



wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:47 PM BST

PDB ID : 3J0C
EMDB ID: : EMD-5275
Title : Models of E1, E2 and CP of Venezuelan Equine Encephalitis Virus TC-83 strain restrained by a near atomic resolution cryo-EM map
Authors : Zhang, R.; Hryc, C.F.; Cong, Y.; Liu, X.; Jakana, J.; Gorchakov, R.; Baker, M.L.; Weaver, S.C.; Chiu, W.
Deposited on : 2011-06-22
Resolution : 4.80 Å(reported)
Based on PDB ID : 1EP5, 3N40

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>

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) : trunk27241

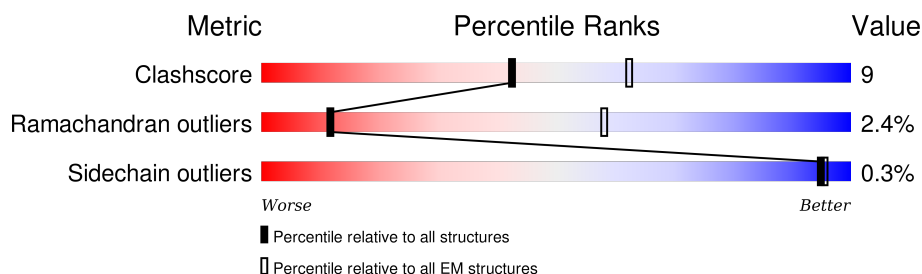
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.80 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.



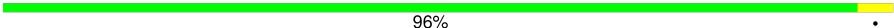
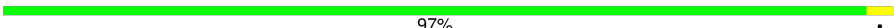
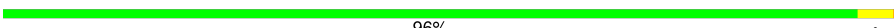
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	442	90% 7% .
1	D	442	90% 7% .
1	G	442	90% 7% .
1	J	442	91% 6% .
2	B	423	85% 12% .
2	E	423	87% 10% .
2	H	423	86% 10% .
2	K	423	85% 12% .
3	C	162	96% .

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	162	 96% .
3	I	162	 97% .
3	L	162	 96% . .

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 31804 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 E1 envelope glycoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	442	Total	C	N	O	S	0	0
			3370	2146	555	648	21		
1	D	442	Total	C	N	O	S	0	0
			3370	2146	555	648	21		
1	G	442	Total	C	N	O	S	0	0
			3370	2146	555	648	21		
1	J	442	Total	C	N	O	S	0	0
			3370	2146	555	648	21		

- Molecule 2 is a protein called E2 envelope glycoprotein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	423	Total	C	N	O	S	0	0
			3304	2086	586	607	25		
2	E	423	Total	C	N	O	S	0	0
			3304	2086	586	607	25		
2	H	423	Total	C	N	O	S	0	0
			3304	2086	586	607	25		
2	K	423	Total	C	N	O	S	0	0
			3304	2086	586	607	25		

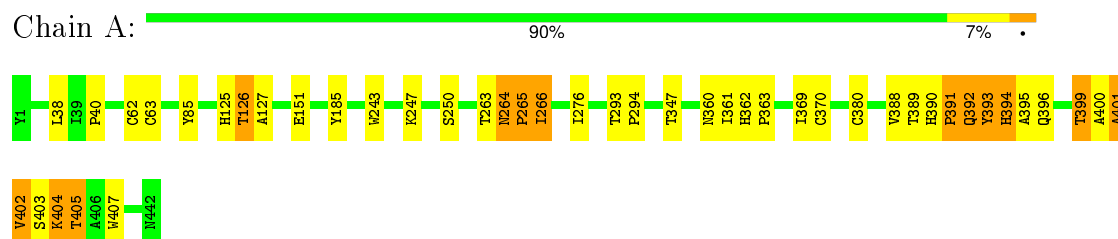
- Molecule 3 is a protein called Capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	162	Total	C	N	O	S	0	0
			1277	807	227	235	8		
3	F	162	Total	C	N	O	S	0	0
			1277	807	227	235	8		
3	I	162	Total	C	N	O	S	0	0
			1277	807	227	235	8		
3	L	162	Total	C	N	O	S	0	0
			1277	807	227	235	8		

