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

Jan 31, 2019 – 07:02 AM EST

PDB ID	:	6IIC
EMDB ID:	:	EMD-9673
Title	:	CryoEM structure of Mud Crab Dicistrovirus
Authors	:	Zhang, Q.; Gao, Y.
Deposited on	:	2018-10-04
Resolution	:	3.30 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
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

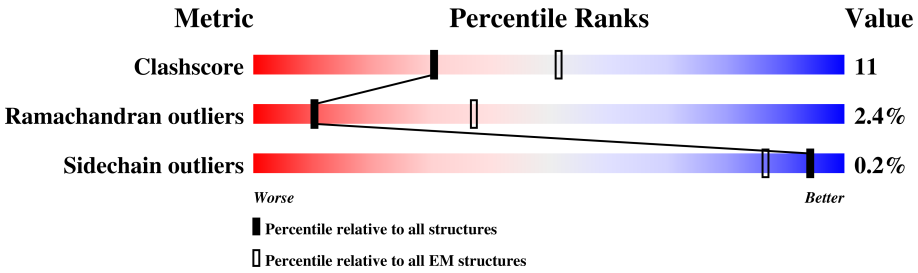
MolProbity	:	4.02b-467
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031633

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 3.30 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	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	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	191	69% 28%
2	B	253	71% 26%
3	C	448	41% 15% 44%
4	D	58	88% 9%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5748 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 of Mud crab dicistrovirus.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	188	1438	900	241	291	6	0	0

- Molecule 2 is a protein called VP2 of Mud crab dicistrovirus.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	250	1945	1217	332	388	8	0	0

- Molecule 3 is a protein called VP3 of Mud crab dicistrovirus.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	253	1966	1264	319	374	9	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	31	PRO	SER	see sequence details	UNP E5G7H9
C	32	LEU	THR	see sequence details	UNP E5G7H9
C	33	SER	PHE	see sequence details	UNP E5G7H9
C	34	LEU	ALA	see sequence details	UNP E5G7H9

- Molecule 4 is a protein called VP4 of Mud crab dicistrovirus.

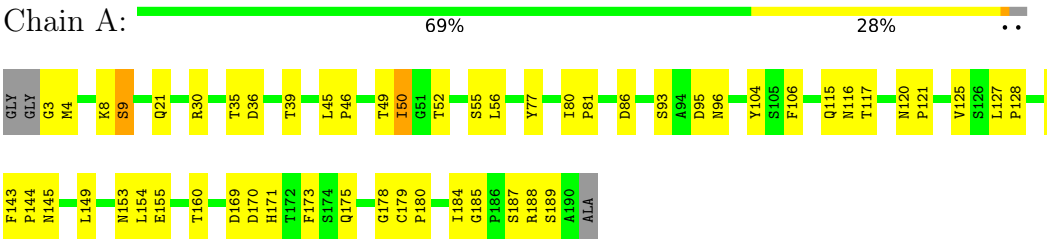
Mol	Chain	Residues	Atoms				AltConf	Trace
			Total	C	N	O		
4	D	56	399	251	66	82	0	0

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.

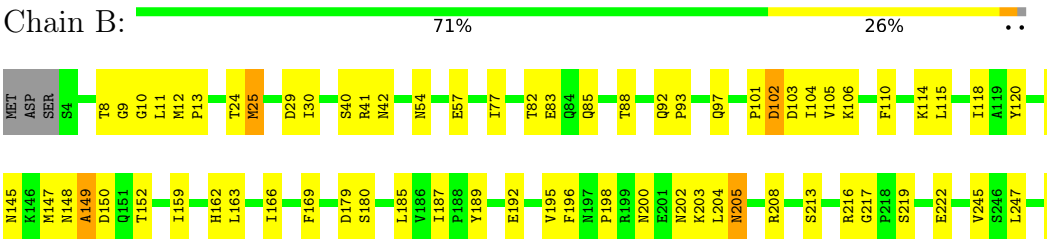
- Molecule 1: VP1 of Mud crab dicistrovirus

Chain A:



- Molecule 2: VP2 of Mud crab dicistrovirus

Chain B:

