wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 02:11 PM BST

PDB ID	:	5ACA
EMDB ID:	:	EMD-3130
Title	:	Structure-based energetics of protein interfaces guide Foot-and-Mouth disease virus vaccine design
Authors	:	Kotecha, A.; Seago, J.; Scott, K.; Burman, A.; Loureiro, S.; Ren, J.; Porta, C.; Ginn, H.M.; Jackson, T.; Perez-Martin, E.; Siebert, C.A.; Paul, G.; Huiskonen, J.T.; Jones, I.M.; Esnouf, R.M.; Fry, E.E.; Maree, F.F.; Charleston, B.; Stuart, D.I.
Deposited on	:	2015-08-14
Resolution	:	3.50 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

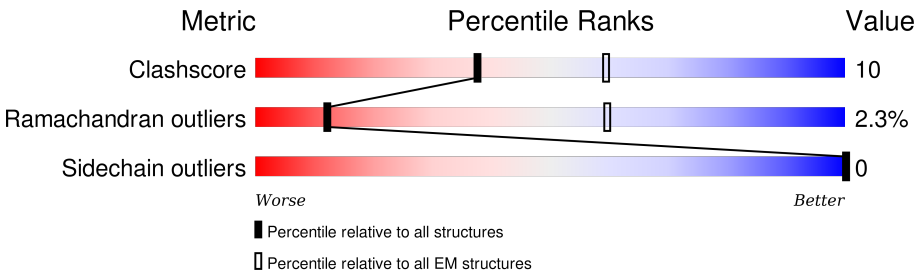
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 3.50 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)	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	1	214	65% 22% 12%
2	2	207	70% 29%
3	3	222	85% 14% •
4	4	85	44% 11% • 45%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5244 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	1	189	1506	949	267	282	8	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
1	64	ALA	GLY	CONFLICT	UNP Q1L764
1	172	TYR	HIS	CONFLICT	UNP Q1L764
1	189	GLU	ALA	CONFLICT	UNP Q1L764
1	190	LEU	VAL	CONFLICT	UNP Q1L764

- Molecule 2 is a protein called VP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	2	207	1638	1037	281	315	5	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
2	15	LEU	VAL	CONFLICT	UNP Q1L764
2	93	TYR	SER	ENGINEERED MUTATION	UNP Q1L764

- Molecule 3 is a protein called VP3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	3	222	1739	1113	285	332	9	0	0

- Molecule 4 is a protein called VP4.

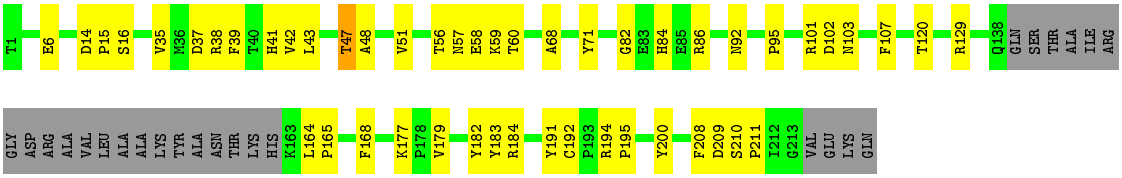
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	4	47	361	225	59	75	2	0	0

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.

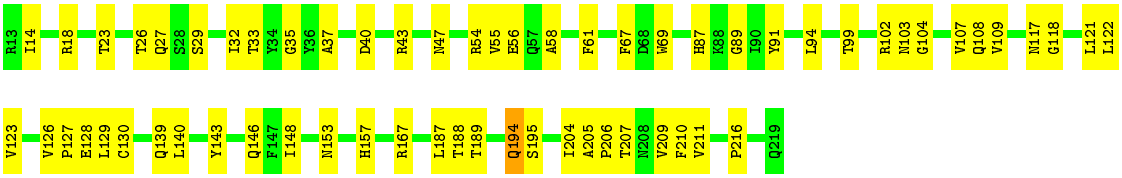
- Molecule 1: VP1

Chain 1:



- Molecule 2: VP2

Chain 2:

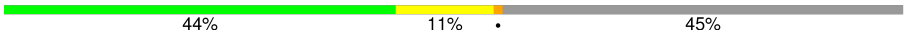


- Molecule 3: VP3

Chain 3:

- Molecule 4: VP4

Chain 4:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH PARTICLE, Not provided	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	25	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	160000	Depositor
Image detector	GATAN K2 (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	1	0.38	0/1546	0.48	0/2106
2	2	0.40	0/1679	0.50	0/2285
3	3	0.42	0/1798	0.50	0/2460
4	4	0.32	0/366	0.40	0/491
All	All	0.40	0/5389	0.49	0/7342

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	1	1506	0	1452	37	0
2	2	1638	0	1593	48	0
3	3	1739	0	1614	21	0
4	4	361	0	333	9	0
All	All	5244	0	4992	100	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 100 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)
3:3:63:VAL:HG11	3:3:73:MET:HB3	1.57	0.86
4:4:24:ASN:HD21	4:4:31:GLN:HE21	1.30	0.76
3:3:133:ARG:HB3	3:3:134:LEU:HA	1.69	0.74
2:2:107:VAL:HG22	2:2:204:ILE:HG12	1.73	0.71
2:2:54:ARG:NH1	2:2:56:GLU:OE2	2.24	0.69

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	1	185/214 (86%)	173 (94%)	6 (3%)	6 (3%)	5 40
2	2	205/207 (99%)	191 (93%)	11 (5%)	3 (2%)	13 56
3	3	220/222 (99%)	203 (92%)	12 (6%)	5 (2%)	8 48
4	4	43/85 (51%)	40 (93%)	2 (5%)	1 (2%)	8 48
All	All	653/728 (90%)	607 (93%)	31 (5%)	15 (2%)	12 48

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	2	194	GLN
1	1	209	ASP
3	3	69	GLY
3	3	70	ASP
3	3	145	TYR

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	1	163/182 (90%)	163 (100%)	0	100	100
2	2	179/179 (100%)	179 (100%)	0	100	100
3	3	186/186 (100%)	186 (100%)	0	100	100
4	4	38/68 (56%)	38 (100%)	0	100	100
All	All	566/615 (92%)	566 (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 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	2	108	GLN
4	4	31	GLN
3	3	85	HIS
2	2	103	ASN
3	3	82	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.