



# Full wwPDB/EMDatabank EM Map/Model Validation Report ⓘ

Jan 31, 2018 – 05:07 PM EST

PDB ID : 5O6A  
EMDB ID: : EMD-3752  
Title : The cryo-EM structure of Tick-borne encephalitis virus mature particle  
Authors : Fuzik, T.; Plevka, P.  
Deposited on : 2017-06-06  
Resolution : 3.90 Å(reported)  
Based on PDB ID : 2SVB

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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/EMValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

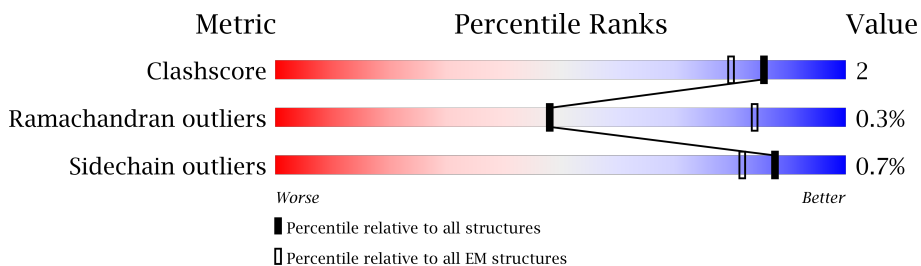
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.90 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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	A	496	92% 7% ..
1	B	496	92% 7% .
1	C	496	92% 7% .
2	D	75	87% 8% 5%
2	E	75	88% 7% 5%
2	F	75	89% 5% 5%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 12942 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.

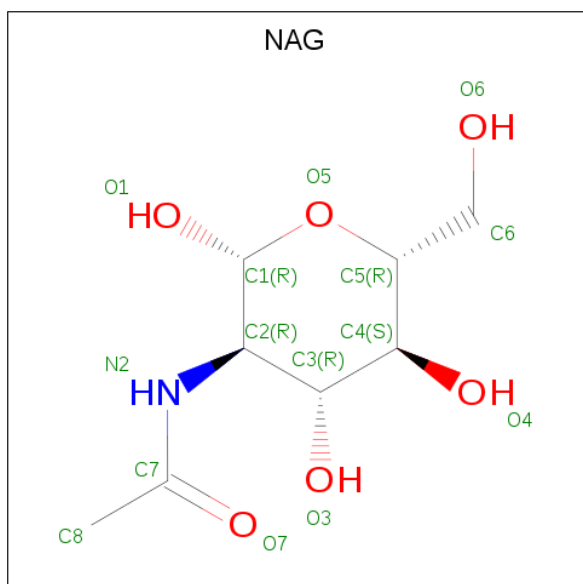
- Molecule 1 is a protein called Envelope protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	492	Total 3743	C 2371	N 654	O 693	S 25	0	0
1	B	492	Total 3743	C 2371	N 654	O 693	S 25	0	0
1	C	492	Total 3743	C 2371	N 654	O 693	S 25	0	0

- Molecule 2 is a protein called Small envelope protein M.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	D	71	Total 557	C 363	N 100	O 92	S 2	0	0
2	E	71	Total 557	C 363	N 100	O 92	S 2	0	0
2	F	71	Total 557	C 363	N 100	O 92	S 2	0	0

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>N<sub>1</sub>O<sub>6</sub>).

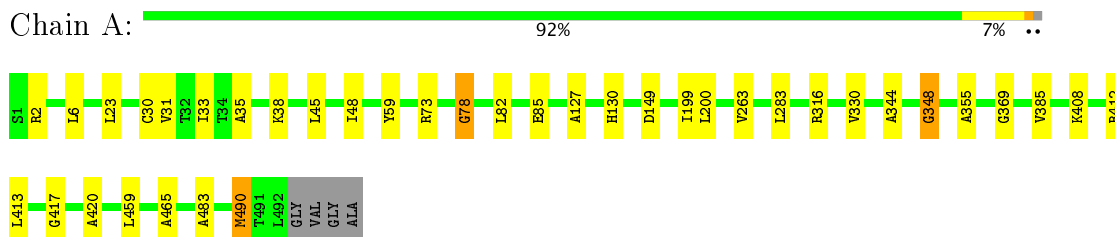


<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>				<b>AltConf</b>
3	A	1	Total 14	C 8	N 1	O 5	0
3	B	1	Total 14	C 8	N 1	O 5	0
3	C	1	Total 14	C 8	N 1	O 5	0