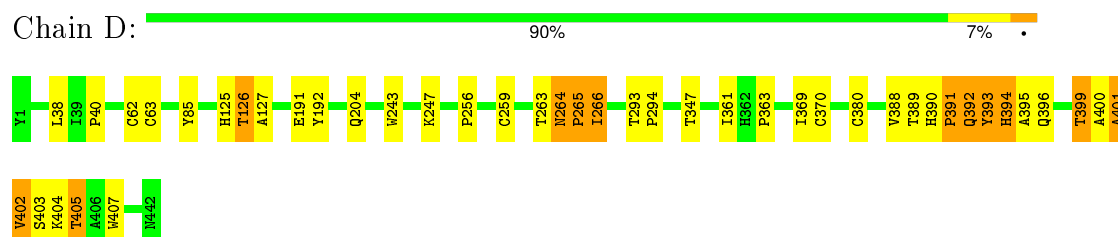
3 Residue-property plots

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.

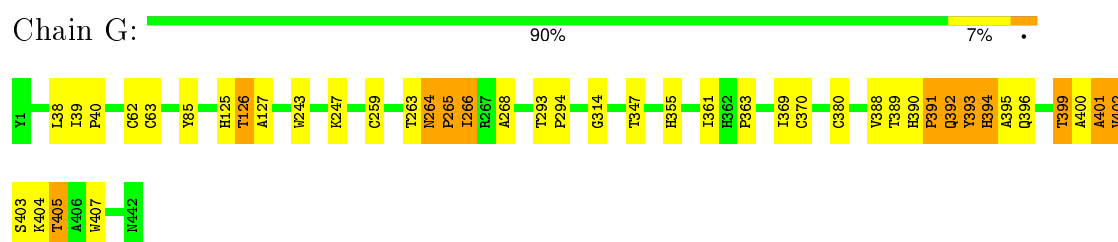
- Molecule 1: E1 envelope glycoprotein



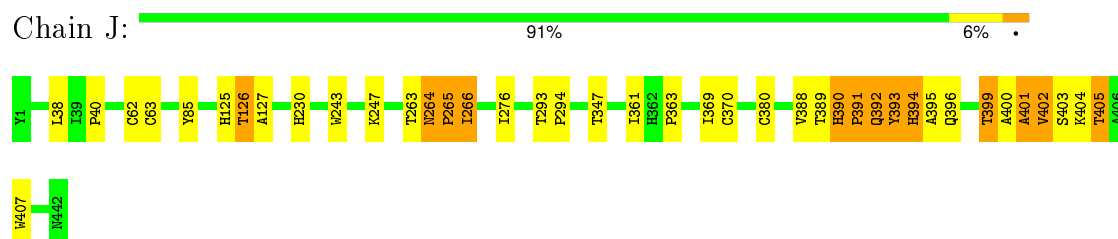
- Molecule 1: E1 envelope glycoprotein




- Molecule 1: E1 envelope glycoprotein

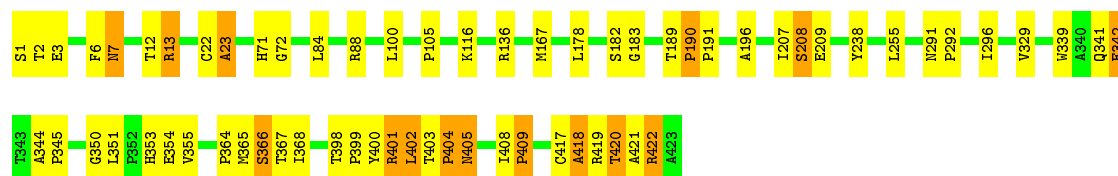


- Molecule 1: E1 envelope glycoprotein




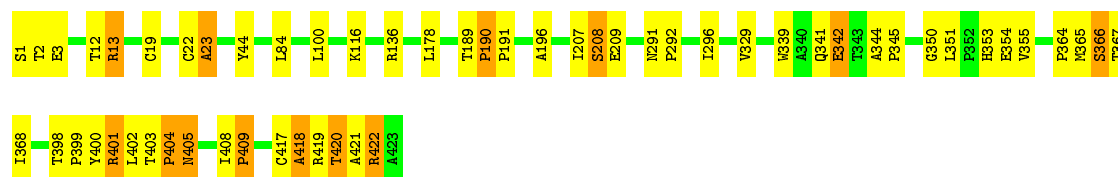
- Molecule 2: E2 envelope glycoprotein

Chain B:  85% 12%




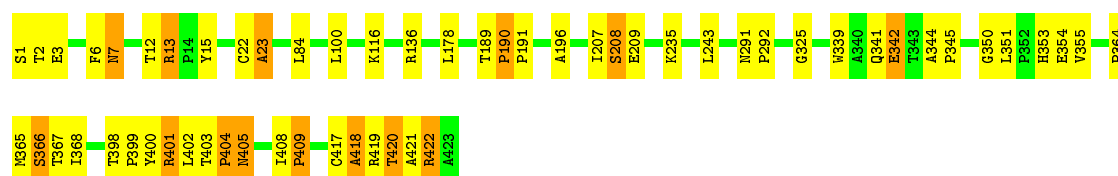
- Molecule 2: E2 envelope glycoprotein

Chain E:  87% 10%




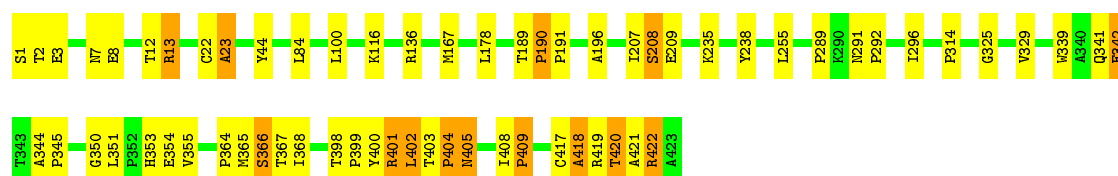
- Molecule 2: E2 envelope glycoprotein

Chain H:  86% 10%



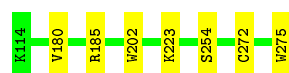
- Molecule 2: E2 envelope glycoprotein

Chain K:  85% 12%



- Molecule 3: Capsid protein

Chain C:  96%



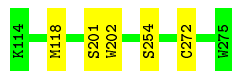
- Molecule 3: Capsid protein

Chain F:  96%



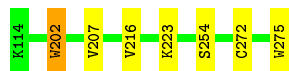
- Molecule 3: Capsid protein

Chain I:  97% .



- Molecule 3: Capsid protein

Chain L:  96% . .



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, 1	Depositor
Number of images	37000	Depositor
Resolution determination method	FSC at 0.5 cut-off	Depositor
