wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Mar 2, 2017 – 11:27 am GMT

PDB ID	:	3J2P
EMDB ID:	:	EMD-5499
Title	:	CryoEM structure of Dengue virus envelope protein heterotetramer
Authors	:	Zhang, X.; Ge, P.; Yu, X.; Brannan, J.M.; Bi, G.; Zhang, Q.; Schein, S.; Zhou, Z.H.
Deposited on	:	2012-11-30
Resolution	:	3.60 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp
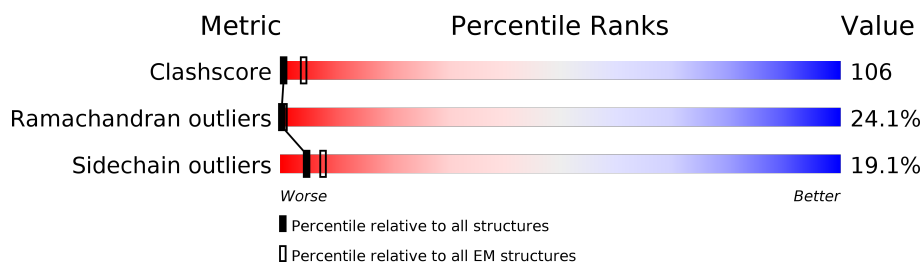with specific help available everywhere you see the ⓘ symbol.

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 3.60 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	Percentile Ranks	Value
Clashscore		106
Ramachandran outliers		24.1%
Sidechain outliers		19.1%

Worse   *Better*

▌ Percentile relative to all structures
◻ Percentile relative to all EM structures

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 >=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	495	13% 55% 28% •
1	C	495	12% 56% 28% •
2	B	75	17% 53% 23% • •
2	D	75	13% 57% 23% • •

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 9890 atoms, of which 1086 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 E.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	495	Total	C	H	N	O	S	0	0
			4276	2410	475	648	711	32		
1	C	495	Total	C	H	N	O	S	0	0
			4276	2410	475	648	711	32		

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

Mol	Chain	Residues	Atoms						AltConf	Trace
2	B	72	Total	C	H	N	O	S	0	0
			627	366	68	95	95	3		
2	D	72	Total	C	H	N	O	S	0	0
			627	366	68	95	95	3		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	15	ALA	ARG	SEE REMARK 999	UNP P14340
D	15	ALA	ARG	SEE REMARK 999	UNP P14340

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).
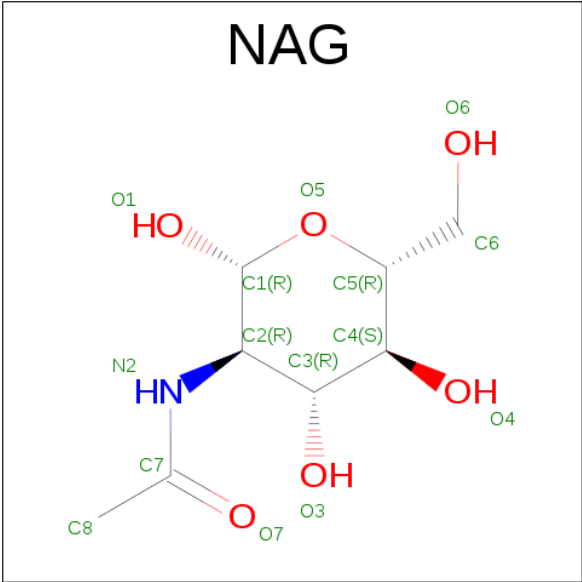
Mol	Chain	Residues	Atoms				AltConf
3	A	1	Total	C	N	O	0
			42	24	3	15	
3	A	1	Total	C	N	O	0
			42	24	3	15	
3	A	1	Total	C	N	O	0
			42	24	3	15	
3	C	1	Total	C	N	O	0
			42	24	3	15	
3	C	1	Total	C	N	O	0
			42	24	3	15	
3	C	1	Total	C	N	O	0
			42	24	3	15	