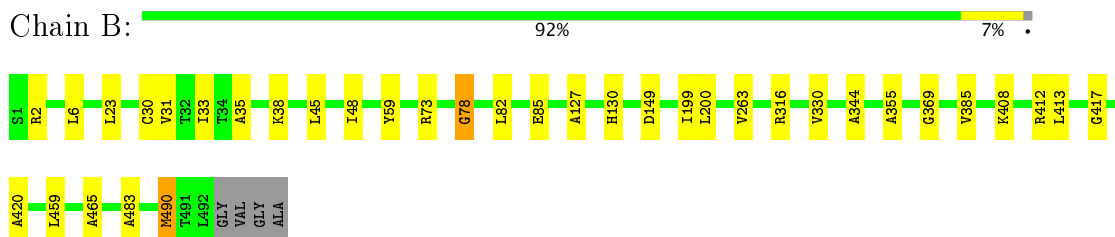
### 3 Residue-property plots [i](#)

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.

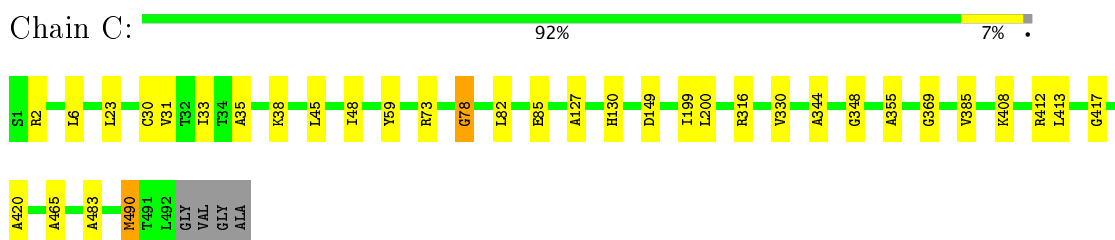
- Molecule 1: Envelope protein



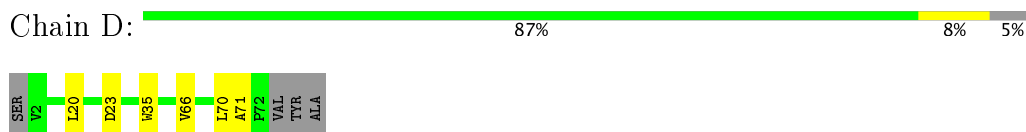
- Molecule 1: Envelope protein



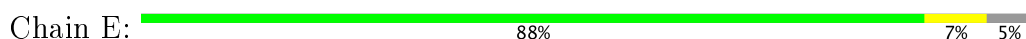
- Molecule 1: Envelope protein



- Molecule 2: Small envelope protein M




- Molecule 2: Small envelope protein M





- Molecule 2: Small envelope protein M

Chain F:  89% 5% 5%



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	11882	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 ( $e^-/\text{\AA}^2$ )	22	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	75000	Depositor
Image detector	FEI FALCON II (4k x 4k)	Depositor

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG

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	0.33	0/3826	0.53	0/5190
1	B	0.33	0/3826	0.53	0/5190
1	C	0.33	0/3826	0.53	0/5190
2	D	0.37	0/569	0.51	0/779
2	E	0.38	0/569	0.51	0/779
2	F	0.36	0/569	0.49	0/779
All	All	0.33	0/13185	0.53	0/17907

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	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	348	GLY	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	3743	0	3732	21	0
1	B	3743	0	3732	20	0
1	C	3743	0	3732	18	0
2	D	557	0	591	4	0
2	E	557	0	591	4	0
2	F	557	0	591	2	0
3	A	14	0	13	0	0
3	B	14	0	13	0	0
3	C	14	0	13	0	0
All	All	12942	0	13008	65	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 2.

