wwPDB EM Validation Summary Report (i)

Dec 13, 2022 – 12:04 AM EST

PDB ID	:	1K4R
Title	:	Structure of Dengue Virus
Authors	:	Kuhn, R.J.; Zhang, W.; Rossmann, M.G.; Pletnev, S.V.; Corver, J.; Lenches, E.; Jones, C.T.; Mukhopadhyay, S.; Chipman, P.R.; Strauss, E.G.; Baker, T.S.; Strauss, J.H.
Deposited on	:	2001-10-08
Resolution	:	24.00 Å(reported)
Based on initial model	:	1SVB

This is a wwPDB EM 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
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:

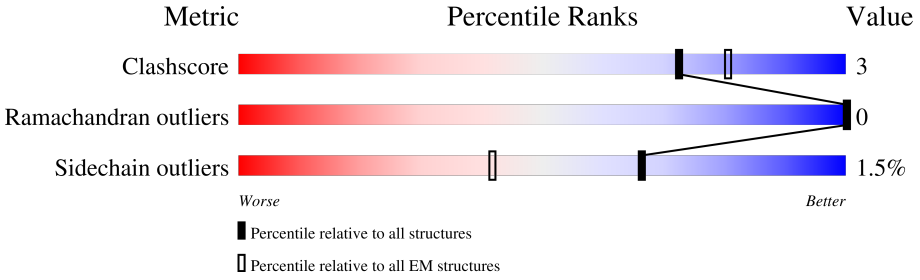
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 24.00 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	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	395	91% 9%
1	B	395	91% 9%
1	C	395	90% 9%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9084 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 MAJOR ENVELOPE PROTEIN E.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	395	3028	1902	533	572	21	0	0
1	B	395	3028	1902	533	572	21	0	0
1	C	395	3028	1902	533	572	21	0	0

There are 39 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	ALA	SER	CONFLICT	UNP C3V005
A	88	GLY	SER	CONFLICT	UNP C3V005
A	115	ALA	THR	CONFLICT	UNP C3V005
A	120	ALA	SER	CONFLICT	UNP C3V005
A	167	ILE	VAL	CONFLICT	UNP C3V005
A	171	LYS	ARG	CONFLICT	UNP C3V005
A	178	GLU	ASP	CONFLICT	UNP C3V005
A	206	VAL	SER	CONFLICT	UNP C3V005
A	267	ALA	SER	CONFLICT	UNP C3V005
A	277	GLU	ASP	CONFLICT	UNP C3V005
A	317	ALA	ILE	CONFLICT	UNP C3V005
A	331	THR	ALA	CONFLICT	UNP C3V005
A	358	ILE	MET	CONFLICT	UNP C3V005
B	47	ALA	SER	CONFLICT	UNP C3V005
B	88	GLY	SER	CONFLICT	UNP C3V005
B	115	ALA	THR	CONFLICT	UNP C3V005
B	120	ALA	SER	CONFLICT	UNP C3V005
B	167	ILE	VAL	CONFLICT	UNP C3V005
B	171	LYS	ARG	CONFLICT	UNP C3V005
B	178	GLU	ASP	CONFLICT	UNP C3V005
B	206	VAL	SER	CONFLICT	UNP C3V005
B	267	ALA	SER	CONFLICT	UNP C3V005
B	277	GLU	ASP	CONFLICT	UNP C3V005
B	317	ALA	ILE	CONFLICT	UNP C3V005

Continued on next page...

Continued from previous page...

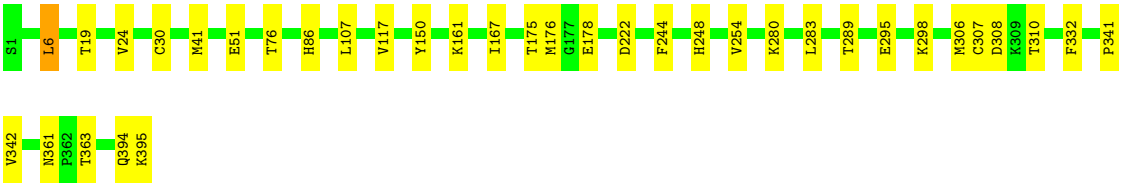
Chain	Residue	Modelled	Actual	Comment	Reference
B	331	THR	ALA	CONFLICT	UNP C3V005
B	358	ILE	MET	CONFLICT	UNP C3V005
C	47	ALA	SER	CONFLICT	UNP C3V005
C	88	GLY	SER	CONFLICT	UNP C3V005
C	115	ALA	THR	CONFLICT	UNP C3V005
C	120	ALA	SER	CONFLICT	UNP C3V005
C	167	ILE	VAL	CONFLICT	UNP C3V005
C	171	LYS	ARG	CONFLICT	UNP C3V005
C	178	GLU	ASP	CONFLICT	UNP C3V005
C	206	VAL	SER	CONFLICT	UNP C3V005
C	267	ALA	SER	CONFLICT	UNP C3V005
C	277	GLU	ASP	CONFLICT	UNP C3V005
C	317	ALA	ILE	CONFLICT	UNP C3V005
C	331	THR	ALA	CONFLICT	UNP C3V005
C	358	ILE	MET	CONFLICT	UNP C3V005

