



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 01:44 PM BST

PDB ID : 2XFB  
EMDB ID: : EMD-1121  
Title : CHIKUNGUNYA E1 E2 ENVELOPE GLYCOPROTEINS FITTED IN  
SINDBIS VIRUS cryo- EM MAP  
Authors : Voss, J.E.; Vaney, M.C.; Duquerroy, S.; Rey, F.A.  
Deposited on : 2010-05-21  
Resolution : 9.00 Å(reported)  
Based on PDB ID : 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](mailto: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

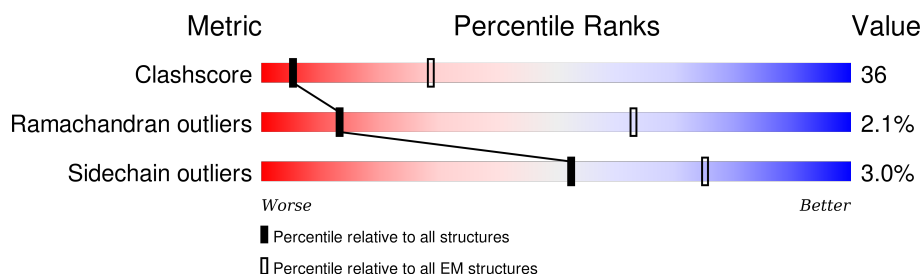
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 9.00 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  $\geq 3$ , 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	391	78% 19% .
1	D	391	79% 19% .
1	F	391	80% 17% .
1	H	391	75% 21% .
2	B	334	85% 11% ..
2	E	334	82% 14% ..
2	G	334	83% 14% ..
2	I	334	81% 16% ..

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 22460 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	391	Total	C	N	O	S	0	0
			2981	1886	499	572	24		
1	D	391	Total	C	N	O	S	0	0
			2981	1886	499	572	24		
1	F	391	Total	C	N	O	S	0	0
			2981	1886	499	572	24		
1	H	391	Total	C	N	O	S	0	0
			2981	1886	499	572	24		

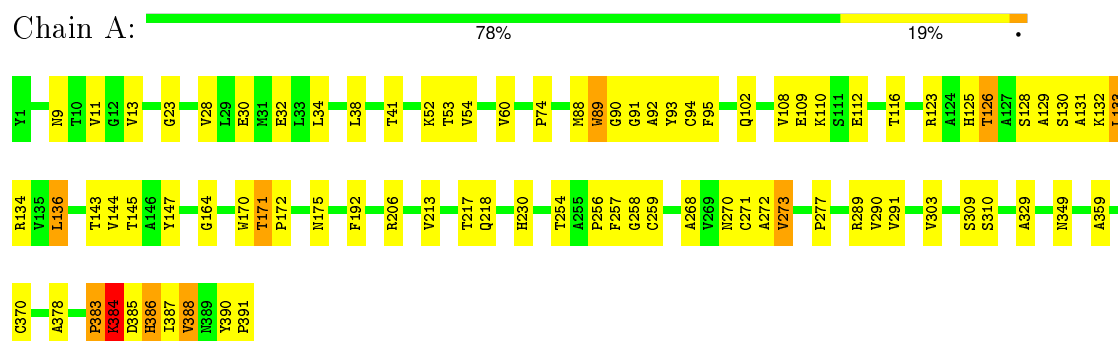
- Molecule 2 is a protein called E2 ENVELOPE GLYCOPROTEIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	334	Total	C	N	O	S	0	0
			2634	1643	477	494	20		
2	E	334	Total	C	N	O	S	0	0
			2634	1643	477	494	20		
2	G	334	Total	C	N	O	S	0	0
			2634	1643	477	494	20		
2	I	334	Total	C	N	O	S	0	0
			2634	1643	477	494	20		