CTF correction method	CTF parameters were determined from particles within each CCD image	Depositor
Microscope	JEOL 3200FSC	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	18	Depositor
Minimum defocus (nm)	400	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	100000	Depositor
Image detector	Gatan 4kx4k CCD Camera	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	0.58	1/3461 (0.0%)	0.50	0/4731
1	D	0.58	1/3461 (0.0%)	0.50	0/4731
1	G	0.58	1/3461 (0.0%)	0.50	0/4731
1	J	0.58	1/3461 (0.0%)	0.50	0/4731
2	B	0.64	1/3395 (0.0%)	0.58	0/4616
2	E	0.64	1/3395 (0.0%)	0.58	0/4616
2	H	0.64	1/3395 (0.0%)	0.58	0/4616
2	K	0.64	1/3395 (0.0%)	0.58	0/4616
3	C	0.64	0/1305	0.51	0/1757
3	F	0.64	0/1305	0.50	0/1757
3	I	0.64	0/1305	0.50	0/1757
3	L	0.64	0/1305	0.51	0/1757
All	All	0.62	8/32644 (0.0%)	0.54	0/44416

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	5
1	D	0	5
1	G	0	5
1	J	0	5
2	B	0	1
2	K	0	1
All	All	0	22

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	391	PRO	N-CD	6.14	1.56	1.47
2	B	409	PRO	N-CD	6.10	1.56	1.47

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	391	PRO	N-CD	6.08	1.56	1.47
1	A	391	PRO	N-CD	5.99	1.56	1.47
2	K	409	PRO	N-CD	5.99	1.56	1.47

There are no bond angle outliers.