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.

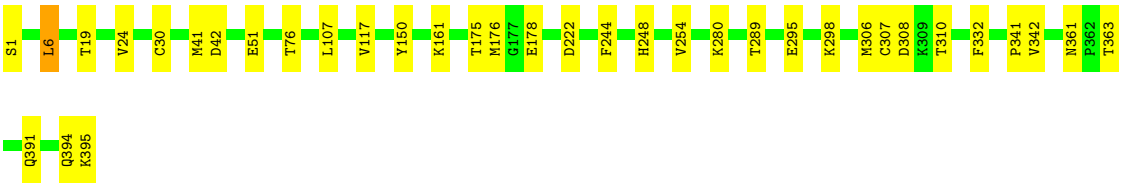
• Molecule 1: MAJOR ENVELOPE PROTEIN E

Chain A: 91% 9%



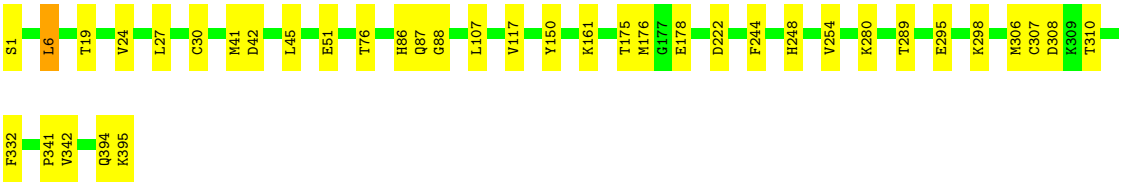
• Molecule 1: MAJOR ENVELOPE PROTEIN E

Chain B: 91% 9%



• Molecule 1: MAJOR ENVELOPE PROTEIN E

Chain C: 90% 9%



GLOBAL-STATISTICS INFOmissingINFO

4 Model quality (i)

4.1 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	0/3096	0.74	0/4205
1	B	0.46	0/3096	0.74	0/4205
1	C	0.46	0/3096	0.74	0/4205
All	All	0.46	0/9288	0.74	0/12615

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
1	B	0	1
1	C	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	150	TYR	Sidechain
1	B	150	TYR	Sidechain
1	C	150	TYR	Sidechain

4.2 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	3028	0	2983	27	0
1	B	3028	0	2983	19	0
1	C	3028	0	2983	25	0
All	All	9084	0	8949	61	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 3.

The worst 5 of 61 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:86:HIS:ND1	1:C:87:GLN:HA	1.65	1.11
1:A:86:HIS:CG	1:C:87:GLN:HA	2.00	0.96
1:B:24:VAL:HG22	1:B:289:THR:HG22	1.53	0.90
1:A:24:VAL:HG22	1:A:289:THR:HG22	1.53	0.90
1:C:24:VAL:HG22	1:C:289:THR:HG22	1.53	0.89

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	393/395 (100%)	381 (97%)	12 (3%)	0	100	100
1	B	393/395 (100%)	381 (97%)	12 (3%)	0	100	100
1	C	393/395 (100%)	381 (97%)	12 (3%)	0	100	100
All	All	1179/1185 (100%)	1143 (97%)	36 (3%)	0	100	100

There are no Ramachandran outliers to report.

4.3.2 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	327/327 (100%)	322 (98%)	5 (2%)	65	80
1	B	327/327 (100%)	322 (98%)	5 (2%)	65	80
1	C	327/327 (100%)	322 (98%)	5 (2%)	65	80
All	All	981/981 (100%)	966 (98%)	15 (2%)	66	80

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	222	ASP
1	C	248	HIS
1	B	248	HIS
1	C	342	VAL
1	C	117	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	214	GLN
1	C	221	ASN
1	B	130	HIS
1	B	214	GLN
1	B	221	ASN

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

There are no ligands in this entry.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.