



wwPDB X-ray Structure Validation Summary Report ⓘ

May 10, 2018 – 04:44 PM EDT

PDB ID : 5KDM
Title : Crystal structure of EBV tegument protein BNRF1 in complex with histone chaperone DAXX and histones H3.3-H4
Authors : Huang, H.; Patel, D.
Deposited on : 2016-06-08
Resolution : 3.50 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)	:	1.13
EDS	:	rb-20031172
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031172

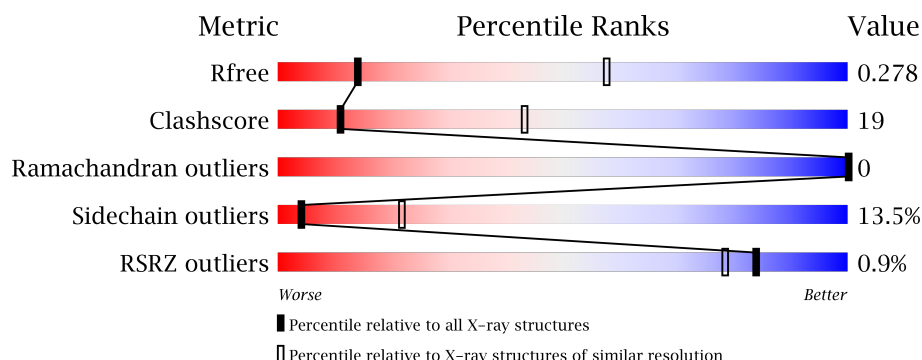
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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}	111664	1391 (3.60-3.40)
Clashscore	122126	1485 (3.60-3.40)
Ramachandran outliers	120053	1446 (3.60-3.40)
Sidechain outliers	120020	1447 (3.60-3.40)
RSRZ outliers	108989	1303 (3.60-3.40)

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 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 $\leq 5\%$. 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	135	
2	B	102	
3	C	212	
4	D	219	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4341 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 H3.3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	99	Total	C	N	O	S	0	0	0
			785	495	152	136	2			

- Molecule 2 is a protein called Histone H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	77	Total	C	N	O	S	0	0	0
			599	379	116	103	1			

- Molecule 3 is a protein called Death domain-associated protein 6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	205	Total	C	N	O	S	0	0	0
			1636	1022	303	305	6			

- Molecule 4 is a protein called Major tegument protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	180	Total	C	N	O	S	0	0	0
			1321	851	217	244	9			

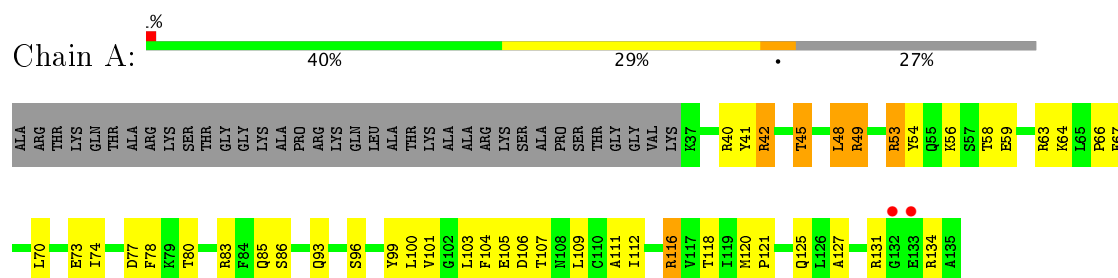
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	587	SER	ARG	conflict	UNP Q1HJVJ0

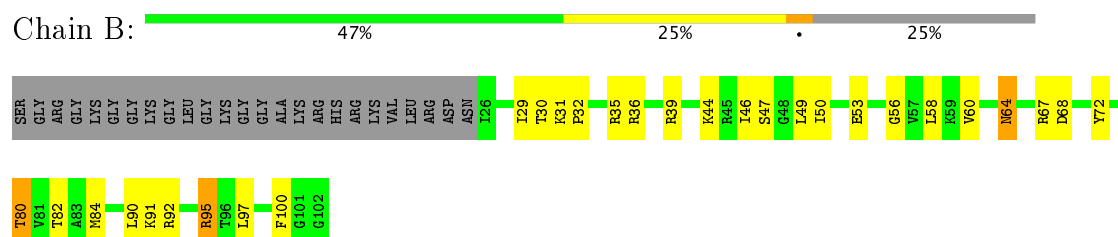
3 Residue-property plots

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 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.