All (65) 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:B:459:LEU:HD21	2:E:70:LEU:HD13	1.86	0.58
1:A:59:TYR:HB2	1:A:127:ALA:HB3	1.90	0.53
1:A:459:LEU:HD21	2:D:70:LEU:HD13	1.91	0.53
1:C:59:TYR:HB2	1:C:127:ALA:HB3	1.91	0.52
1:C:23:LEU:HD11	1:C:33:ILE:HD11	1.91	0.52
1:B:59:TYR:HB2	1:B:127:ALA:HB3	1.90	0.52
1:A:23:LEU:HD11	1:A:33:ILE:HD11	1.92	0.51
1:B:23:LEU:HD11	1:B:33:ILE:HD11	1.93	0.50
1:C:413:LEU:HA	1:C:417:GLY:HA3	1.93	0.50
1:A:413:LEU:HA	1:A:417:GLY:HA3	1.94	0.49
1:B:413:LEU:HA	1:B:417:GLY:HA3	1.93	0.49
1:B:23:LEU:HD22	1:B:31:VAL:HG11	1.96	0.48
1:B:6:LEU:HD21	1:B:30:CYS:HB2	1.94	0.48
1:A:6:LEU:HD21	1:A:30:CYS:HB2	1.95	0.48
1:C:23:LEU:HD22	1:C:31:VAL:HG11	1.96	0.47
1:C:6:LEU:HD21	1:C:30:CYS:HB2	1.95	0.47
2:F:20:LEU:HB3	2:F:23:ASP:HB2	1.97	0.46
2:F:35:TRP:NE1	2:F:71:ALA:O	2.48	0.46
1:B:490:MET:N	1:B:490:MET:SD	2.88	0.46
2:D:35:TRP:NE1	2:D:71:ALA:O	2.49	0.46
2:E:35:TRP:NE1	2:E:71:ALA:O	2.49	0.46
1:B:330:VAL:HG11	1:B:385:VAL:HG11	1.99	0.45
1:A:127:ALA:HB1	1:A:200:LEU:HD11	1.97	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:344:ALA:HB3	1:A:355:ALA:HB2	1.99	0.45
1:C:82:LEU:HB2	1:C:85:GLU:HG3	1.99	0.45
2:D:66:VAL:HG13	2:E:70:LEU:HD11	1.99	0.44
1:A:408:LYS:O	1:A:412:ARG:N	2.47	0.44
1:A:82:LEU:HB2	1:A:85:GLU:HG3	1.99	0.44
1:A:330:VAL:HG11	1:A:385:VAL:HG11	1.99	0.44
1:A:45:LEU:HD21	1:A:48:ILE:HD11	1.99	0.44
1:B:127:ALA:HB1	1:B:200:LEU:HD11	2.00	0.44
1:C:73:ARG:NH1	1:C:78:GLY:O	2.50	0.44
1:C:45:LEU:HD21	1:C:48:ILE:HD11	2.00	0.44
1:B:408:LYS:O	1:B:412:ARG:N	2.47	0.43
2:D:20:LEU:HB3	2:D:23:ASP:HB2	1.99	0.43
1:A:130:HIS:HB2	1:A:199:ILE:HB	2.00	0.43
1:B:149:ASP:OD1	1:B:369:GLY:N	2.51	0.43
1:C:408:LYS:O	1:C:412:ARG:N	2.47	0.43
1:A:490:MET:SD	1:A:490:MET:N	2.89	0.43
1:B:417:GLY:HA2	1:B:420:ALA:HB3	2.01	0.43
1:C:149:ASP:OD1	1:C:369:GLY:N	2.51	0.43
2:E:20:LEU:HB3	2:E:23:ASP:HB2	1.99	0.43
1:A:149:ASP:OD1	1:A:369:GLY:N	2.52	0.43
1:A:73:ARG:NH1	1:A:78:GLY:O	2.51	0.43
1:A:465:ALA:HB2	1:A:483:ALA:HB2	2.00	0.43
1:C:35:ALA:HB3	1:C:38:LYS:HB2	2.01	0.43
1:C:127:ALA:HB1	1:C:200:LEU:HD11	1.99	0.43
1:A:417:GLY:HA2	1:A:420:ALA:HB3	2.00	0.43
1:A:23:LEU:HD22	1:A:31:VAL:HG11	1.99	0.43
1:B:130:HIS:HB2	1:B:199:ILE:HB	2.01	0.43
1:B:73:ARG:NH1	1:B:78:GLY:O	2.52	0.43
1:C:344:ALA:HB3	1:C:355:ALA:HB2	2.01	0.43
1:A:35:ALA:HB3	1:A:38:LYS:HB2	2.01	0.42
1:B:45:LEU:HD21	1:B:48:ILE:HD11	2.02	0.42
1:B:82:LEU:HB2	1:B:85:GLU:HG3	2.00	0.42
1:A:199:ILE:HD11	1:A:283:LEU:HD21	2.02	0.42
1:C:130:HIS:HB2	1:C:199:ILE:HB	2.00	0.42
1:A:263:VAL:HG22	1:B:263:VAL:HG22	2.01	0.42
1:C:465:ALA:HB2	1:C:483:ALA:HB2	2.02	0.42
1:B:35:ALA:HB3	1:B:38:LYS:HB2	2.01	0.41
1:B:465:ALA:HB2	1:B:483:ALA:HB2	2.01	0.41
1:C:417:GLY:HA2	1:C:420:ALA:HB3	2.03	0.41
1:C:330:VAL:HG11	1:C:385:VAL:HG11	2.03	0.41
1:B:344:ALA:HB3	1:B:355:ALA:HB2	2.02	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:490:MET:SD	1:C:490:MET:N	2.88	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	490/496 (99%)	446 (91%)	42 (9%)	2 (0%)	38	76
1	B	490/496 (99%)	448 (91%)	41 (8%)	1 (0%)	51	84
1	C	490/496 (99%)	449 (92%)	39 (8%)	2 (0%)	38	76
2	D	69/75 (92%)	57 (83%)	12 (17%)	0	100	100
2	E	69/75 (92%)	57 (83%)	12 (17%)	0	100	100
2	F	69/75 (92%)	58 (84%)	11 (16%)	0	100	100
All	All	1677/1713 (98%)	1515 (90%)	157 (9%)	5 (0%)	48	80