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 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 E
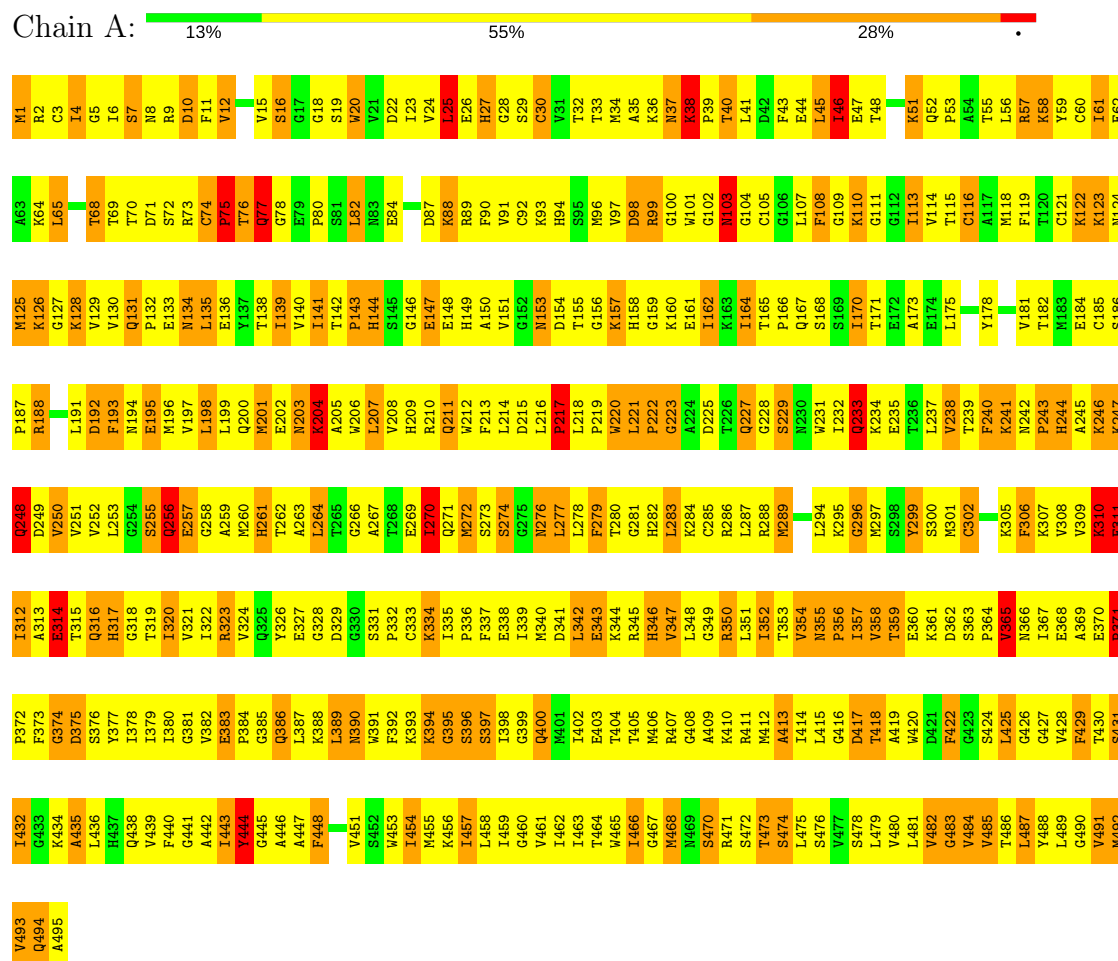
- Molecule 1: Envelope protein E
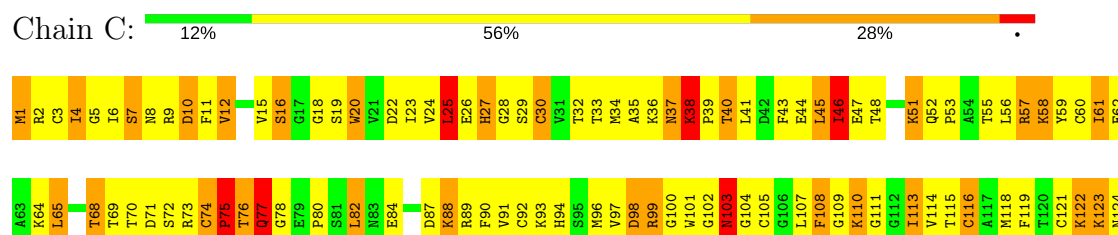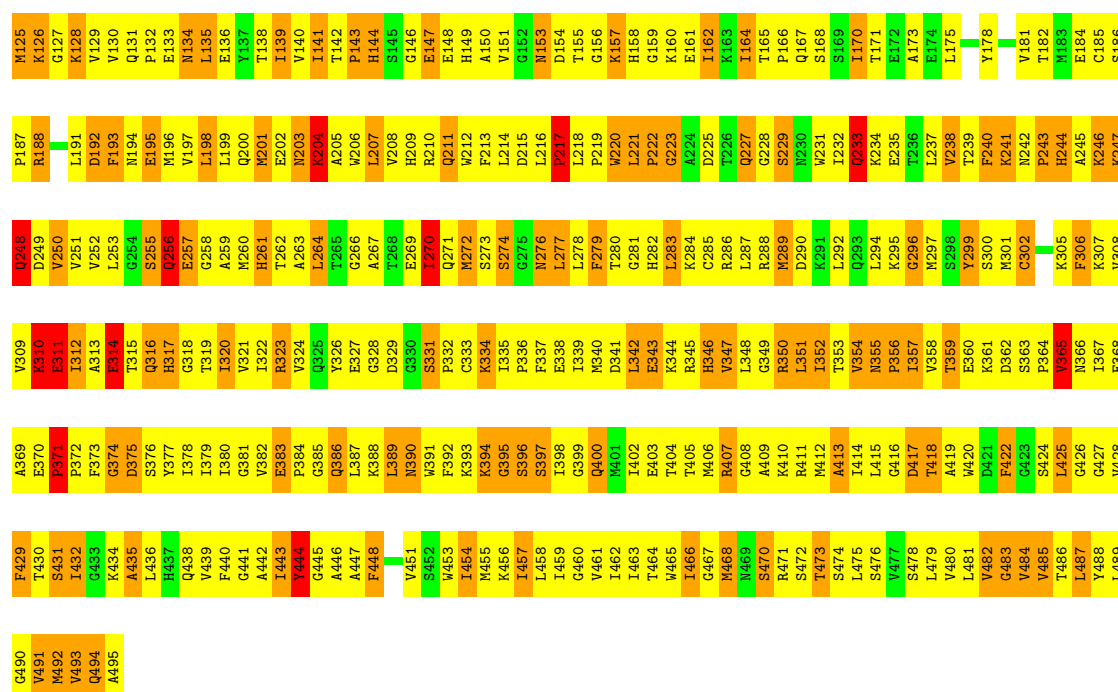
• Molecule 2: Small envelope protein M

