



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:10 PM BST

PDB ID : 5AC9  
EMDB ID: : EMD-3129  
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.; PerezMartin, 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 : 3.20 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](mailto: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

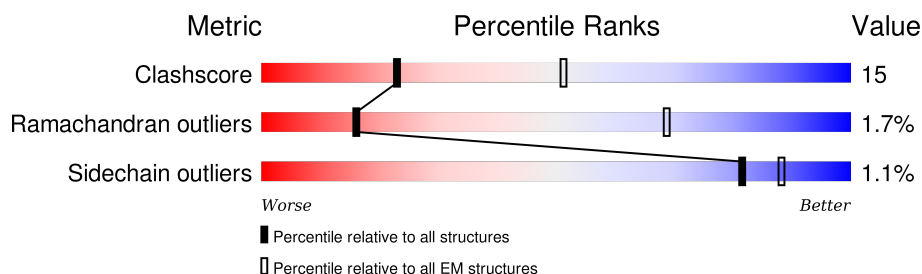
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.20 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  $\geq 3$ , 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	1	208	62% 25% • 11%
2	2	218	68% 25% • 5%
3	3	220	76% 22% •
4	4	85	40% 12% • 47%

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5115 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
1	1	186	Total	C	N	O	S	0	0
			1452	918	260	271	3		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
1	133	GLU	ASN	CONFLICT	UNP Q6PMW3

- Molecule 2 is a protein called VP3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	2	207	Total	C	N	O	S	0	0
			1631	1040	279	305	7		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
2	93	TYR	SER	ENGINEERED MUTATION	UNP Q6PMW3

- Molecule 3 is a protein called VP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	3	220	Total	C	N	O	S	0	0
			1687	1081	274	323	9		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
3	168	THR	ALA	CONFLICT	UNP Q6PMW3
3	174	VAL	THR	CONFLICT	UNP Q6PMW3

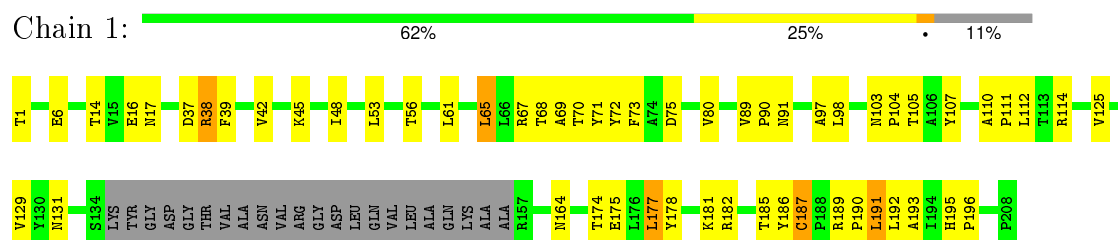
- Molecule 4 is a protein called VP4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	4	45	Total	C	N	O	S	0	0
			345	218	55	70	2		

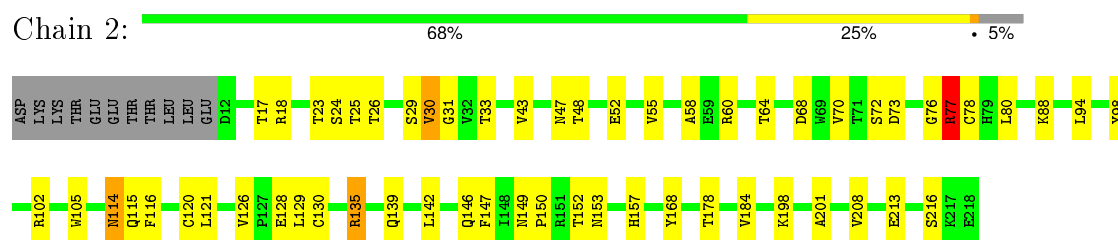
### 3 Residue-property plots [i](#)

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.

