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

Feb 1, 2017 – 04:43 PM EST

PDB ID	:	5LWG
EMDB ID:	:	EMD-4114
Title	:	Israeli acute paralysis virus heated to 63 degree - full particle
Authors	:	Mullapudi, E.; Fuzik, T.; Pridal, A.; Plevka, P.
Deposited on	:	2016-09-16
Resolution	:	3.20 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.

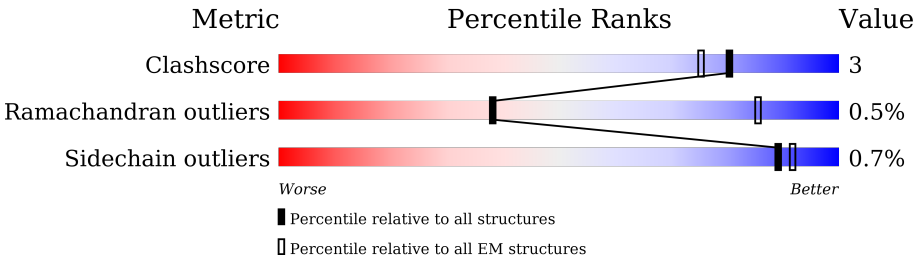
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

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 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 >=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	208	87% 13%
2	C	300	95% 5%
3	B	247	87% 11% .
4	D	57	82% 12% 5%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 6373 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	A	208	Total	C	N	O	S	0	0
			1676	1061	287	322	6		

- Molecule 2 is a protein called VP3.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	300	Total	C	N	O	S	0	0
			2337	1492	388	443	14		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	106	ARG	SER	conflict	UNP G0Z733
C	109	ALA	GLU	conflict	UNP G0Z733
C	118	MET	ILE	conflict	UNP G0Z733
C	177	ARG	GLN	conflict	UNP G0Z733

- Molecule 3 is a protein called VP2.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	B	247	Total	C	N	O	S	0	0
			1944	1226	324	388	6		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	115	ARG	LYS	conflict	UNP B3TZF1
B	255	ILE	MET	conflict	UNP B3TZF1

- Molecule 4 is a protein called VP4.

Mol	Chain	Residues	Atoms				AltConf	Trace
4	D	57	Total	C	N	O	0	0
			416	270	68	78		

There is a discrepancy between the modelled and reference sequences:

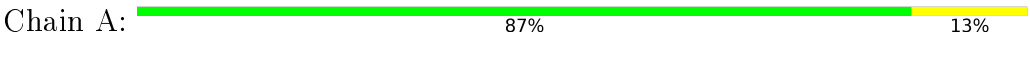
Chain	Residue	Modelled	Actual	Comment	Reference
D	19	ILE	VAL	conflict	UNP B3TZL5

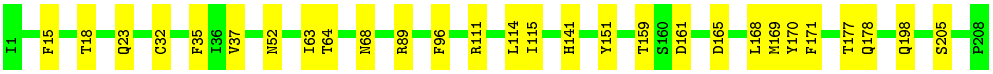
3 Residue-property plots (i)

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

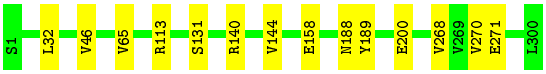
Chain A:





- Molecule 2: VP3

Chain C:

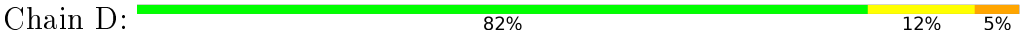


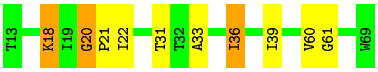
- Molecule 3: VP2

Chain B:

- Molecule 4: VP4

Chain D:





4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	9614	Depositor
Resolution determination method	FSC 0.143 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)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	75000	Depositor
Image detector	Not provided	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	0.39	0/1718	0.61	0/2337
2	C	0.35	0/2400	0.57	0/3287
3	B	0.37	0/1984	0.63	0/2708
4	D	0.47	0/425	0.81	1/582 (0.2%)
All	All	0.38	0/6527	0.62	1/8914 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
3	B	0	2
4	D	0	1
All	All	0	3

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	18	LYS	CD-CE-NZ	7.81	129.66	111.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	33	THR	Peptide
3	B	49	GLN	Peptide
4	D	22	ILE	Peptide

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	1676	0	1618	19	0
2	C	2337	0	2300	11	0
3	B	1944	0	1908	17	0
4	D	416	0	429	4	0
All	All	6373	0	6255	41	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 3.

