wwPDB X-ray Structure Validation Summary Report (i)

Mar 19, 2024 – 12:53 pm GMT

PDB ID	:	8Q36
Title	:	Structure of Nucleosome Core with a Bound Metallopeptide Conjugate (Foamy Virus GAG Peptide-Au[I] Compound)
Authors	:	De Falco, L.; Batchelor, L.K.; Dyson, P.J.; Davey, C.A.
Deposited on	:	2023-08-03
Resolution	:	2.60 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 (1)) were used in the production of this report:

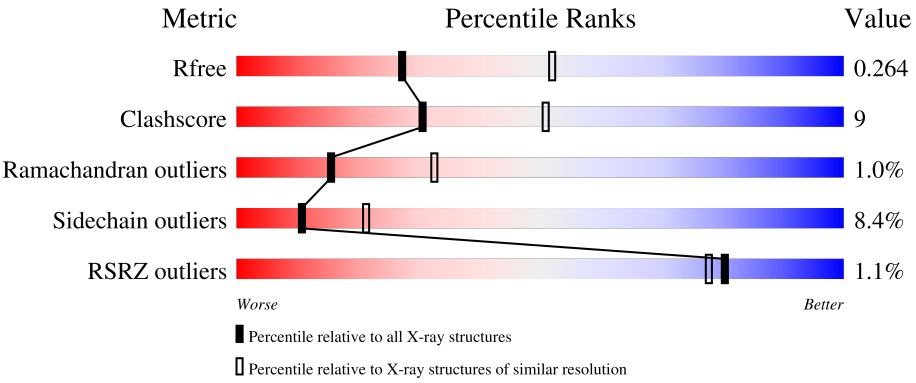
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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% 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	MMM	18	6% 72% 22% 6%
2	AAA	98	76% 22% ·
2	EEE	98	86% 12% ·
3	BBB	87	84% 10% 6%
3	FFF	87	79% 17% ··

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	CCC	107	2% 78% 19%
4	GGG	107	% 80% 18%
5	DDD	95	% 81% 18%
5	HHH	95	2% 71% 28%
6	III	145	% 48% 46% 6%
7	JJJ	145	2% 46% 53%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 12197 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 GAG structural protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	MMM	18	Total	C	N	O	0	0	1
			141	87	30	24			

- Molecule 2 is a protein called Histone H3.1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	AAA	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			
2	EEE	98	Total	C	N	O	S	0	0	0
			807	508	156	139	4			

- Molecule 3 is a protein called Histone H4.

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

- Molecule 4 is a protein called Histone H2A type 1-B/E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	CCC	107	Total	C	N	O	0	0	0
			828	523	162	143			
4	GGG	107	Total	C	N	O	0	0	0
			828	523	162	143			

- Molecule 5 is a protein called Histone H2B type 1-K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	DDD	95	Total	C	N	O	S	0	0	0
			745	467	136	140	2			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	HHH	95	Total	C	N	O	S	0	0	0
			745	467	136	140	2			

- Molecule 6 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	III	145	Total	C	N	O	P	0	0	0
			2970	1421	538	867	144			

- Molecule 7 is a DNA chain called DNA (145-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	JJJ	145	Total	C	N	O	P	0	0	0
			2969	1421	535	869	144			

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

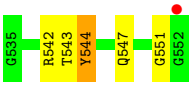
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	EEE	1	Total	Mg	0	0
			1	1		

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 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.

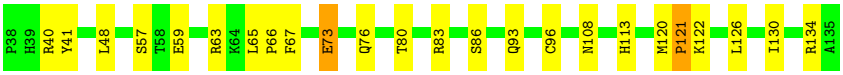
• Molecule 1: GAG structural protein

Chain MMM:



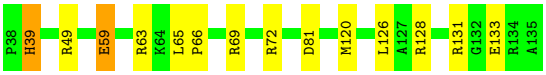
• Molecule 2: Histone H3.1

Chain AAA:



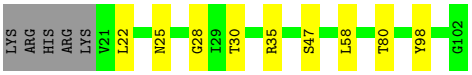
• Molecule 2: Histone H3.1

Chain EEE:



• Molecule 3: Histone H4

Chain BBB:



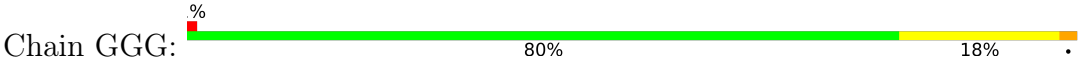
• Molecule 3: Histone H4

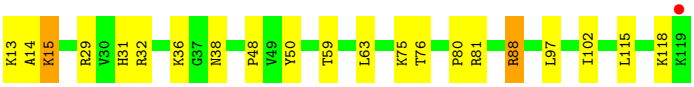
Chain FFF:

• Molecule 4: Histone H2A type 1-B/E

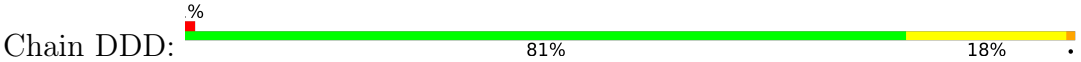
Chain CCC:

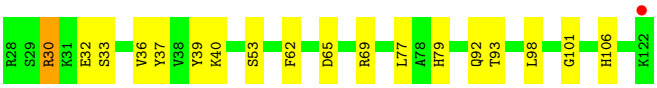
• Molecule 4: Histone H2A type 1-B/E



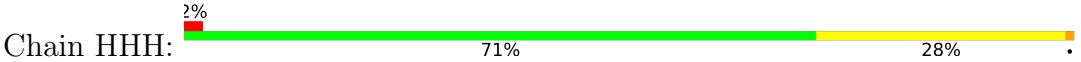


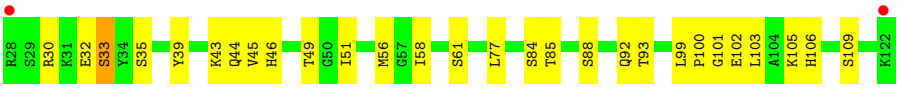
• Molecule 5: Histone H2B type 1-K



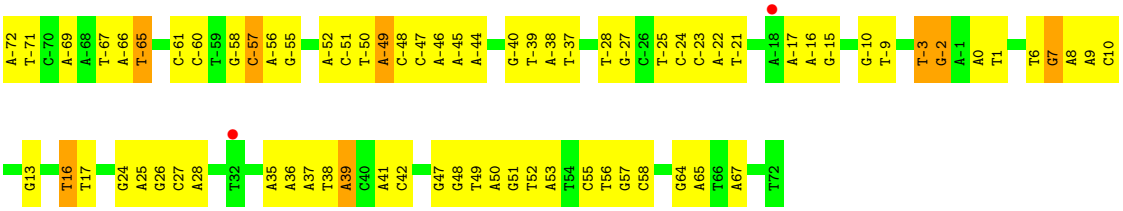


• Molecule 5: Histone H2B type 1-K

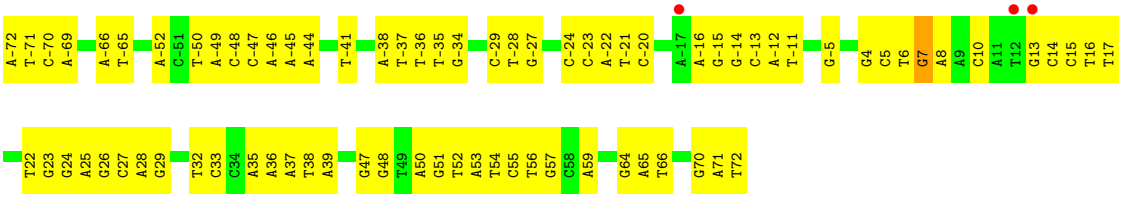




• Molecule 6: DNA (145-MER)



• Molecule 7: DNA (145-MER)



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	107.19Å 109.56Å 183.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.83 – 2.60 48.78 – 2.60	Depositor EDS
% Data completeness (in resolution range)	93.6 (48.83-2.60) 93.6 (48.78-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.05 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.213 , 0.262 0.215 , 0.264	Depositor DCC
R_{free} test set	1262 reflections (2.01%)	wwPDB-VP
Wilson B-factor (Å^2)	74.8	Xtriage
Anisotropy	0.441	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 50.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.025 for k,h,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	12197	wwPDB-VP
Average B, all atoms (Å^2)	104.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% 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: MG

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	MMM	0.71	0/145	1.02	0/195
2	AAA	0.70	0/819	0.97	0/1097
2	EEE	0.77	0/819	1.08	1/1097 (0.1%)
3	BBB	0.75	0/660	0.96	0/883
3	FFF	0.74	0/711	1.06	2/948 (0.2%)
4	CCC	0.72	0/838	1.03	1/1129 (0.1%)
4	GGG	0.77	0/838	0.92	0/1129
5	DDD	0.73	0/756	0.93	0/1014
5	HHH	0.78	0/756	0.94	0/1014
6	III	0.54	2/3332 (0.1%)	0.96	16/5141 (0.3%)
7	JJJ	0.55	2/3330 (0.1%)	0.91	5/5138 (0.1%)
All	All	0.65	4/13004 (0.0%)	0.96	25/18785 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	III	-3	DT	O3'-P	-6.07	1.53	1.61
7	JJJ	47	DG	O3'-P	-5.86	1.54	1.61
6	III	-57	DC	O3'-P	-5.39	1.54	1.61
7	JJJ	-37	DT	O3'-P	-5.05	1.55	1.61

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	JJJ	-5	DG	O5'-P-OP2	-10.86	95.93	105.70
6	III	-49	DA	C1'-O4'-C4'	-9.30	100.80	110.10
6	III	9	DA	O5'-P-OP2	-8.87	97.72	105.70
6	III	-2	DG	O5'-P-OP2	-8.78	97.80	105.70
7	JJJ	-65	DT	O5'-P-OP1	-8.56	98.00	105.70

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	MMM	141	0	131	1	0
2	AAA	807	0	844	17	0
2	EEE	807	0	844	11	0
3	BBB	653	0	696	8	0
3	FFF	703	0	755	13	0
4	CCC	828	0	892	18	0
4	GGG	828	0	892	15	0
5	DDD	745	0	769	14	0
5	HHH	745	0	769	12	0
6	III	2970	0	1640	68	0
7	JJJ	2969	0	1641	73	0
8	EEE	1	0	0	0	0
All	All	12197	0	9873	202	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 9.