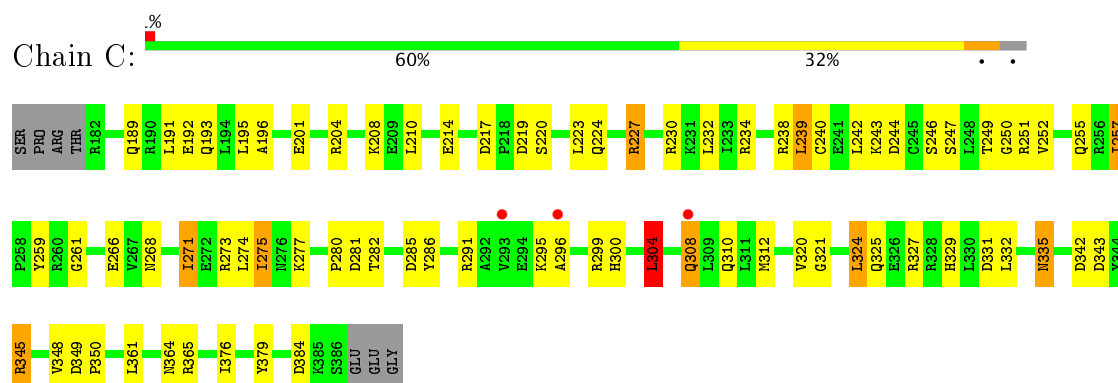
• Molecule 1: Histone H3.3



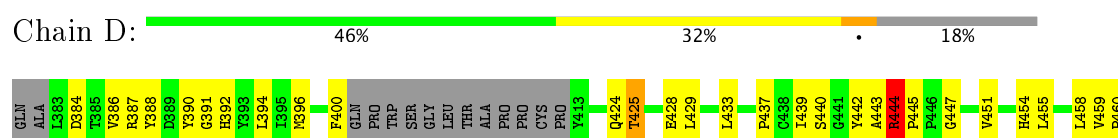
• Molecule 2: Histone H4



• Molecule 3: Death domain-associated protein 6



• Molecule 4: Major tegument protein



K461	L464	L465	L466	S469	H470	L471	P472	D473	D474	Y475	L479	G480	GLU	MET	GLY	PRO	ALA	ARG	ALA	T488	G489	P490	G491	N492	Q493	S497	S498	T499	F500	L501	N502	P503	A504	C505	S506	N507	N508	F509	L510	T511	V512	R513	N520	T523	Q526	A527	L528	A529	R530	A531	C532	D533
M534	A535	G543	S544	T545	V546	P547	L551	V554	N555	D556	V561	S562	E565	M566	M567	D568	D569	F570	F573	F574	I581	GLN	GLU	GLU	GLY	ALA	SER	SER	PRO	VAL	PRO	LEU	ASP	VAL	ASP	F509	GLU	SER	MET	ASP	N520	T523	Q526	A527	L528	A529	R530	A531	C532	D533		

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	161.22Å 161.22Å 117.77Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.16 – 3.50 48.16 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.2 (48.16-3.50) 87.0 (48.16-3.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.64 (at 3.48Å)	Xtriage
Refinement program	PHENIX dev_1839	Depositor
R, R_{free}	0.234 , 0.280 0.234 , 0.278	Depositor DCC
R_{free} test set	1135 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	114.8	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 68.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.045 for -h,-k,l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	4341	wwPDB-VP
Average B, all atoms (Å ²)	130.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% 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 [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	0.57	0/796	0.71	0/1072
2	B	0.52	0/606	0.78	0/813
3	C	0.56	0/1660	0.78	2/2239 (0.1%)
4	D	0.66	0/1356	0.85	0/1851
All	All	0.59	0/4418	0.79	2/5975 (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
3	C	0	2
4	D	0	1
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	304	LEU	CA-CB-CG	5.48	127.91	115.30
3	C	239	LEU	CB-CG-CD2	-5.07	102.39	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	246	SER	Peptide
3	C	273	ARG	Peptide
4	D	444	ARG	Peptide

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	785	0	793	43	0
2	B	599	0	627	31	0
3	C	1636	0	1637	69	0
4	D	1321	0	1260	52	0
All	All	4341	0	4317	167	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 19.

The worst 5 of 167 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:A:120:MET:HG2	2:B:47:SER:HB3	1.59	0.83
1:A:40:ARG:HH12	1:A:42:ARG:HH22	1.27	0.80
4:D:469:SER:H	4:D:556:ASP:HB3	1.46	0.80
3:C:201:GLU:OE1	3:C:204:ARG:NH1	2.19	0.76
1:A:106:ASP:OD1	1:A:131:ARG:NH1	2.22	0.72

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	97/135 (72%)	92 (95%)	5 (5%)	0	100	100
2	B	75/102 (74%)	69 (92%)	6 (8%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	203/212 (96%)	194 (96%)	9 (4%)	0	100	100
4	D	174/219 (80%)	160 (92%)	14 (8%)	0	100	100
All	All	549/668 (82%)	515 (94%)	34 (6%)	0	100	100

There are no Ramachandran outliers to report.

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	77/109 (71%)	68 (88%)	9 (12%)	6	27
2	B	59/78 (76%)	54 (92%)	5 (8%)	12	42
3	C	172/187 (92%)	150 (87%)	22 (13%)	5	23
4	D	137/178 (77%)	113 (82%)	24 (18%)	2	11
All	All	445/552 (81%)	385 (86%)	60 (14%)	4	22

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	285	ASP
3	C	345	ARG
4	D	546	VAL
3	C	324	LEU
3	C	384	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	308	GLN
3	C	314	GLN
4	D	520	ASN

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](#)

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	99/135 (73%)	-0.04	2 (2%) 65 59	103, 121, 164, 193	0
2	B	77/102 (75%)	-0.12	0 100 100	105, 122, 149, 178	0
3	C	205/212 (96%)	-0.19	3 (1%) 73 67	102, 129, 192, 206	0
4	D	180/219 (82%)	-0.15	0 100 100	96, 121, 160, 179	0
All	All	561/668 (83%)	-0.14	5 (0%) 84 78	96, 124, 178, 206	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	308	GLN	2.9
3	C	293	VAL	2.5
1	A	133	GLU	2.2
1	A	132	GLY	2.1
3	C	296	ALA	2.1

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.