There are no chirality outliers.

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	391	PRO	Peptide
1	A	392	GLN	Peptide
1	A	393	TYR	Peptide
1	A	394	HIS	Peptide
1	A	399	THR	Peptide

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	3370	0	3274	63	0
1	D	3370	0	3274	62	0
1	G	3370	0	3274	63	0
1	J	3370	0	3274	60	0
2	B	3304	0	3244	76	0
2	E	3304	0	3244	71	0
2	H	3304	0	3244	73	0
2	K	3304	0	3244	75	0
3	C	1277	0	1272	3	0
3	F	1277	0	1272	3	0
3	I	1277	0	1272	2	0
3	L	1277	0	1272	4	0
All	All	31804	0	31160	550	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 9.

The worst 5 of 550 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:190:PRO:HB2	2:B:191:PRO:C	1.38	1.45
2:K:190:PRO:HB2	2:K:191:PRO:C	1.40	1.40
2:E:190:PRO:HB2	2:E:191:PRO:C	1.39	1.38
2:H:190:PRO:HB2	2:H:191:PRO:C	1.37	1.38
2:K:190:PRO:HB2	2:K:191:PRO:CA	1.57	1.31

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	440/442 (100%)	412 (94%)	19 (4%)	9 (2%)	9	52
1	D	440/442 (100%)	412 (94%)	20 (4%)	8 (2%)	11	54
1	G	440/442 (100%)	415 (94%)	17 (4%)	8 (2%)	11	54
1	J	440/442 (100%)	412 (94%)	19 (4%)	9 (2%)	9	52
2	B	421/423 (100%)	382 (91%)	23 (6%)	16 (4%)	4	38
2	E	421/423 (100%)	381 (90%)	25 (6%)	15 (4%)	4	40
2	H	421/423 (100%)	383 (91%)	22 (5%)	16 (4%)	4	38
2	K	421/423 (100%)	381 (90%)	25 (6%)	15 (4%)	4	40
3	C	160/162 (99%)	154 (96%)	6 (4%)	0	100	100
3	F	160/162 (99%)	155 (97%)	5 (3%)	0	100	100
3	I	160/162 (99%)	156 (98%)	4 (2%)	0	100	100
3	L	160/162 (99%)	155 (97%)	5 (3%)	0	100	100
All	All	4084/4108 (99%)	3798 (93%)	190 (5%)	96 (2%)	12	48

5 of 96 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	264	ASN
1	A	265	PRO
1	A	266	ILE
1	A	402	VAL
2	B	13	ARG

5.3.2 Protein sidechains ⓘ

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	368/368 (100%)	367 (100%)	1 (0%)	94	96
1	D	368/368 (100%)	368 (100%)	0	100	100
1	G	368/368 (100%)	368 (100%)	0	100	100
1	J	368/368 (100%)	367 (100%)	1 (0%)	94	96
2	B	364/364 (100%)	363 (100%)	1 (0%)	94	96
2	E	364/364 (100%)	363 (100%)	1 (0%)	94	96
2	H	364/364 (100%)	363 (100%)	1 (0%)	94	96
2	K	364/364 (100%)	363 (100%)	1 (0%)	94	96
3	C	135/135 (100%)	134 (99%)	1 (1%)	88	94
3	F	135/135 (100%)	134 (99%)	1 (1%)	88	94
3	I	135/135 (100%)	134 (99%)	1 (1%)	88	94
3	L	135/135 (100%)	134 (99%)	1 (1%)	88	94
All	All	3468/3468 (100%)	3458 (100%)	10 (0%)	95	96

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

Mol	Chain	Res	Type
3	F	202	TRP
2	H	136	ARG
1	J	230	HIS
2	E	136	ARG
3	I	202	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	442	ASN
1	G	222	GLN
1	G	442	ASN
1	J	235	GLN
1	J	439	GLN

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.