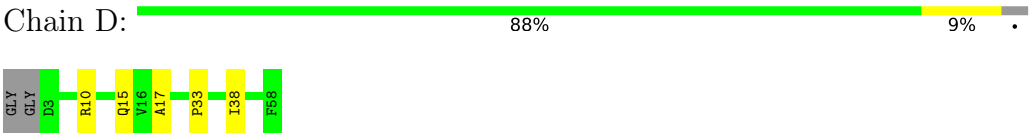


- Molecule 3: VP3 of Mud crab dicistrovirus

Chain C:

• Molecule 4: VP4 of Mud crab dicistrovirus

Chain D:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	31801	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	20	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	96000	Depositor
Image detector	GATAN ULTRASCAN 4000 (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	0.43	0/1464	0.59	0/1994
2	B	0.41	0/1977	0.56	0/2685
3	C	0.44	0/2021	0.60	0/2764
4	D	0.39	0/404	0.55	0/553
All	All	0.42	0/5866	0.58	0/7996

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	1438	0	1426	35	0
2	B	1945	0	1951	51	0
3	C	1966	0	1928	49	0
4	D	399	0	401	2	0
All	All	5748	0	5706	123	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 11.

All (123) 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)
2:B:102:ASP:OD1	2:B:106:LYS:HE2	1.62	0.98
2:B:196:PHE:CD2	2:B:202:ASN:ND2	2.45	0.85
2:B:102:ASP:CG	2:B:106:LYS:HE2	1.97	0.84
2:B:101:PRO:HB3	2:B:125:TYR:OH	1.79	0.83
2:B:196:PHE:HD2	2:B:202:ASN:HD22	1.28	0.78
2:B:102:ASP:OD2	2:B:106:LYS:CE	2.31	0.78
2:B:196:PHE:HD2	2:B:202:ASN:ND2	1.83	0.76
1:A:179:CYS:SG	3:C:110:PRO:HA	2.27	0.74
1:A:127:LEU:HD12	1:A:128:PRO:HD2	1.72	0.71
2:B:216:ARG:HB3	3:C:220:ASN:HD22	1.55	0.71
3:C:100:ILE:HG22	3:C:101:ARG:HG3	1.73	0.70
1:A:45:LEU:HD22	1:A:149:LEU:HD23	1.74	0.70
2:B:102:ASP:CG	2:B:106:LYS:CE	2.62	0.68
3:C:115:VAL:HG12	3:C:236:LEU:HD21	1.77	0.67
3:C:73:GLY:O	3:C:74:ASN:HB2	1.97	0.63
2:B:97:GLN:OE1	2:B:208:ARG:NH1	2.32	0.63
2:B:145:ASN:ND2	2:B:187:ILE:HD12	2.14	0.61
3:C:220:ASN:OD1	3:C:221:ASP:N	2.33	0.61
2:B:77:ILE:HD11	2:B:163:LEU:HD13	1.84	0.60
2:B:102:ASP:OD2	2:B:106:LYS:HE2	1.99	0.59
2:B:145:ASN:HD22	2:B:187:ILE:HD12	1.67	0.59
2:B:150:ASP:OD2	2:B:203:LYS:NZ	2.36	0.59
2:B:217:GLY:HA3	2:B:222:GLU:HG3	1.82	0.59
1:A:178:GLY:H	2:B:169:PHE:HE1	1.51	0.59
3:C:126:HIS:ND1	3:C:176:CYS:SG	2.71	0.59
1:A:95:ASP:OD2	1:A:145:ASN:ND2	2.36	0.58
2:B:82:THR:H	2:B:85:GLN:HE21	1.50	0.58
1:A:39:THR:HG22	1:A:160:THR:HG22	1.85	0.58
3:C:88:ILE:HD12	3:C:125:ILE:HD12	1.85	0.57
3:C:143:VAL:HG22	3:C:207:ILE:HG12	1.87	0.56
1:A:49:THR:O	1:A:50:ILE:HG13	2.05	0.56
3:C:72:GLY:O	3:C:221:ASP:HB3	2.05	0.56
1:A:3:GLY:O	1:A:4:MET:HB2	2.06	0.56
1:A:81:PRO:HG3	1:A:117:THR:HG21	1.88	0.56
1:A:180:PRO:HB3	2:B:159:ILE:HD12	1.89	0.55
1:A:187:SER:HB2	3:C:103:TYR:CZ	2.42	0.55
