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

Feb 19, 2018 – 02:09 am GMT

PDB ID	:	5L7Q
EMDB ID:	:	EMD-4009
Title	:	Structure of deformed wing virus, a honeybee pathogen
Authors	:	Skubnik, K.; Novacek, J.; Fuzik, T.; Pridal, A.; Paxton, R.; Plevka, P.
Deposited on	:	2016-06-03
Resolution	:	3.44 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
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

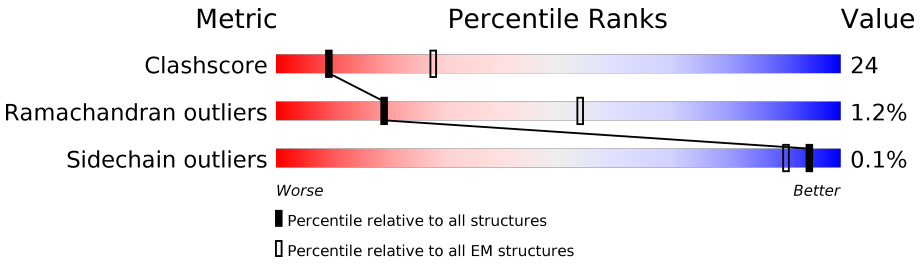
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30686

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.44 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 <=5%

Mol	Chain	Length	Quality of chain
1	A	258	50% 48% .
2	B	253	53% 45% ..
3	C	416	45% 50% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7091 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	252	1976	1256	336	374	10	0	0

- Molecule 2 is a protein called vp2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	250	1973	1253	335	378	7	0	0

- Molecule 3 is a protein called vp3.

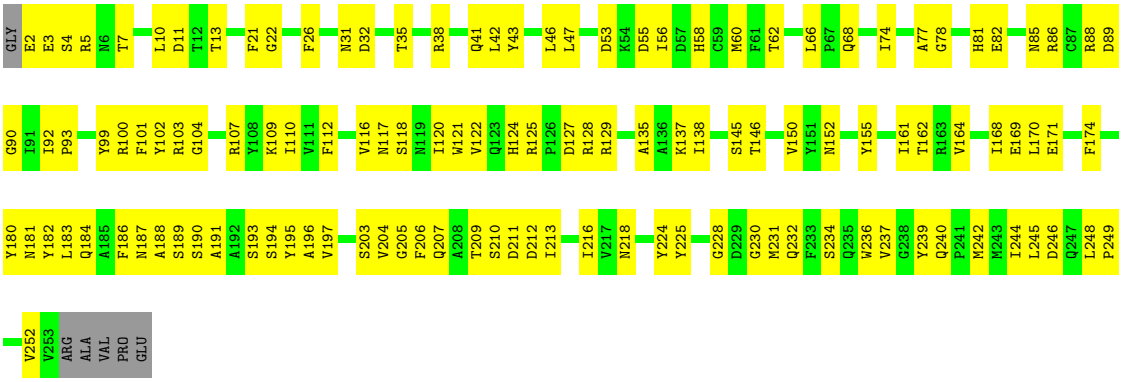
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	397	3142	2015	541	574	12	0	0

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 the various outlier classes 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.

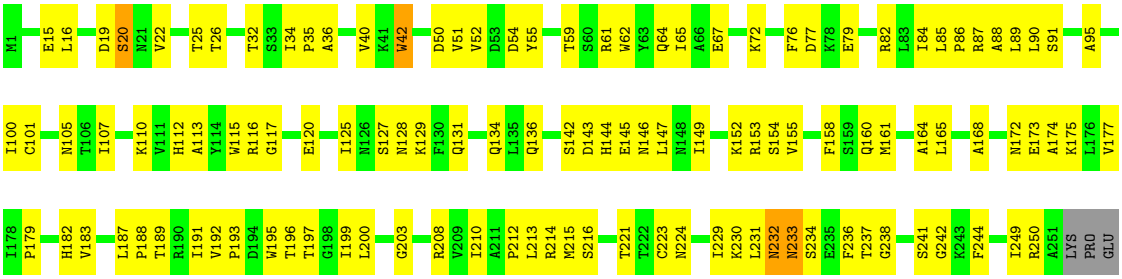
- Molecule 1: VP1

Chain A:



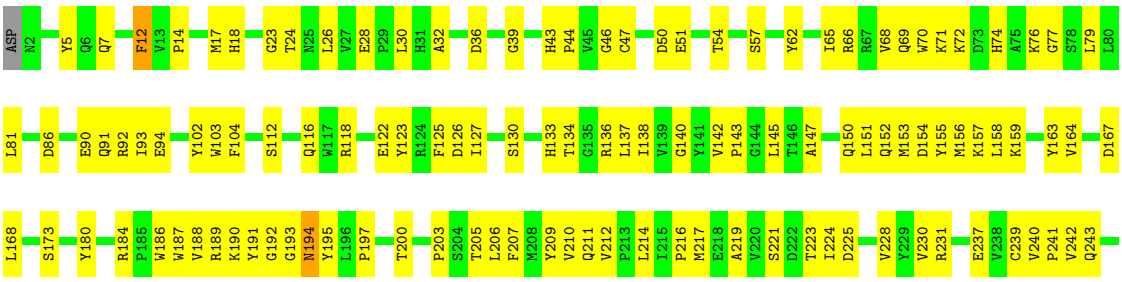
- Molecule 2: vp2

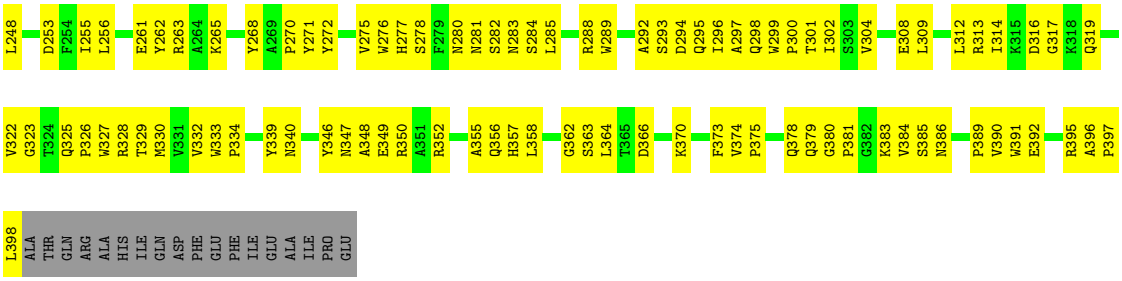
Chain B:



- Molecule 3: vp3

Chain C:




L248 D253 F254 I255 L256 E261 Y262 R263 A264 K265 Y268 A269 P270 Y271 Y272 V275 W276 H277 S278 F279 N280 N281 S282 N283 S284 L285 R288 W289 A292 S293 D294 Q295 I296 A297 Q298 W299 P300 T301 I302 S303 V304 E308 L309 L312 R313 I314 K315 D316 G317 K318 Q319
V322 G323 T324 Q325 P326 W327 R328 T329 M330 V331 V332 W333 P334 Y339 N340 Y346 N347 A348 E349 R350 A351 R352 A355 Q356 H357 L358 G362 S363 L364 T365 D366 K370 F373 V374 P375 Q378 Q379 G380 P381 G382 K383 V384 S385 N386 P389 V390 W391 E392 R395 A396 P397
L398 ALA THR GLN ARG ALA HIS ILE GLN ASP PHE GLU PHE ILE GLU ALA ILE PRO GLU

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	136828	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	15	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	74235	Depositor
Image detector	FEI FALCON II (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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	0.61	0/2024	0.59	0/2754
2	B	0.61	0/2024	0.62	0/2763
3	C	0.59	0/3238	0.56	0/4418
All	All	0.60	0/7286	0.58	0/9935

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

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	1976	0	1911	113	0
2	B	1973	0	1928	96	0
3	C	3142	0	3051	172	0
All	All	7091	0	6890	342	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 24.

The worst 5 of 342 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:86:ARG:HD2	3:C:194:ASN:HD22	1.36	0.89

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:150:GLN:HE22	3:C:205:THR:HG21	1.37	0.88
2:B:84:ILE:HG21	2:B:87:ARG:HD3	1.55	0.86
2:B:42:TRP:HE1	3:C:46:GLY:HA3	1.44	0.82
1:A:187:ASN:HB3	2:B:147:LEU:HB3	1.62	0.81

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	250/258 (97%)	227 (91%)	20 (8%)	3 (1%)	14 53
2	B	248/253 (98%)	221 (89%)	21 (8%)	6 (2%)	6 38
3	C	395/416 (95%)	362 (92%)	31 (8%)	2 (0%)	31 71
All	All	893/927 (96%)	810 (91%)	72 (8%)	11 (1%)	19 53

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	76	PHE
2	B	233	ASN
3	C	12	PHE
3	C	194	ASN
1	A	252	VAL

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	215/219 (98%)	215 (100%)	0	100 100
2	B	221/224 (99%)	220 (100%)	1 (0%)	90 95
3	C	338/354 (96%)	338 (100%)	0	100 100
All	All	774/797 (97%)	773 (100%)	1 (0%)	94 98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	195	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	136	GLN
2	B	146	ASN
3	C	243	GLN
2	B	134	GLN
3	C	280	ASN

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.