The worst 5 of 202 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)
4:GGG:50:TYR:OH	5:HHH:92:GLN:NE2	2.04	0.90
4:CCC:50:TYR:OH	5:DDD:92:GLN:NE2	2.07	0.85
7:JJJ:-49:DA:H4'	7:JJJ:-48:DC:OP1	1.73	0.85
4:GGG:102:ILE:HG23	5:HHH:58:ILE:HD13	1.59	0.83
7:JJJ:-22:DA:H2"	7:JJJ:-21:DT:O5'	1.80	0.80

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	MMM	16/18 (89%)	14 (88%)	1 (6%)	1 (6%)	1 1
2	AAA	96/98 (98%)	86 (90%)	9 (9%)	1 (1%)	15 32
2	EEE	96/98 (98%)	95 (99%)	1 (1%)	0	100 100
3	BBB	80/87 (92%)	76 (95%)	4 (5%)	0	100 100
3	FFF	85/87 (98%)	83 (98%)	1 (1%)	1 (1%)	13 27
4	CCC	105/107 (98%)	99 (94%)	4 (4%)	2 (2%)	8 15
4	GGG	105/107 (98%)	96 (91%)	9 (9%)	0	100 100
5	DDD	93/95 (98%)	79 (85%)	13 (14%)	1 (1%)	14 30
5	HHH	93/95 (98%)	85 (91%)	6 (6%)	2 (2%)	6 12
All	All	769/792 (97%)	713 (93%)	48 (6%)	8 (1%)	15 32

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	FFF	18	HIS
4	CCC	118	LYS
5	DDD	101	GLY
5	HHH	51	ILE
5	HHH	101	GLY

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	MMM	13/13 (100%)	9 (69%)	4 (31%)	0 0
2	AAA	85/85 (100%)	78 (92%)	7 (8%)	11 22
2	EEE	85/85 (100%)	81 (95%)	4 (5%)	26 50
3	BBB	67/72 (93%)	65 (97%)	2 (3%)	41 67
3	FFF	72/72 (100%)	67 (93%)	5 (7%)	15 31
4	CCC	85/85 (100%)	75 (88%)	10 (12%)	5 9
4	GGG	85/85 (100%)	79 (93%)	6 (7%)	14 29
5	DDD	81/81 (100%)	76 (94%)	5 (6%)	18 37
5	HHH	81/81 (100%)	69 (85%)	12 (15%)	3 5
All	All	654/659 (99%)	599 (92%)	55 (8%)	11 21

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	EEE	39	HIS
3	FFF	23	ARG
5	HHH	109	SER
5	HHH	88	SER
2	EEE	59	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 1 ligands modelled in this entry, 1 is 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

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, 95^{th} 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	MMM	18/18 (100%)	0.13	1 (5%) 24 19	70, 98, 134, 162	0
2	AAA	98/98 (100%)	-0.08	0 100 100	50, 69, 113, 157	0
2	EEE	98/98 (100%)	-0.09	0 100 100	39, 55, 92, 136	0
3	BBB	82/87 (94%)	-0.13	0 100 100	48, 68, 93, 170	0
3	FFF	87/87 (100%)	0.06	0 100 100	43, 56, 94, 155	0
4	CCC	107/107 (100%)	0.01	2 (1%) 66 62	45, 62, 109, 168	0
4	GGG	107/107 (100%)	0.05	1 (0%) 84 82	51, 74, 125, 159	0
5	DDD	95/95 (100%)	-0.00	1 (1%) 80 78	50, 68, 120, 162	0
5	HHH	95/95 (100%)	0.07	2 (2%) 63 58	51, 78, 125, 164	0
6	III	145/145 (100%)	-0.29	2 (1%) 75 71	71, 137, 185, 240	0
7	JJJ	145/145 (100%)	-0.36	3 (2%) 63 58	79, 133, 184, 216	0
All	All	1077/1082 (99%)	-0.09	12 (1%) 80 78	39, 77, 168, 240	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	CCC	119	LYS	9.9
4	GGG	119	LYS	5.4
1	MMM	552	GLY	4.7
5	DDD	122	LYS	4.3
4	CCC	14	ALA	3.5

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
8	MG	EEE	201	1/1	0.97	0.33	57,57,57,57	0

6.5 Other polymers (i)

There are no such residues in this entry.