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

Jan 30, 2017 – 06:41 PM EST

PDB ID	:	5WTE
EMDB ID:	:	EMD-6686
Title	:	Cryo-EM structure for Hepatitis A virus full particle
Authors	:	Wang, X.; Zhu, L.; Dang, M.; Hu, Z.; Gao, Q.; Yuan, S.; Sun, Y.; Zhang, B.; Ren, J.; Walter, T.S.; Wang, J.; Fry, E.E.; Stuart, D.I.; Rao, Z.
Deposited on	:	2016-12-11
Resolution	:	3.40 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary 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.

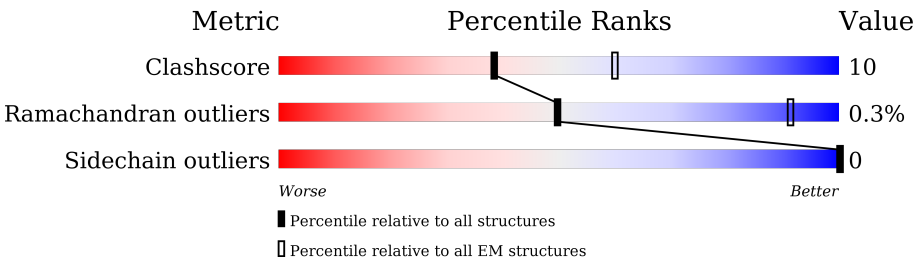
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

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 3.40 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
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	278	69% 24% • 6%
2	B	222	81% 17% •
3	C	246	78% 22%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5706 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 VP1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	260	2035	1300	332	395	8	0	0

- Molecule 2 is a protein called VP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	217	1715	1102	292	317	4	0	0

- Molecule 3 is a protein called VP3.

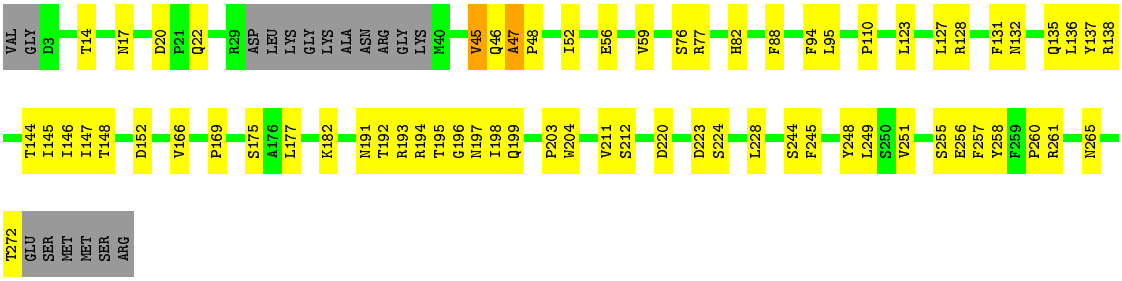
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	246	1956	1248	331	363	14	0	0

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 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.

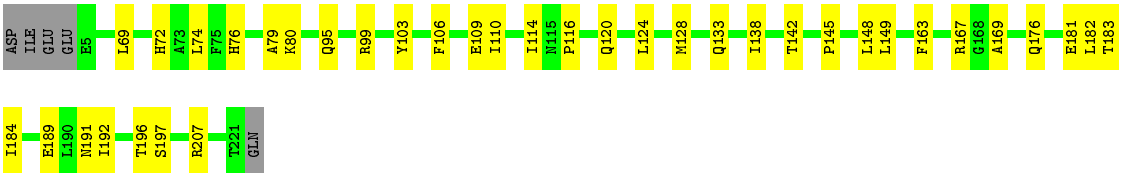
- Molecule 1: VP1

Chain A:



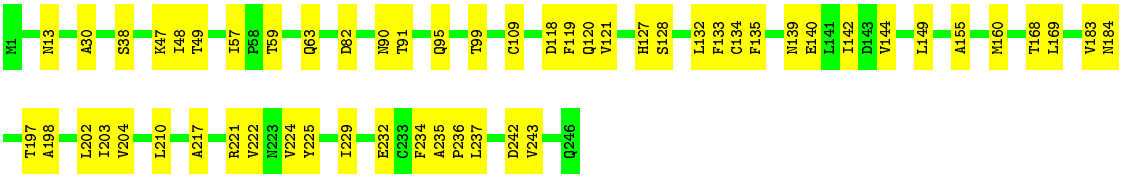
- Molecule 2: VP2

Chain B:



- Molecule 3: VP3

Chain C:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	2500	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	3000	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).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| >2$	RMSZ	$\#\|Z\| >2$
1	A	0.50	0/2090	0.55	1/2851 (0.0%)
2	B	0.51	0/1759	0.58	0/2397
3	C	0.52	0/2007	0.55	0/2733
All	All	0.51	0/5856	0.56	1/7981 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	47	ALA	C-N-CD	5.04	138.97	128.40

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	2035	0	1962	56	0
2	B	1715	0	1699	29	0
3	C	1956	0	1911	46	0
All	All	5706	0	5572	108	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 10.

The worst 5 of 108 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:265:ASN:HD22	3:C:95:GLN:HB2	1.34	0.93
2:B:69:LEU:H	2:B:72:HIS:HD2	1.22	0.85
2:B:95:GLN:HE21	2:B:99:ARG:HB2	1.48	0.78
1:A:56:GLU:OE2	3:C:120:GLN:NE2	2.17	0.77
2:B:120:GLN:NE2	2:B:197:SER:O	2.21	0.73

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	256/278 (92%)	240 (94%)	14 (6%)	2 (1%)	24 67
2	B	215/222 (97%)	203 (94%)	12 (6%)	0	100 100
3	C	244/246 (99%)	236 (97%)	8 (3%)	0	100 100
All	All	715/746 (96%)	679 (95%)	34 (5%)	2 (0%)	50 82

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	45	VAL
1	A	48	PRO

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	228/242 (94%)	228 (100%)	0	100	100
2	B	189/194 (97%)	189 (100%)	0	100	100
3	C	217/217 (100%)	217 (100%)	0	100	100
All	All	634/653 (97%)	634 (100%)	0	100	100

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

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	64	HIS
3	C	63	GLN
2	B	76	HIS
1	A	237	ASN
2	B	72	HIS

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.