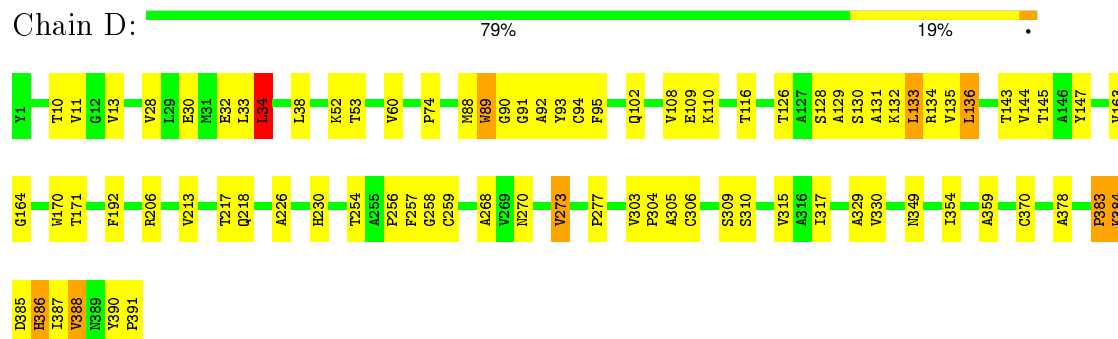
### 3 Residue-property plots

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.

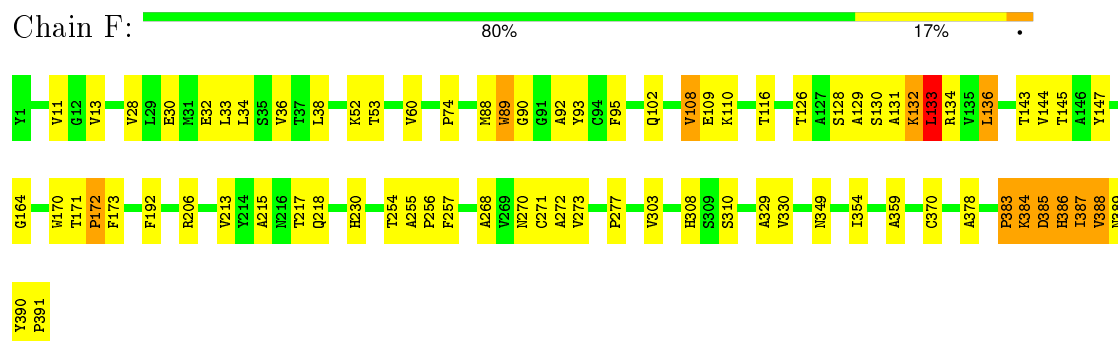
#### • Molecule 1: E1 ENVELOPE GLYCOPROTEIN



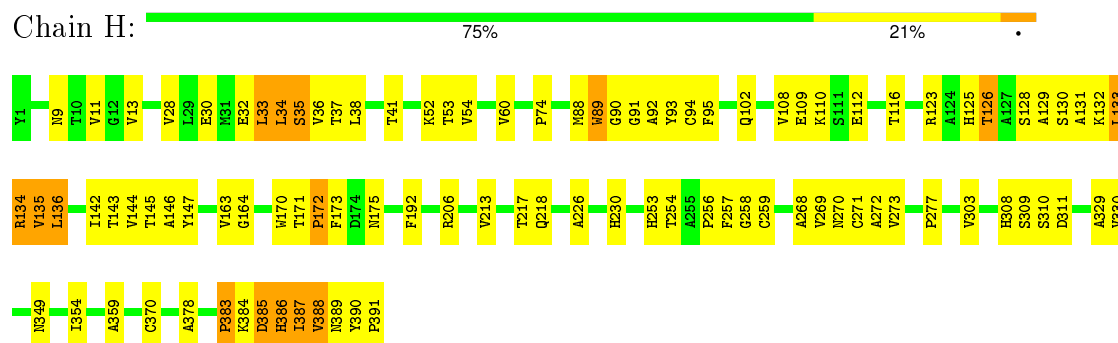
#### • Molecule 1: E1 ENVELOPE GLYCOPROTEIN