Chain B:
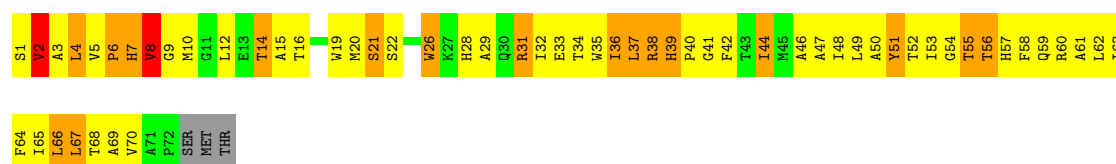
• Molecule 2: Small envelope protein M

Chain D:
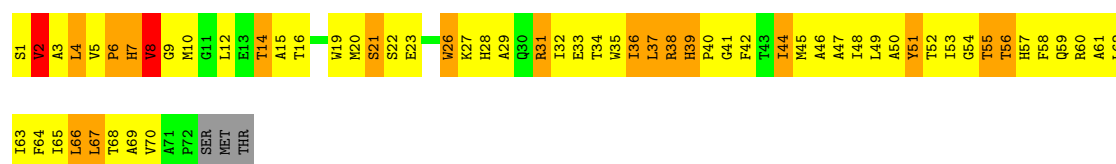
4 Experimental information ⓘ

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	9288	Depositor
Resolution determination method	FSC at 0.143 cut-off	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose $(e^-/\text{Å}^2)$	25	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	2400	Depositor
Magnification	57518	Depositor
Image detector	Kodak SO163 film	Depositor

