wwPDB X-ray Structure Validation Summary Report (i)

May 27, 2020 – 10:39 pm BST

PDB ID	:	3AAD
Title	:	Structure of the histone chaperone CIA/ASF1-double bromodomain complex linking histone modifications and site-specific histone eviction
Authors	:	Akai, Y.; Adachi, N.; Hayashi, Y.; Eitoku, M.; Sano, N.; Natsume, R.; Kudo, N.; Tanokura, M.; Senda, T.; Horikoshi, M.
Deposited on	:	2009-11-16
Resolution	:	3.30 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 following versions of software and data (see references (1)) were used in the production of this report:

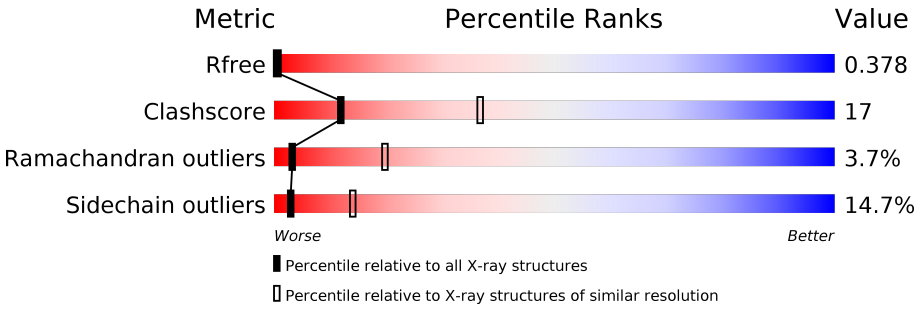
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.30 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)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	292	53% 32% • 10%
2	B	158	44% 41% 10% • •
2	D	158	57% 34% 5% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SO4	A	400	-	-	X	-

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4617 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 Transcription initiation factor TFIID subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	264	Total	C	N	O	S	0	0	0
			2164	1369	367	417	11			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1338	GLY	-	EXPRESSION TAG	UNP P21675
A	1339	SER	-	EXPRESSION TAG	UNP P21675
A	1340	HIS	-	EXPRESSION TAG	UNP P21675
A	1341	MET	-	EXPRESSION TAG	UNP P21675

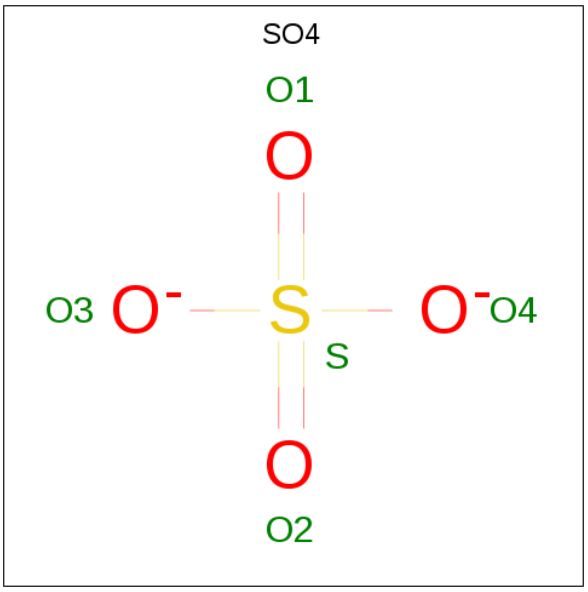
- Molecule 2 is a protein called Histone chaperone ASF1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	153	Total	C	N	O	S	0	0	0
			1231	791	201	235	4			
2	D	152	Total	C	N	O	S	0	0	0
			1217	780	199	234	4			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLY	-	EXPRESSION TAG	UNP Q9Y294
B	-1	SER	-	EXPRESSION TAG	UNP Q9Y294
B	0	HIS	-	EXPRESSION TAG	UNP Q9Y294
D	-2	GLY	-	EXPRESSION TAG	UNP Q9Y294
D	-1	SER	-	EXPRESSION TAG	UNP Q9Y294
D	0	HIS	-	EXPRESSION TAG	UNP Q9Y294

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



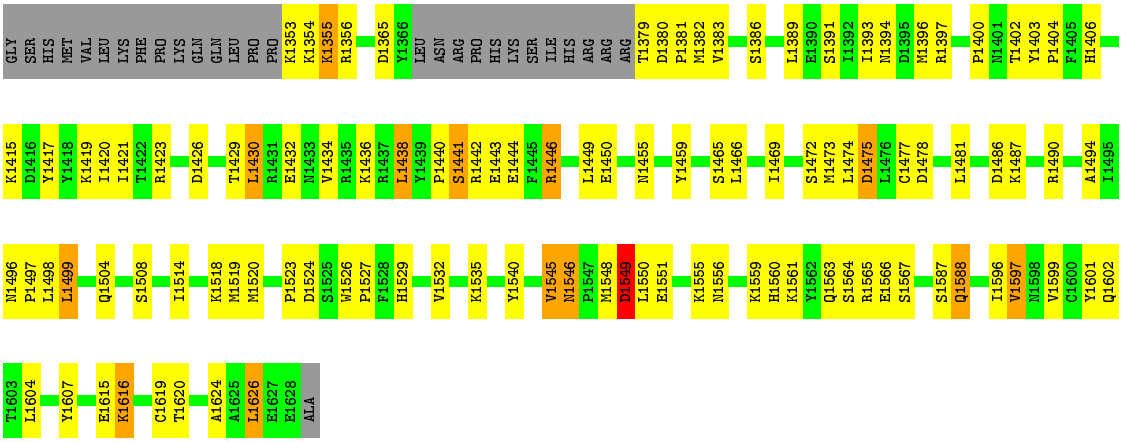
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0

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 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.