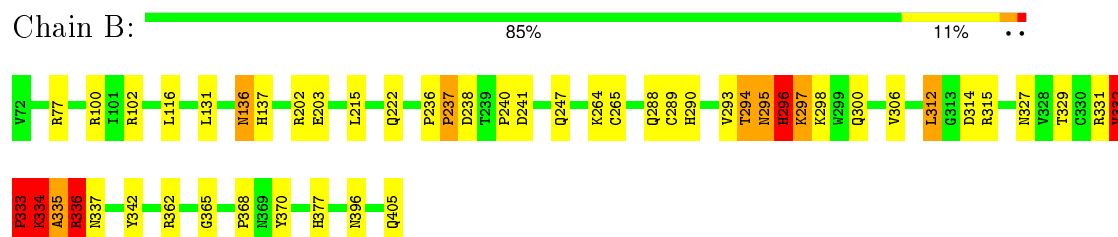
#### • Molecule 1: E1 ENVELOPE GLYCOPROTEIN



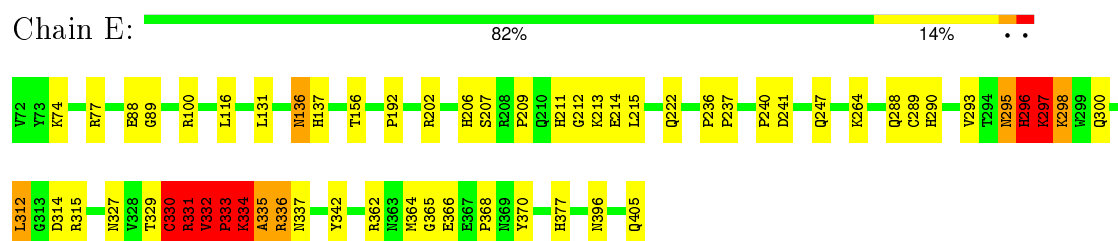
#### • Molecule 1: E1 ENVELOPE GLYCOPROTEIN



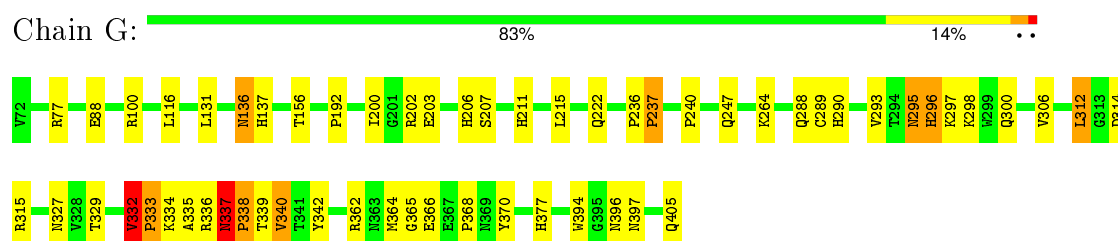
• Molecule 2: E2 ENVELOPE GLYCOPROTEIN



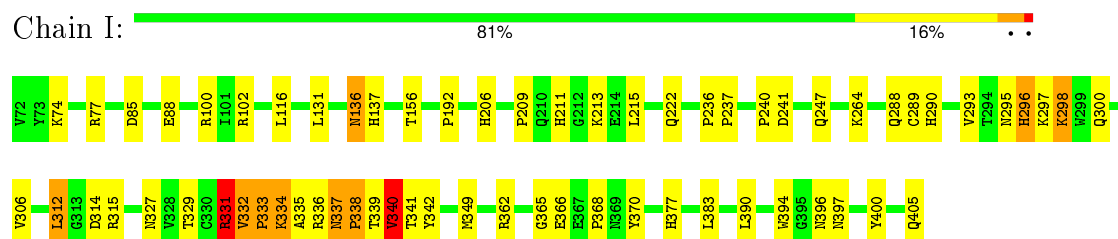
• Molecule 2: E2 ENVELOPE GLYCOPROTEIN



• Molecule 2: E2 ENVELOPE GLYCOPROTEIN



• Molecule 2: E2 ENVELOPE GLYCOPROTEIN



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Not provided	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	100	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	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  > 2$	RMSZ	# $ Z  > 2$
1	A	0.56	1/3053 (0.0%)	0.77	4/4157 (0.1%)
1	D	0.37	1/3054 (0.0%)	0.88	4/4160 (0.1%)
1	F	0.62	1/3053 (0.0%)	0.59	1/4157 (0.0%)
1	H	0.34	0/3054	0.57	0/4160
2	B	0.57	9/2705 (0.3%)	0.69	7/3682 (0.2%)
2	E	0.48	2/2704 (0.1%)	0.66	7/3678 (0.2%)
2	G	0.38	0/2705	0.63	2/3682 (0.1%)
2	I	0.35	0/2704	0.60	2/3678 (0.1%)
All	All	0.47	14/23032 (0.1%)	0.68	27/31354 (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	2
1	D	0	2
2	E	0	1
All	All	0	5