All (41) 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)
4:D:60:VAL:H	4:D:61:GLY:HA3	1.67	0.59
1:A:96:PHE:HB2	1:A:141:HIS:HB3	1.85	0.59
2:C:65:VAL:H	3:B:175:THR:HG21	1.69	0.58
1:A:170:TYR:OH	1:A:178:GLN:NE2	2.38	0.56
3:B:74:ASP:OD2	3:B:105:ARG:NH1	2.42	0.53
1:A:52:ASN:OD1	1:A:177:THR:OG1	2.25	0.52
3:B:100:SER:OG	3:B:168:GLN:NE2	2.43	0.51
3:B:82:THR:HG22	3:B:84:ALA:H	1.75	0.51
1:A:114:LEU:HD11	2:C:32:LEU:HB3	1.92	0.50
1:A:37:VAL:HG11	4:D:33:ALA:HB1	1.94	0.50
1:A:89:ARG:HD2	2:C:46:VAL:HG11	1.93	0.49
1:A:115:ILE:HG13	1:A:169:MET:HB2	1.94	0.48
2:C:131:SER:HB2	2:C:271:GLU:HB3	1.95	0.48
1:A:35:PHE:HE1	2:C:268:VAL:HG12	1.79	0.47
1:A:159:THR:HG23	1:A:161:ASP:H	1.80	0.47
1:A:151:TYR:OH	1:A:161:ASP:OD2	2.33	0.46
2:C:140:ARG:NH1	2:C:200:GLU:OE2	2.48	0.46
3:B:33:THR:HA	3:B:34:PHE:HB2	1.97	0.46
1:A:111:ARG:HB2	1:A:171:PHE:HB3	1.98	0.46
3:B:71:VAL:HG11	3:B:117:GLY:HA3	1.98	0.45
3:B:50:ASP:OD2	3:B:51:THR:OG1	2.31	0.45
1:A:165:ASP:OD1	1:A:165:ASP:N	2.50	0.45
1:A:64:THR:O	1:A:68:ASN:ND2	2.50	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:162:THR:HG22	3:B:168:GLN:HG3	1.99	0.44
2:C:144:VAL:HG22	3:B:148:ARG:HE	1.82	0.44
1:A:63:ILE:HG12	1:A:168:LEU:HB3	2.00	0.44
3:B:129:LEU:HD13	3:B:248:LEU:HD13	1.99	0.43
1:A:15:PHE:CD1	2:C:189:TYR:HE2	2.36	0.43
3:B:187:ASP:OD1	3:B:187:ASP:N	2.51	0.42
1:A:18:THR:HB	1:A:23:GLN:HE21	1.85	0.42
3:B:114:LEU:HD22	3:B:123:LEU:HD21	2.01	0.42
1:A:198:GLN:HE21	3:B:170:SER:N	2.18	0.42
3:B:109:LEU:HB2	3:B:213:VAL:HG13	2.03	0.41
3:B:60:THR:H	3:B:61:HIS:HA	1.85	0.41
4:D:20:GLY:HA2	4:D:21:PRO:HD3	1.94	0.41
3:B:59:ASP:HA	3:B:60:THR:HA	1.72	0.41
4:D:31:THR:HG21	4:D:39:ILE:HD12	2.03	0.41
1:A:32:CYS:SG	2:C:270:VAL:HG13	2.61	0.41
3:B:196:TYR:CZ	3:B:198:SER:HB3	2.56	0.41
2:C:158:GLU:OE1	2:C:188:ASN:ND2	2.47	0.40
1:A:205:SER:HA	2:C:113:ARG:HA	2.03	0.40

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	206/208 (99%)	183 (89%)	23 (11%)	0	100 100
2	C	298/300 (99%)	284 (95%)	14 (5%)	0	100 100
3	B	245/247 (99%)	215 (88%)	28 (11%)	2 (1%)	24 69
4	D	55/57 (96%)	35 (64%)	18 (33%)	2 (4%)	4 30
All	All	804/812 (99%)	717 (89%)	83 (10%)	4 (0%)	38 78

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	20	GLY
3	B	202	ALA
3	B	50	ASP
4	D	36	ILE

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	186/186 (100%)	186 (100%)	0	100 100
2	C	262/262 (100%)	262 (100%)	0	100 100
3	B	222/222 (100%)	219 (99%)	3 (1%)	74 92
4	D	44/44 (100%)	42 (96%)	2 (4%)	34 74
All	All	714/714 (100%)	709 (99%)	5 (1%)	89 97

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	21	THR
3	B	33	THR
3	B	51	THR
4	D	18	LYS
4	D	36	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	5	ASN
1	A	141	HIS
1	A	178	GLN
1	A	198	GLN
2	C	67	HIS
2	C	83	ASN
2	C	255	ASN
3	B	56	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	B	168	GLN

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.