1:A:125:VAL:HG13	3:C:32:LEU:HD13	1.90	0.54
2:B:185:LEU:HD21	2:B:187:ILE:HD11	1.89	0.54
2:B:148:ASN:O	2:B:150:ASP:N	2.41	0.54
3:C:142:LEU:HD23	3:C:144:GLN:HE22	1.73	0.54
2:B:29:ASP:OD1	2:B:30:ILE:N	2.41	0.53
2:B:101:PRO:HB3	2:B:125:TYR:CZ	2.44	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:40:SER:OG	2:B:41:ARG:N	2.42	0.52
1:A:21:GLN:NE2	3:C:235:GLU:OE2	2.43	0.52
2:B:110:PHE:O	2:B:114:LYS:HG2	2.11	0.51
2:B:85:GLN:HA	2:B:88:THR:HG22	1.91	0.51
1:A:170:ASP:OD1	3:C:40:ILE:HG13	2.10	0.51
3:C:65:LEU:HG	3:C:227:PRO:HG2	1.91	0.51
2:B:120:TYR:HB3	2:B:195:VAL:HB	1.91	0.51
3:C:190:ASP:HB3	3:C:200:CYS:SG	2.51	0.50
3:C:27:ASP:O	3:C:28:TYR:HB2	2.11	0.49
1:A:55:SER:OG	1:A:56:LEU:N	2.44	0.49
3:C:109:ALA:O	3:C:112:SER:OG	2.21	0.49
2:B:162:HIS:O	2:B:166:ILE:HG12	2.12	0.49
2:B:105:VAL:HG11	2:B:198:PRO:HB2	1.93	0.49
2:B:204:LEU:HD12	2:B:205:ASN:OD1	2.12	0.49
3:C:127:LEU:HD23	3:C:229:LEU:HB3	1.95	0.49
3:C:142:LEU:HD23	3:C:144:GLN:NE2	2.28	0.48
3:C:168:VAL:HG13	3:C:173:ILE:HG12	1.96	0.48
1:A:35:THR:OG1	1:A:36:ASP:N	2.39	0.48
1:A:8:LYS:O	1:A:9:SER:OG	2.27	0.48
3:C:142:LEU:HD13	3:C:210:LEU:HB2	1.95	0.48
3:C:66:ASP:OD1	3:C:67:ARG:N	2.46	0.48
2:B:149:ALA:O	2:B:152:THR:HG22	2.13	0.48
1:A:104:TYR:CE1	1:A:106:PHE:HB2	2.49	0.47
1:A:52:THR:HG21	1:A:139:ASP:OD2	2.13	0.47
1:A:137:VAL:HG12	2:B:192:GLU:OE2	2.15	0.47
2:B:115:LEU:HD22	2:B:118:ILE:HD11	1.97	0.47
3:C:71:LYS:HB2	3:C:72:GLY:HA3	1.97	0.47
3:C:88:ILE:HB	3:C:202:PRO:HG2	1.96	0.46
3:C:46:ASN:ND2	3:C:48:GLU:O	2.50	0.45
2:B:247:LEU:HD22	2:B:251:GLN:OE1	2.17	0.45
3:C:154:SER:OG	3:C:155:LEU:N	2.48	0.45
2:B:103:ASP:OD1	2:B:104:ILE:N	2.49	0.45
1:A:184:ILE:HG13	1:A:185:GLY:N	2.32	0.45
3:C:98:SER:O	3:C:105:THR:HG22	2.17	0.45
3:C:139:CYS:HB3	3:C:213:LEU:HA	1.99	0.45
2:B:213:SER:HB2	3:C:226:MET:HE1	1.99	0.45
1:A:171:HIS:HE1	1:A:173:PHE:CE1	2.35	0.45
2:B:123:CYS:O	2:B:189:TYR:HB2	2.17	0.44
2:B:24:THR:OG1	2:B:25:MET:N	2.51	0.44
4:D:33:PRO:HB2	4:D:38:ILE:HD11	2.00	0.44
2:B:54:ASN:HB2	2:B:57:GLU:HB3	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:188:ARG:HH21	3:C:99:ILE:HD12	1.83	0.43
2:B:102:ASP:OD2	2:B:106:LYS:NZ	2.50	0.43
1:A:46:PRO:O	1:A:77:TYR:OH	2.36	0.43
2:B:8:THR:O	2:B:10:GLY:N	2.50	0.43
2:B:213:SER:HB2	3:C:226:MET:CE	2.49	0.43
3:C:65:LEU:HD12	3:C:66:ASP:N	2.32	0.43
3:C:94:SER:OG	3:C:95:LEU:N	2.51	0.43
2:B:200:ASN:O	2:B:202:ASN:OD1	2.37	0.43
1:A:80:ILE:HA	1:A:81:PRO:HD2	1.88	0.43