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	108	VAL	C-N	28.82	2.00	1.34
1	A	171	THR	C-N	-23.03	0.90	1.34
2	B	333	PRO	CA-C	8.69	1.70	1.52
2	B	336	ARG	N-CA	7.42	1.61	1.46
2	B	335	ALA	N-CA	7.36	1.61	1.46

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	171	THR	O-C-N	27.45	173.25	121.10
1	D	171	THR	C-N-CD	23.30	177.34	128.40
1	A	171	THR	O-C-N	19.99	159.09	121.10
1	D	171	THR	CA-C-N	-17.56	67.94	117.10
1	A	171	THR	C-N-CD	17.28	164.68	128.40

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	383	PRO	Mainchain,Peptide
1	D	383	PRO	Mainchain,Peptide
2	E	330	CYS	Mainchain

## 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	2981	0	2877	385	0
1	D	2981	0	2879	308	0
1	F	2981	0	2877	298	0
1	H	2981	0	2880	374	0
2	B	2634	0	2556	176	0
2	E	2634	0	2554	168	0
2	G	2634	0	2553	179	0
2	I	2634	0	2550	255	0
All	All	22460	0	21726	1595	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 36.

The worst 5 of 1595 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:125:HIS:CE1	1:H:41:THR:HB	1.27	1.67
1:A:41:THR:HB	1:H:125:HIS:CE1	1.26	1.67

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:132:LYS:CB	1:D:145:THR:HA	1.25	1.64
1:H:257:PHE:CZ	2:I:366:GLU:HA	1.14	1.63
1:F:256:PRO:HG3	2:G:368:PRO:CG	1.28	1.63

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	379/391 (97%)	361 (95%)	13 (3%)	5 (1%)	15	60
1	D	381/391 (97%)	363 (95%)	12 (3%)	6 (2%)	12	56
1	F	379/391 (97%)	362 (96%)	9 (2%)	8 (2%)	9	50
1	H	381/391 (97%)	358 (94%)	13 (3%)	10 (3%)	7	45
2	B	332/334 (99%)	304 (92%)	21 (6%)	7 (2%)	9	50
2	E	330/334 (99%)	305 (92%)	18 (6%)	7 (2%)	9	50
2	G	332/334 (99%)	302 (91%)	21 (6%)	9 (3%)	6	45
2	I	330/334 (99%)	300 (91%)	23 (7%)	7 (2%)	9	50
All	All	2844/2900 (98%)	2655 (93%)	130 (5%)	59 (2%)	13	50

5 of 59 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	387	ILE
2	B	296	HIS
2	B	332	VAL
2	B	333	PRO
1	D	387	ILE

### 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	328/328 (100%)	318 (97%)	10 (3%)	48	77
1	D	328/328 (100%)	318 (97%)	10 (3%)	48	77
1	F	328/328 (100%)	319 (97%)	9 (3%)	52	79
1	H	328/328 (100%)	319 (97%)	9 (3%)	52	79
2	B	296/296 (100%)	286 (97%)	10 (3%)	44	75
2	E	296/296 (100%)	284 (96%)	12 (4%)	37	71
2	G	296/296 (100%)	290 (98%)	6 (2%)	63	85
2	I	296/296 (100%)	286 (97%)	10 (3%)	44	75
All	All	2496/2496 (100%)	2420 (97%)	76 (3%)	52	77

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	312	LEU
1	F	89	TRP
2	I	314	ASP
2	E	314	ASP
2	E	332	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	222	GLN
2	G	296	HIS
2	I	296	HIS
1	F	373	GLN
2	G	300	GLN

### 5.3.3 RNA ⓘ

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.