



Full wwPDB/EMDatabank EM Map/Model Validation Report ⓘ

Feb 20, 2018 – 01:27 pm GMT

PDB ID : 5KHF
EMDB ID: : EMD-8250
Title : Fitted structure of rubella virus capsid protein
Authors : Mangala Prasad, V.; Klose, T.; Rossmann, M.G.
Deposited on : 2016-06-14
Resolution : 35.00 Å(reported)

This is a Full wwPDB/EMDatabank EM Map/Model Validation Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30686

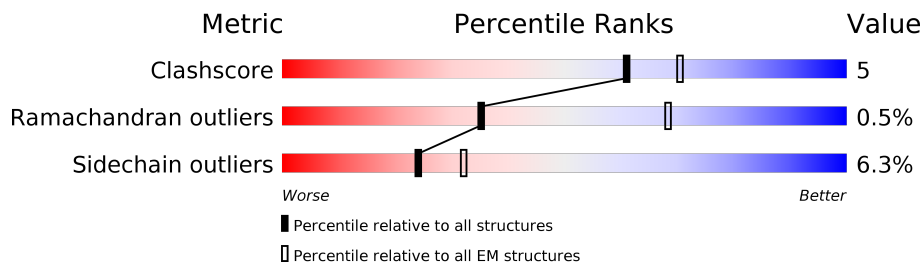
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 35.00 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)	EM structures (#Entries)
Clashscore	136279	1886
Ramachandran outliers	132675	1663
Sidechain outliers	132484	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	269	
1	B	269	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1599 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	98	781	496	142	139	4	0	0
1	B	98	783	497	142	140	4	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		AltConf
			Total	O	
2	A	35	35	35	0

4 Experimental information

Property	Value	Source
Reconstruction method	SUBTOMOGRAM AVERAGING	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of subtomograms used	18	Depositor
Resolution determination method	OTHER	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	1.0	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	11000	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

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 >2	RMSZ	# Z >2
1	A	0.43	0/809	0.60	0/1107
1	B	0.45	0/811	0.62	0/1109
All	All	0.44	0/1620	0.61	0/2216

There are no bond length outliers.

There are no bond angle outliers.

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	781	0	726	11	0
1	B	783	0	725	9	0
2	A	35	0	0	3	0
All	All	1599	0	1451	15	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 5.

All (15) 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:99:ARG:NH1	1:B:74:GLU:OE2	2.12	0.82

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:43:ARG:NH2	1:B:68:TYR:O	2.27	0.68
1:B:59:ASP:H	1:B:60:GLY:HA2	1.62	0.64
1:A:63:ASP:OD1	2:A:201:HOH:O	2.17	0.60
1:A:88:GLY:N	2:A:203:HOH:O	2.34	0.60
1:B:59:ASP:N	1:B:60:GLY:HA2	2.22	0.55
1:B:95:GLY:HA3	1:B:99:ARG:O	2.10	0.52
1:A:74:GLU:OE2	1:B:118:ARG:HD3	2.15	0.47
1:B:75:PRO:HG2	1:B:80:VAL:HG22	1.98	0.46
1:A:100:THR:H	1:A:118:ARG:HH21	1.63	0.46
1:A:75:PRO:HD3	1:B:27:CYS:N	2.31	0.45
1:A:87:ALA:HB3	1:A:104:GLN:OE1	2.19	0.43
1:A:91:ARG:HD3	1:A:105:ASP:HB3	2.01	0.42
1:A:81:ARG:NH1	1:B:120:PRO:O	2.49	0.42
1:A:53:THR:HG23	2:A:232:HOH:O	2.21	0.41

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 PDB entries followed by that with respect to all EM entries.

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	96/269 (36%)	86 (90%)	9 (9%)	1 (1%)	17	60
1	B	96/269 (36%)	91 (95%)	5 (5%)	0	100	100
All	All	192/538 (36%)	177 (92%)	14 (7%)	1 (0%)	35	74

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	88	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 PDB entries followed by that with respect to all EM entries.

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	79/219 (36%)	73 (92%)	6 (8%)	14	45
1	B	79/219 (36%)	75 (95%)	4 (5%)	26	57
All	All	158/438 (36%)	148 (94%)	10 (6%)	24	50

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	GLU
1	A	96	LYS
1	A	100	THR
1	A	116	LEU
1	A	118	ARG
1	A	121	VAL
1	B	37	GLU
1	B	47	HIS
1	B	94	TRP
1	B	96	LYS

Some 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 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.