Full wwPDB/EMDataBank EM Map/Model Validation Report (i)

Jan 16, 2017 – 02:23 PM EST

PDB ID	:	5H32
EMDB ID:	:	EMD-9574
Title	:	Cryo-EM structure of zika virus complexed with Fab C10 at pH 5.0
Authors	:	Zhang, S.; Kostyuchenko, V.; Ng, T.-S.; Lok, S.-M.
Deposited on	:	2016-10-20
Resolution	:	12.00 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
http://wwpdb.org/validation/2016/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

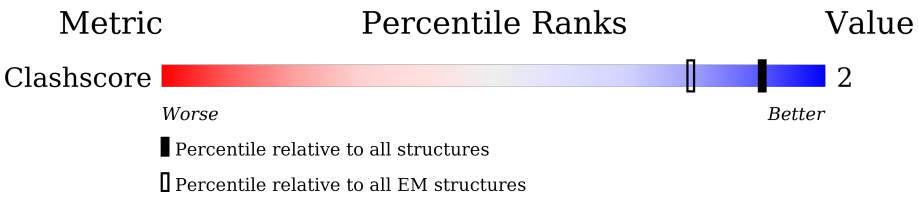
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 12.00 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	114402	924

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	403	100%
1	B	403	100%
1	C	403	100%
2	G	127	100%
2	I	127	100%
2	K	127	100%
3	H	109	98% .
3	L	109	96% .
3	M	109	100%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1917 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 structural protein E.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
1	A	403	403	403	0	403
1	C	403	403	403	0	403
1	B	403	403	403	0	403

- Molecule 2 is a protein called C10 IgG heavy chain variable region.

Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
2	G	127	127	127	0	127
2	K	127	127	127	0	127
2	I	127	127	127	0	127

- Molecule 3 is a protein called C10 IgG light chain variable region.

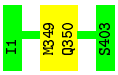
Mol	Chain	Residues	Atoms		AltConf	Trace
			Total	C		
3	H	109	109	109	0	109
3	L	109	109	109	0	109
3	M	109	109	109	0	109

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. 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: structural protein E

Chain A: 100%



- Molecule 1: structural protein E

Chain C: 100%

There are no outlier residues recorded for this chain.

- Molecule 1: structural protein E

Chain B: 100%

There are no outlier residues recorded for this chain.

- Molecule 2: C10 IgG heavy chain variable region

Chain G: 100%

There are no outlier residues recorded for this chain.

- Molecule 2: C10 IgG heavy chain variable region

Chain K: 100%

There are no outlier residues recorded for this chain.

- Molecule 2: C10 IgG heavy chain variable region

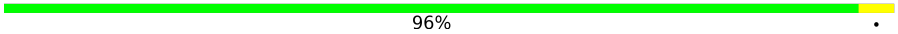
Chain I: 100%

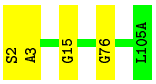
There are no outlier residues recorded for this chain.

- Molecule 3: C10 IgG light chain variable region

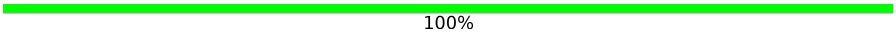
Chain H: 98%

• Molecule 3: C10 IgG light chain variable region

Chain L: 



• Molecule 3: C10 IgG light chain variable region

Chain M: 

There are no outlier residues recorded for this chain.

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	23810	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	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).

There are no protein, RNA or DNA chains available to summarize Z scores of covalent bonds and angles.

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	403	0	0	1	0
1	B	403	0	0	0	0
1	C	403	0	0	0	0
2	G	127	0	0	0	0
2	I	127	0	0	0	0
2	K	127	0	0	0	0
3	H	109	0	0	1	0
3	L	109	0	0	2	0
3	M	109	0	0	0	0
All	All	1917	0	0	4	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 2.

All (4) 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:L:15:GLY:CA	3:L:76:GLY:CA	2.35	1.04
1:A:349:MET:CA	1:A:350:GLN:CA	2.85	0.54
3:H:2:SER:CA	3:H:3:ALA:CA	2.91	0.48
3:L:2:SER:CA	3:L:3:ALA:CA	2.94	0.46

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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.