wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 02:31 AM GMT

PDB ID	:	2HIO
Title	:	HISTONE OCTAMER (CHICKEN), CHROMOSOMAL PROTEIN
Authors	:	Arents, G.; Moudrianakis, E.N.
Deposited on	:	1999-06-15
Resolution	:	3.10 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

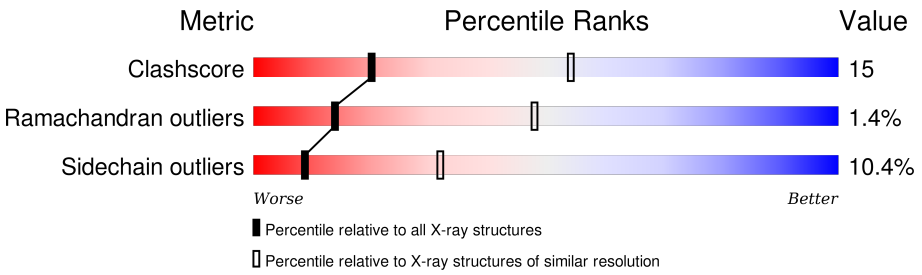
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 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)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)

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. 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% 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	128	48% 24% 5% • 21%
2	B	125	48% 21% • 28%
3	C	136	45% 19% • 32%
4	D	103	49% 24% • 24%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2850 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 PROTEIN (HISTONE H2A).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	101	777	490	151	136	0	0	0

- Molecule 2 is a protein called PROTEIN (HISTONE H2B).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	90	700	442	124	132	2	0	0	0

- Molecule 3 is a protein called PROTEIN (HISTONE H3).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	93	755	476	143	133	3	0	0	0

- Molecule 4 is a protein called PROTEIN (HISTONE H4).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	78	618	391	120	106	1	0	0	0

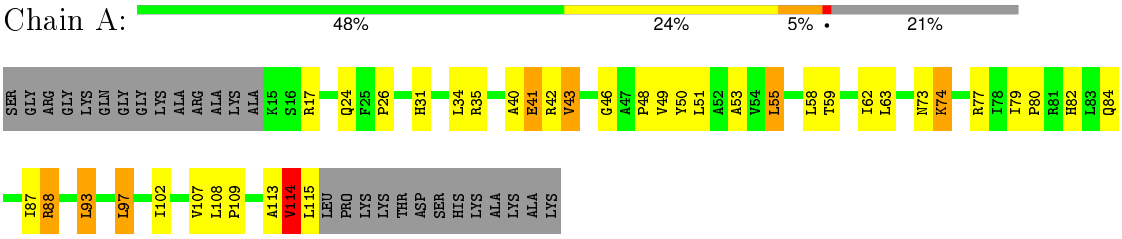
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.

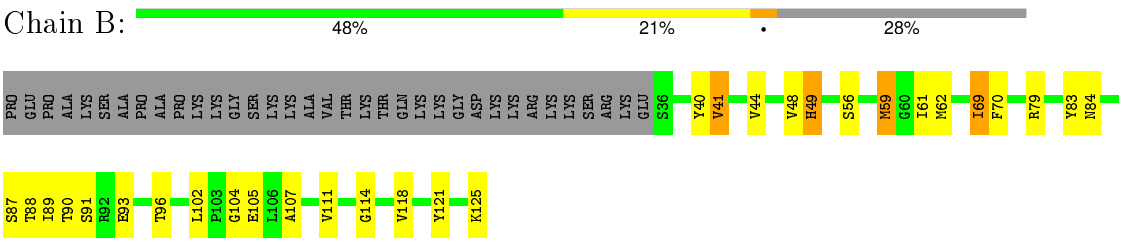
- Molecule 1: PROTEIN (HISTONE H2A)

Chain A:



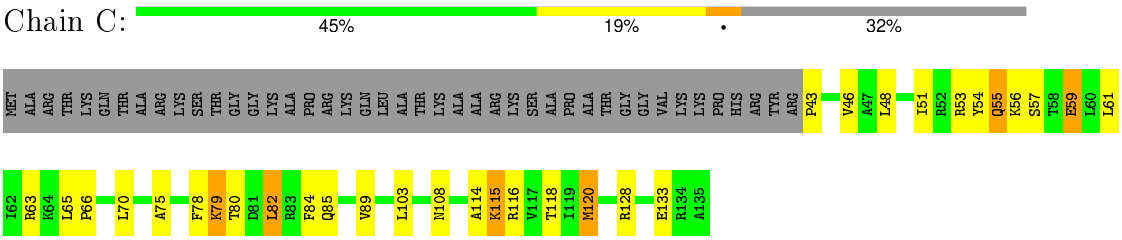
- Molecule 2: PROTEIN (HISTONE H2B)

Chain B:



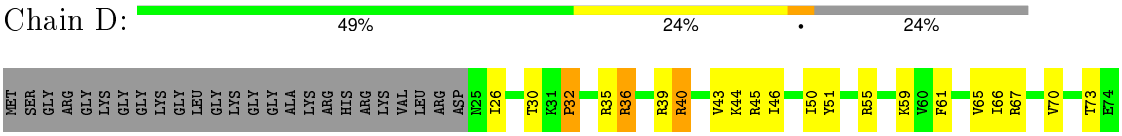
- Molecule 3: PROTEIN (HISTONE H3)

Chain C:



- Molecule 4: PROTEIN (HISTONE H4)

Chain D:



H75
A76
K77
R78
M84
L90
R95
G101
G102

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	118.82Å 118.82Å 102.93Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	10.00 – 3.10	Depositor
% Data completeness (in resolution range)	80.0 (10.00-3.10)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROFFT	Depositor
R, R_{free}	0.255 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2850	wwPDB-VP
Average B, all atoms ($Å^2$)	47.0	wwPDB-VP

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| >5	RMSZ	#|Z| >5
1	A	0.44	0/786	0.70	0/1060
2	B	0.46	0/711	0.69	0/957
3	C	0.47	0/764	0.68	0/1024
4	D	0.47	0/625	0.72	1/837 (0.1%)
All	All	0.46	0/2886	0.70	1/3878 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	32	PRO	CA-N-CD	-5.21	104.21	111.50

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	777	0	828	32	0
2	B	700	0	721	31	0
3	C	755	0	795	24	0
4	D	618	0	657	23	0
All	All	2850	0	3001	86	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 15.

The worst 5 of 86 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)
3:C:115:LYS:CD	3:C:115:LYS:O	1.82	1.27
3:C:115:LYS:HD2	3:C:115:LYS:O	1.38	1.15
3:C:115:LYS:HD3	3:C:115:LYS:O	1.62	0.95
2:B:84:ASN:HD21	4:D:76:ALA:HB2	1.31	0.94
3:C:79:LYS:HB3	3:C:82:LEU:HD22	1.53	0.91

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
1	A	99/128 (77%)	85 (86%)	11 (11%)	3 (3%)	5 28
2	B	88/125 (70%)	80 (91%)	7 (8%)	1 (1%)	17 55
3	C	91/136 (67%)	85 (93%)	5 (6%)	1 (1%)	17 55
4	D	76/103 (74%)	70 (92%)	6 (8%)	0	100 100
All	All	354/492 (72%)	320 (90%)	29 (8%)	5 (1%)	14 48

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	114	VAL
2	B	104	GLY
1	A	74	LYS
1	A	73	ASN
3	C	114	ALA

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
1	A	79/97 (81%)	67 (85%)	12 (15%)	3 14
2	B	77/106 (73%)	71 (92%)	6 (8%)	16 49
3	C	80/111 (72%)	69 (86%)	11 (14%)	4 19
4	D	63/80 (79%)	61 (97%)	2 (3%)	46 79
All	All	299/394 (76%)	268 (90%)	31 (10%)	9 32

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	56	SER
2	B	105	GLU
3	C	120	MET
2	B	59	MET
3	C	53	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	104	GLN
1	A	110	ASN
3	C	93	GLN
1	A	94	ASN
3	C	76	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)

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 will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.