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	348	GLY
1	A	78	GLY
1	C	78	GLY
1	C	348	GLY
1	B	78	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	400/401 (100%)	397 (99%)	3 (1%)	85	93
1	B	400/401 (100%)	397 (99%)	3 (1%)	85	93
1	C	400/401 (100%)	397 (99%)	3 (1%)	85	93
2	D	61/64 (95%)	61 (100%)	0	100	100
2	E	61/64 (95%)	61 (100%)	0	100	100
2	F	61/64 (95%)	61 (100%)	0	100	100
All	All	1383/1395 (99%)	1374 (99%)	9 (1%)	87	94

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	ARG
1	A	316	ARG
1	A	490	MET
1	B	2	ARG
1	B	316	ARG
1	B	490	MET
1	C	2	ARG
1	C	316	ARG
1	C	490	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	A	500	1	14,14,15	0.50	0	15,19,21	1.18	1 (6%)
3	NAG	B	500	1	14,14,15	0.28	0	15,19,21	0.74	0
3	NAG	C	500	1	14,14,15	0.28	0	15,19,21	0.77	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	A	500	1	-	0/6/23/26	0/1/1/1
3	NAG	B	500	1	-	0/6/23/26	0/1/1/1
3	NAG	C	500	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	500	NAG	C1-O5-C5	3.90	117.54	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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.