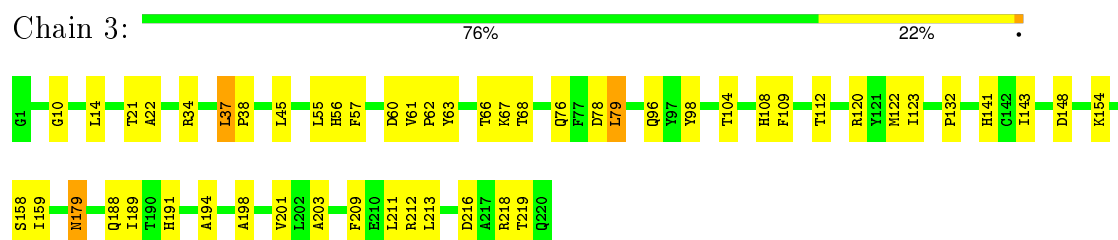
#### • Molecule 1: VP1



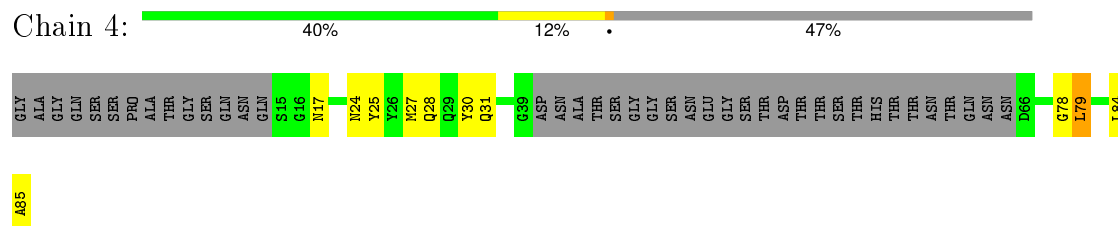
#### • Molecule 2: VP3



#### • Molecule 3: VP2



#### • Molecule 4: VP4



## 4 Experimental information

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	INDIVIDUAL PARTICLES, Not provided	Depositor
Microscope	FEI TECNAI F30	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	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	0.87	1/1486 (0.1%)	0.98	3/2031 (0.1%)
2	2	0.91	1/1676 (0.1%)	0.93	2/2289 (0.1%)
3	3	0.96	0/1737	0.96	3/2373 (0.1%)
4	4	0.93	1/351 (0.3%)	0.99	1/470 (0.2%)
All	All	0.92	3/5250 (0.1%)	0.96	9/7163 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	4	30	TYR	CE2-CZ	5.47	1.45	1.38
2	2	105	TRP	CB-CG	-5.35	1.40	1.50
1	1	114	ARG	CZ-NH1	5.33	1.40	1.33

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	1	114	ARG	NE-CZ-NH2	-7.84	116.38	120.30
3	3	45	LEU	CB-CG-CD1	-6.51	99.93	111.00
2	2	142	LEU	CB-CG-CD1	-6.25	100.38	111.00
4	4	79	LEU	CA-CB-CG	5.83	128.70	115.30
2	2	121	LEU	CB-CG-CD1	-5.66	101.38	111.00

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	1452	0	1451	52	0
2	2	1631	0	1582	44	0
3	3	1687	0	1614	61	0
4	4	345	0	318	6	0
All	All	5115	0	4965	148	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 15.

The worst 5 of 148 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)
1:1:69:ALA:CB	1:1:185:THR:HG23	1.68	1.23
1:1:69:ALA:HB2	1:1:185:THR:CG2	1.70	1.20
3:3:66:THR:CG2	3:3:188:GLN:NE2	2.15	1.10
1:1:69:ALA:HB2	1:1:185:THR:HG23	1.05	1.02
3:3:66:THR:CG2	3:3:188:GLN:HE22	1.70	1.02

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	182/208 (88%)	166 (91%)	13 (7%)	3 (2%)	12	54
2	2	205/218 (94%)	178 (87%)	23 (11%)	4 (2%)	9	48
3	3	218/220 (99%)	199 (91%)	17 (8%)	2 (1%)	21	67
4	4	41/85 (48%)	36 (88%)	3 (7%)	2 (5%)	3	22
All	All	646/731 (88%)	579 (90%)	56 (9%)	11 (2%)	16	52

5 of 11 Ramachandran outliers are listed below:



Mol	Chain	Res	Type
2	2	24	SER
2	2	30	VAL
2	2	77	ARG
4	4	28	GLN
1	1	104	PRO

### 5.3.2 Protein sidechains ⓘ

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	154/170 (91%)	151 (98%)	3 (2%)	65	89
2	2	179/190 (94%)	177 (99%)	2 (1%)	80	94
3	3	177/177 (100%)	176 (99%)	1 (1%)	90	97
4	4	36/67 (54%)	36 (100%)	0	100	100
All	All	546/604 (90%)	540 (99%)	6 (1%)	81	94

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	1	174	THR
3	3	37	LEU
2	2	77	ARG
1	1	65	LEU
2	2	135	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
3	3	88	ASN
3	3	141	HIS
3	3	179	ASN
3	3	56	HIS
3	3	152	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

### 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 5.6 Ligand geometry ⓘ

There are no ligands in this entry.

### 5.7 Other polymers ⓘ

There are no such residues in this entry.

### 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.