYS
ARG
ARG
LYS
T29
R30
K31
E32
S33
I36
K43
Q44
V45
I58
R69
E73
G101
E102
L103
H106
K122

4 Data and refinement statistics (i)

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	105.40Å 181.54Å 109.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.00	Depositor
% Data completeness (in resolution range)	99.9 (6.00-2.00)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.240 , 0.275	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	13023	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

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	0.35	0/3354	0.70	0/5175
1	J	0.39	0/3354	0.71	1/5175 (0.0%)
2	A	0.52	0/820	0.67	0/1099
2	E	0.69	0/820	0.80	1/1099 (0.1%)
3	B	0.56	0/660	0.72	1/883 (0.1%)
3	F	0.69	0/711	0.83	1/948 (0.1%)
4	C	0.67	0/835	0.83	2/1127 (0.2%)
4	G	0.51	0/828	0.69	2/1117 (0.2%)
5	D	0.66	0/747	0.73	0/1004
5	H	0.55	0/747	0.67	0/1004
All	All	0.50	0/12876	0.72	8/18631 (0.0%)

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	J	0	2
5	D	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	81	ARG	NE-CZ-NH1	9.46	125.03	120.30
4	C	81	ARG	NE-CZ-NH2	-8.80	115.90	120.30
4	G	88	ARG	NE-CZ-NH1	6.80	123.70	120.30
2	E	128	ARG	NE-CZ-NH2	-6.79	116.90	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	29	ILE	N-CA-C	-5.36	96.54	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
5	D	39	TYR	Sidechain
1	J	-12	DT	Sidechain
1	J	-6	DG	Sidechain

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	I	2990	0	1652	75	0
1	J	2990	0	1652	92	0
2	A	808	0	846	40	0
2	E	808	0	846	41	0
3	B	653	0	696	16	0
3	F	703	0	755	25	0
4	C	825	0	884	22	0
4	G	818	0	877	25	0
5	D	736	0	760	22	0
5	H	736	0	760	15	0
6	E	1	0	0	0	0
6	I	6	0	0	0	0
6	J	6	0	0	0	0
7	A	74	0	0	3	0
7	B	62	0	0	1	0
7	C	107	0	0	3	0
7	D	77	0	0	4	0
7	E	116	0	0	6	0
7	F	90	0	0	3	0
7	G	98	0	0	2	0
7	H	45	0	0	3	0
7	I	141	0	0	8	0
7	J	133	0	0	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	13023	0	9728	312	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 14.

The worst 5 of 312 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:-46:DA:H2"	1:J:-45:DA:H5"	1.27	1.17
2:E:120:MET:HE3	2:E:122:LYS:HD3	1.39	1.03
1:I:26:DA:H2"	1:I:27:DG:H5'	1.41	1.02
2:A:128:ARG:HD3	2:A:134:ARG:HH12	1.22	1.02
2:E:49:ARG:HG3	2:E:49:ARG:HH11	1.27	0.99

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 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	96/135 (71%)	96 (100%)	0	0	100 100
2	E	96/135 (71%)	95 (99%)	1 (1%)	0	100 100
3	B	80/102 (78%)	78 (98%)	0	2 (2%)	5 2
3	F	85/102 (83%)	81 (95%)	1 (1%)	3 (4%)	3 1
4	C	105/128 (82%)	101 (96%)	4 (4%)	0	100 100
4	G	104/128 (81%)	102 (98%)	2 (2%)	0	100 100
5	D	92/125 (74%)	90 (98%)	1 (1%)	1 (1%)	14 8
5	H	92/125 (74%)	88 (96%)	2 (2%)	2 (2%)	6 2
All	All	750/980 (76%)	731 (98%)	11 (2%)	8 (1%)	14 8

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	F	19	ARG
3	B	22	LEU
5	D	101	GLY
5	H	101	GLY
3	F	20	LYS

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 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	85/110 (77%)	83 (98%)	2 (2%)	49	51
2	E	85/110 (77%)	83 (98%)	2 (2%)	49	51
3	B	67/78 (86%)	66 (98%)	1 (2%)	65	69
3	F	72/78 (92%)	71 (99%)	1 (1%)	67	72
4	C	85/101 (84%)	80 (94%)	5 (6%)	19	15
4	G	84/101 (83%)	83 (99%)	1 (1%)	71	76
5	D	80/105 (76%)	79 (99%)	1 (1%)	69	74
5	H	80/105 (76%)	77 (96%)	3 (4%)	33	31
All	All	638/788 (81%)	622 (98%)	16 (2%)	47	49

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	H	33	SER
5	H	31	LYS
5	D	33	SER
4	G	88	ARG
4	C	119	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	18	HIS
4	G	31	HIS
5	H	106	HIS
5	H	79	HIS
4	C	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 13 ligands modelled in this entry, 13 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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.