5   Model quality ⓘ

5.1   Standard geometry ⓘ

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	0.46	0/3876	0.82	1/5237 (0.0%)
1	C	0.46	0/3876	0.82	1/5237 (0.0%)
2	B	0.49	0/575	0.77	0/786
2	D	0.50	0/575	0.77	0/786
All	All	0.46	0/8902	0.82	2/12046 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	25	LEU	CA-CB-CG	-5.64	102.33	115.30
1	C	25	LEU	CA-CB-CG	-5.56	102.51	115.30

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	3801	475	3844	838	0
1	C	3801	475	3844	820	0
2	B	559	68	569	117	0
2	D	559	68	569	119	0
3	A	42	0	38	3	0

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK
EMDataBank
Unified Data Resource for 3DEM

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	42	0	38	3	0
All	All	8804	1086	8902	1831	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 106.

The worst 5 of 1831 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:D:51:TYR:HA	2:D:60:ARG:HD3	1.18	1.17
2:B:51:TYR:HA	2:B:60:ARG:HD3	1.17	1.14
2:B:55:THR:HA	2:B:60:ARG:HH11	1.01	1.14
2:B:55:THR:HA	2:B:60:ARG:NH1	1.65	1.10
2:D:55:THR:HA	2:D:60:ARG:HH11	1.01	1.08

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	493/495 (100%)	228 (46%)	144 (29%)	121 (24%)	0	1
1	C	493/495 (100%)	227 (46%)	144 (29%)	122 (25%)	0	1
2	B	70/75 (93%)	33 (47%)	23 (33%)	14 (20%)	0	2
2	D	70/75 (93%)	33 (47%)	23 (33%)	14 (20%)	0	2
All	All	1126/1140 (99%)	521 (46%)	334 (30%)	271 (24%)	0	1

5 of 271 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	16	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	30	CYS
1	A	57	ARG
1	A	74	CYS
1	A	77	GLN

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	420/420 (100%)	340 (81%)	80 (19%)	2	11
1	C	420/420 (100%)	340 (81%)	80 (19%)	2	11
2	B	57/60 (95%)	46 (81%)	11 (19%)	1	11
2	D	57/60 (95%)	46 (81%)	11 (19%)	1	11
All	All	954/960 (99%)	772 (81%)	182 (19%)	5	11

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	6	PRO
1	C	65	LEU
1	C	457	ILE
2	B	14	THR
1	C	20	TRP

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

Mol	Chain	Res	Type
2	B	7	HIS
1	C	83	ASN
1	C	390	ASN
1	C	27	HIS
1	C	131	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

6 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	501	1	14,14,15	0.51	0	15,19,21	0.89	1 (6%)
3	NAG	A	502	1,3	14,14,15	0.56	0	15,19,21	0.83	0
3	NAG	A	503	3	14,14,15	0.58	0	15,19,21	0.85	1 (6%)
3	NAG	C	501	1	14,14,15	0.50	0	15,19,21	0.90	1 (6%)
3	NAG	C	502	1,3	14,14,15	0.57	0	15,19,21	0.82	0
3	NAG	C	503	3	14,14,15	0.57	0	15,19,21	0.84	1 (6%)

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	501	1	-	0/6/23/26	0/1/1/1
3	NAG	A	502	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	503	3	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	C	501	1	-	0/6/23/26	0/1/1/1
3	NAG	C	502	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	503	3	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	501	NAG	C2-N2-C7	-2.59	119.17	122.94
3	A	501	NAG	C2-N2-C7	-2.54	119.23	122.94
3	A	503	NAG	C2-N2-C7	-2.11	119.86	122.94
3	C	503	NAG	C2-N2-C7	-2.09	119.90	122.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	NAG	3	0
3	C	502	NAG	3	0

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.