3:C:126:HIS:CE1	3:C:174:ASP:HB3	2.54	0.42
1:A:86:ASP:HA	1:A:116:ASN:HA	2.01	0.42
2:B:125:TYR:HB2	2:B:187:ILE:HB	2.00	0.42
1:A:120:ASN:HA	1:A:121:PRO:HD2	1.91	0.42
1:A:93:SER:OG	1:A:96:ASN:HB2	2.18	0.42
4:D:10:ARG:NH1	4:D:15:GLN:HG2	2.33	0.42
1:A:170:ASP:OD2	3:C:40:ILE:HG12	2.19	0.42
3:C:169:ASP:CG	3:C:170:GLU:N	2.73	0.42
3:C:146:LEU:HA	3:C:147:PRO:HD2	1.91	0.42
2:B:83:GLU:OE2	3:C:100:ILE:HG21	2.20	0.42
3:C:86:ILE:HB	3:C:205:ILE:HG23	2.01	0.42
3:C:126:HIS:HE1	3:C:174:ASP:HB3	1.85	0.41
3:C:139:CYS:SG	3:C:165:ILE:HG13	2.60	0.41
3:C:41:ASP:HA	3:C:42:PRO:HD2	1.81	0.41
2:B:179:ASP:OD1	2:B:180:SER:N	2.53	0.41
1:A:30:ARG:NH1	1:A:169:ASP:HA	2.36	0.41
3:C:104:ARG:HG3	3:C:104:ARG:O	2.21	0.41
3:C:137:HIS:HE1	3:C:220:ASN:O	2.04	0.41
1:A:153:ASN:OD1	1:A:154:LEU:N	2.54	0.41
1:A:179:CYS:HA	1:A:180:PRO:HD2	1.82	0.41
2:B:12:MET:N	2:B:13:PRO:HD3	2.35	0.41
1:A:115:GLN:HE21	1:A:120:ASN:CB	2.33	0.41
2:B:92:GLN:HG3	2:B:93:PRO:HD2	2.02	0.41
2:B:219:SER:HB2	2:B:222:GLU:OE2	2.21	0.41
3:C:63:TYR:CZ	3:C:229:LEU:HD11	2.56	0.41
3:C:59:ALA:O	3:C:230:THR:OG1	2.37	0.40
1:A:154:LEU:O	1:A:155:GLU:HB3	2.21	0.40
1:A:143:PHE:HA	1:A:144:PRO:HD2	1.90	0.40
2:B:8:THR:HG21	2:B:11:LEU:HD12	2.04	0.40
2:B:200:ASN:HB3	2:B:245:VAL:HG13	2.04	0.40
3:C:65:LEU:HD23	3:C:229:LEU:HD23	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	186/191 (97%)	170 (91%)	12 (6%)	4 (2%)	7 34
2	B	248/253 (98%)	235 (95%)	7 (3%)	6 (2%)	6 32
3	C	251/448 (56%)	217 (86%)	27 (11%)	7 (3%)	5 29
4	D	54/58 (93%)	50 (93%)	3 (6%)	1 (2%)	9 38
All	All	739/950 (78%)	672 (91%)	49 (7%)	18 (2%)	10 32

All (18) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	50	ILE
2	B	205	ASN
3	C	74	ASN
3	C	148	TYR
1	A	189	SER
2	B	149	ALA
3	C	73	GLY
2	B	25	MET
3	C	197	THR
4	D	17	ALA
1	A	9	SER
1	A	175	GLN
2	B	42	ASN
2	B	147	MET
3	C	26	PHE
3	C	35	PHE
3	C	72	GLY
2	B	9	GLY

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	169/169 (100%)	169 (100%)	0	100 100
2	B	227/230 (99%)	226 (100%)	1 (0%)	92 95
3	C	217/381 (57%)	217 (100%)	0	100 100
4	D	47/47 (100%)	47 (100%)	0	100 100
All	All	660/827 (80%)	659 (100%)	1 (0%)	94 97

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

Mol	Chain	Res	Type
2	B	102	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	58	GLN
1	A	102	ASN
2	B	85	GLN
2	B	162	HIS
2	B	177	GLN
2	B	240	HIS
3	C	38	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.