wwPDB EM Validation Summary Report (i)

May 14, 2024 – 05:49 pm BST

PDB ID	:	5MM2
EMDB ID	:	EMD-3528
Title	:	nora virus structure
Authors	:	Laurinmaki, P.; Shakeel, S.; Ekstrom, J.-O.; Butcher, S.J.
Deposited on	:	2016-12-08
Resolution	:	2.70 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB 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.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

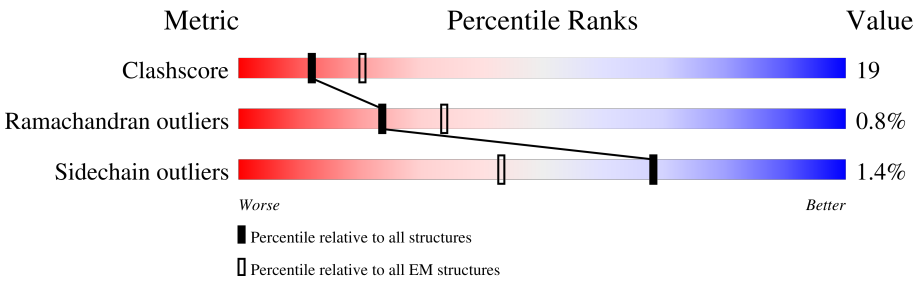
EMDB validation analysis	:	0.0.1.dev92
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 2.70 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	416	74% 13% 12%
2	B	251	69% 25% • •
3	C	264	75% 18% • 6%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6605 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 capsid protein VP4C.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	364	2772	1763	471	531	7	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	388	ILE	THR	conflict	UNP Q27YG7

- Molecule 2 is a protein called capsid protein VP4B.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	241	1900	1201	330	362	7	0	0

- Molecule 3 is a protein called Capsid protein VP4A.

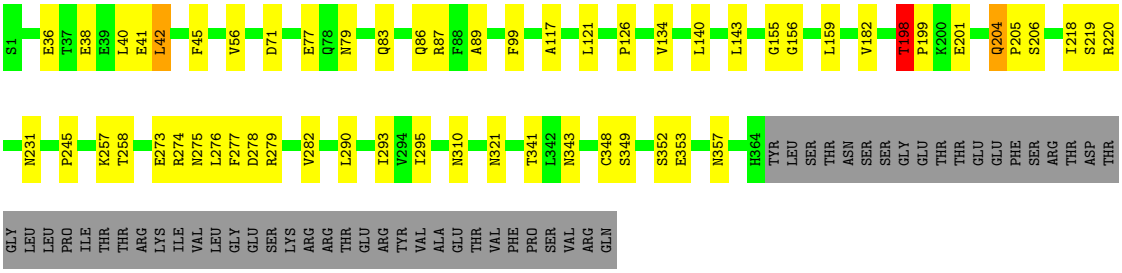
Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	249	1933	1237	324	365	7	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

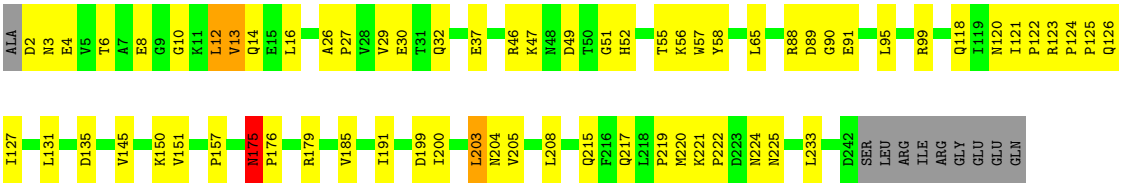
- Molecule 1: capsid protein VP4C

Chain A:



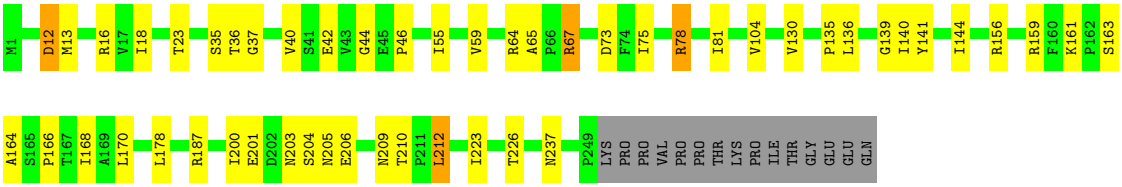
- Molecule 2: capsid protein VP4B

Chain B:



- Molecule 3: Capsid protein VP4A

Chain C:



4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	16131	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 ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	1500	Depositor
Magnification	47170	Depositor
Image detector	GATAN K2 QUANTUM (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	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.46	0/2827	0.71	3/3854 (0.1%)
2	B	0.53	1/1943 (0.1%)	0.71	3/2653 (0.1%)
3	C	0.55	1/1986 (0.1%)	0.64	3/2722 (0.1%)
All	All	0.51	2/6756 (0.0%)	0.69	9/9229 (0.1%)

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
1	A	0	3
2	B	0	1
3	C	0	1
All	All	0	5

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	175	ASN	C-N	-9.37	1.16	1.34
3	C	104	VAL	C-N	-6.56	1.19	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	12	LEU	CA-CB-CG	8.99	135.98	115.30
1	A	290	LEU	CA-CB-CG	7.71	133.03	115.30
3	C	209	ASN	C-N-CA	7.42	140.25	121.70
2	B	220	MET	C-N-CA	6.32	137.51	121.70
3	C	212	LEU	CA-CB-CG	6.05	129.22	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	156	GLY	Peptide
1	A	198	THR	Peptide
1	A	204	GLN	Peptide
2	B	179	ARG	Peptide
3	C	12	ASP	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	2772	0	2786	131	0
2	B	1900	0	1894	221	0
3	C	1933	0	1898	114	0
All	All	6605	0	6578	255	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 19.

The worst 5 of 255 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:40:LEU:HD11	2:B:99:ARG:NH1	1.26	1.44
1:A:40:LEU:CD1	2:B:99:ARG:NH1	1.95	1.29
1:A:41:GLU:HB2	2:B:46:ARG:NH2	1.44	1.28
1:A:77:GLU:N	2:B:224:ASN:OD1	1.67	1.27
2:B:200:ILE:HD13	3:C:204:SER:O	1.35	1.27

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	362/416 (87%)	309 (85%)	49 (14%)	4 (1%)	14	34
2	B	239/251 (95%)	202 (84%)	36 (15%)	1 (0%)	34	60
3	C	247/264 (94%)	227 (92%)	18 (7%)	2 (1%)	19	43
All	All	848/931 (91%)	738 (87%)	103 (12%)	7 (1%)	24	43

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	199	PRO
1	A	205	PRO
2	B	13	VAL
3	C	135	PRO
1	A	198	THR

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	306/354 (86%)	305 (100%)	1 (0%)	92	98
2	B	215/223 (96%)	209 (97%)	6 (3%)	43	73
3	C	214/228 (94%)	211 (99%)	3 (1%)	67	86
All	All	735/805 (91%)	725 (99%)	10 (1%)	68	86

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	67	ARG
3	C	78	ARG
3	C	237	ASN
2	B	175	ASN
2	B	203	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	175	ASN
2	B	225	ASN
3	C	237	ASN
3	C	205	ASN
3	C	222	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 monosaccharides 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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	C	1
2	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	104:VAL	C	105:ASP	N	1.19
1	B	175:ASN	C	176:PRO	N	1.16

6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-3528. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

This section was not generated.

6.2 Central slices (i)

This section was not generated.

6.3 Largest variance slices (i)

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color) (i)

This section was not generated.

6.5 Orthogonal surface views (i)

This section was not generated.

6.6 Mask visualisation (i)

This section was not generated. No masks/segmentation were deposited.

7 Map analysis (i)

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)

This section was not generated.

7.2 Volume estimate versus contour level (i)

This section was not generated.

7.3 Rotationally averaged power spectrum (i)

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit (i)

This section was not generated.