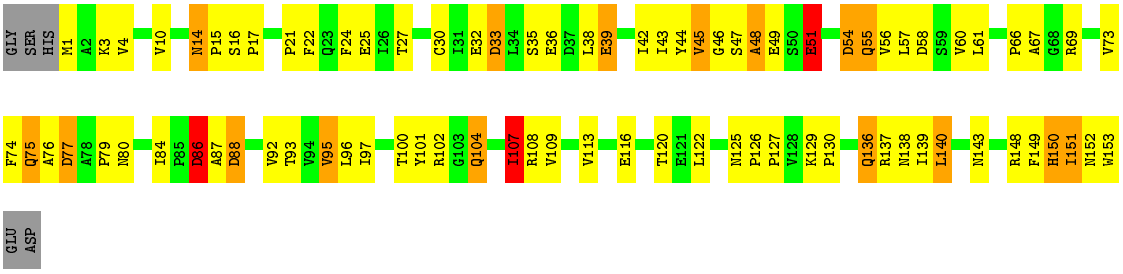
- Molecule 1: Transcription initiation factor TFIID subunit 1

Chain A:



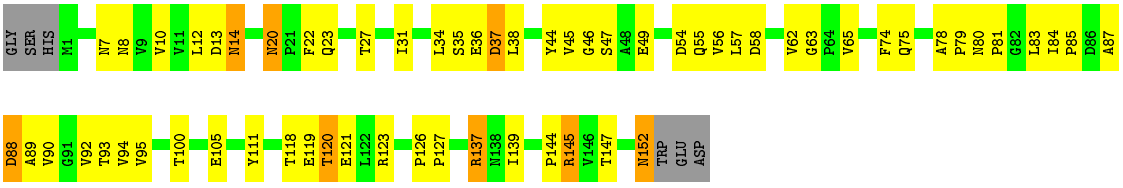
- Molecule 2: Histone chaperone ASF1A

Chain B:



- Molecule 2: Histone chaperone ASF1A

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	102.12Å 102.12Å 271.92Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	31.65 – 3.30 31.65 – 3.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (31.65-3.30) 97.4 (31.65-3.30)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	14.12 (at 3.31Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.237 , 0.293 0.328 , 0.378	Depositor DCC
R_{free} test set	647 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	85.6	Xtriage
Anisotropy	0.088	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 4.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	4617	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.40% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: SO4

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	0.62	0/2206	0.71	0/2985
2	B	0.98	7/1264 (0.6%)	0.93	4/1729 (0.2%)
2	D	0.58	1/1248 (0.1%)	0.67	2/1706 (0.1%)
All	All	0.72	8/4718 (0.2%)	0.76	6/6420 (0.1%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	150	HIS	CE1-NE2	11.51	1.59	1.32
2	B	150	HIS	CG-ND1	9.68	1.60	1.38
2	B	150	HIS	CG-CD2	9.50	1.51	1.35
2	B	51	GLU	CD-OE2	7.49	1.33	1.25
2	B	51	GLU	CD-OE1	6.17	1.32	1.25

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	137	ARG	NE-CZ-NH1	6.71	123.66	120.30
2	B	137	ARG	NE-CZ-NH2	-6.45	117.07	120.30
2	B	148	ARG	NE-CZ-NH1	5.97	123.29	120.30
2	D	145	ARG	NE-CZ-NH2	-5.54	117.53	120.30
2	B	150	HIS	CG-CD2-NE2	-5.53	98.69	109.20

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	2164	0	2165	65	0
2	B	1231	0	1188	55	0
2	D	1217	0	1178	40	0
3	A	5	0	0	2	0
All	All	4617	0	4531	158	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 17.

The worst 5 of 158 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)
2:B:104:GLN:HG2	2:B:151:ILE:CG1	1.76	1.16
2:B:104:GLN:HG2	2:B:151:ILE:HG12	1.07	1.06
2:B:104:GLN:CG	2:B:151:ILE:HG12	1.93	0.98
2:B:95:VAL:O	2:B:96:LEU:HD23	1.69	0.91
2:B:47:SER:O	2:B:49:GLU:N	2.08	0.86

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	260/292 (89%)	218 (84%)	32 (12%)	10 (4%)	3 19
2	B	151/158 (96%)	124 (82%)	19 (13%)	8 (5%)	2 12

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	D	150/158 (95%)	125 (83%)	22 (15%)	3 (2%)	7 32
All	All	561/608 (92%)	467 (83%)	73 (13%)	21 (4%)	3 20

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1523	PRO
2	B	33	ASP
2	B	36	GLU
2	B	48	ALA
2	B	86	ASP

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	250/276 (91%)	222 (89%)	28 (11%)	6 23
2	B	138/142 (97%)	106 (77%)	32 (23%)	1 3
2	D	137/142 (96%)	120 (88%)	17 (12%)	4 19
All	All	525/560 (94%)	448 (85%)	77 (15%)	3 14

5 of 77 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	51	GLU
2	B	77	ASP
2	D	90	VAL
2	B	54	ASP
2	B	60	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1588	GLN
2	B	55	GLN
2	D	20	ASN
1	A	1602	GLN
2	B	20	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	SO4	A	400	-	4,4,4	0.48	0	6,6,6	0.28	0

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.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	400	SO4	2	0

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.