



## Full wwPDB EM Validation Report ⓘ

Mar 11, 2025 – 06:05 am GMT

PDB ID : 8S41  
EMDB ID : EMD-19708  
Title : The structure of the copia retrotransposon icosahedral capsid (T=9)  
Authors : Klumpe, S.; Beck, F.; Briggs, J.A.G.; Beck, M.; Plitzko, J.M.  
Deposited on : 2024-02-20  
Resolution : 7.70 Å(reported)

This is a Full wwPDB EM Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

EMDB validation analysis : **FAILED**  
MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
MapQ : **FAILED**  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.41

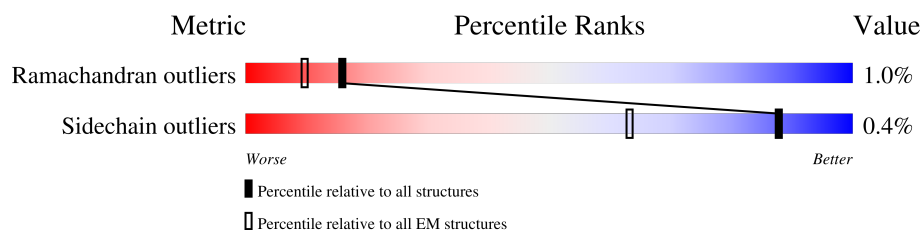
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 7.70 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)
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

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  $\geq 3$ , 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	187	97% .
1	F	187	96% . .
1	G	187	95% 5%
1	L	187	98% .
1	Q	187	97% .
1	V	187	96% . .
1	X	187	97% .
1	a	187	96% .
1	d	187	97% .

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 13509 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 Copia VLP protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		
1	F	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		
1	G	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		
1	L	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		
1	Q	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		
1	V	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		
1	X	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		
1	a	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		
1	d	187	Total	C	N	O	S	0	0
			1501	958	252	285	6		

### 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: Copia VLP protein

Chain A:  97% .



- Molecule 1: Copia VLP protein

Chain F:  96% ..



- Molecule 1: Copia VLP protein

Chain G:  95% 5%



- Molecule 1: Copia VLP protein

Chain L:  98% .



- Molecule 1: Copia VLP protein

Chain Q:  97% .

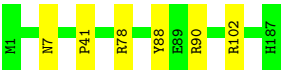


- Molecule 1: Copia VLP protein

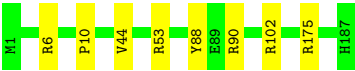
Chain V:  96% ..



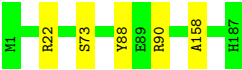
• Molecule 1: Copia VLP protein



• Molecule 1: Copia VLP protein



• Molecule 1: Copia VLP protein



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SUBTOMOGRAM AVERAGING	Depositor
Imposed symmetry	POINT, I	Depositor
Number of subtomograms used	735	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 ( $e^-/\text{\AA}^2$ )	120	Depositor
Minimum defocus (nm)	3500	Depositor
Maximum defocus (nm)	5000	Depositor
Magnification	Not provided	
Image detector	FEI FALCON IV (4k x 4k)	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	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.61	0/1522	1.02	2/2048 (0.1%)
1	F	0.60	0/1522	1.00	6/2048 (0.3%)
1	G	0.63	0/1522	1.04	6/2048 (0.3%)
1	L	0.61	0/1522	0.96	2/2048 (0.1%)
1	Q	0.60	0/1522	0.96	2/2048 (0.1%)
1	V	0.61	0/1522	1.02	4/2048 (0.2%)
1	X	0.62	0/1522	0.97	2/2048 (0.1%)
1	a	0.61	0/1522	0.96	4/2048 (0.2%)
1	d	0.62	0/1522	0.96	2/2048 (0.1%)
All	All	0.61	0/13698	0.99	30/18432 (0.2%)

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	F	0	1
1	Q	0	1
1	X	0	1
1	a	0	2
1	d	0	2
All	All	0	8

There are no bond length outliers.

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	78	ARG	NE-CZ-NH1	10.76	125.68	120.30
1	V	90	ARG	NE-CZ-NH1	8.86	124.73	120.30
1	F	53	ARG	NE-CZ-NH1	8.29	124.45	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	16	TYR	CB-CG-CD2	-8.14	116.12	121.00
1	V	78	ARG	NE-CZ-NH1	8.02	124.31	120.30
1	G	6	ARG	NE-CZ-NH1	7.86	124.23	120.30
1	Q	22	ARG	NE-CZ-NH1	7.73	124.16	120.30
1	G	24	ARG	NE-CZ-NH1	7.21	123.90	120.30
1	a	90	ARG	NE-CZ-NH1	6.98	123.79	120.30
1	F	22	ARG	NE-CZ-NH1	6.67	123.64	120.30
1	G	53	ARG	NE-CZ-NH1	6.42	123.51	120.30
1	a	6	ARG	NE-CZ-NH2	6.07	123.34	120.30
1	F	62	TYR	CB-CG-CD2	-5.93	117.44	121.00
1	F	24	ARG	NE-CZ-NH1	5.89	123.25	120.30
1	G	78	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	a	102	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	F	102	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	V	24	ARG	NE-CZ-NH1	5.61	123.11	120.30
1	X	102	ARG	NE-CZ-NH1	5.54	123.07	120.30
1	X	78	ARG	NE-CZ-NH1	5.49	123.05	120.30
1	Q	90	ARG	NE-CZ-NH1	5.32	122.96	120.30
1	d	22	ARG	NE-CZ-NH1	5.28	122.94	120.30
1	V	22	ARG	NE-CZ-NH1	5.27	122.94	120.30
1	A	22	ARG	NE-CZ-NH1	5.23	122.92	120.30
1	F	100	ARG	NE-CZ-NH1	5.22	122.91	120.30
1	a	175	ARG	NE-CZ-NH1	5.18	122.89	120.30
1	d	90	ARG	NE-CZ-NH1	5.17	122.89	120.30
1	L	90	ARG	NE-CZ-NH2	5.12	122.86	120.30
1	G	102	ARG	NE-CZ-NH1	5.10	122.85	120.30
1	A	152	TYR	CB-CG-CD2	-5.02	117.99	121.00

There are no chirality outliers.

All (8) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	88	TYR	Sidechain
1	F	62	TYR	Sidechain
1	Q	175	ARG	Sidechain
1	X	88	TYR	Sidechain
1	a	53	ARG	Sidechain
1	a	88	TYR	Sidechain
1	d	158	ALA	Mainchain
1	d	88	TYR	Sidechain



## 5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

## 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	185/187 (99%)	168 (91%)	17 (9%)	0	100	100
1	F	185/187 (99%)	172 (93%)	13 (7%)	0	100	100
1	G	185/187 (99%)	174 (94%)	8 (4%)	3 (2%)	8	38
1	L	185/187 (99%)	170 (92%)	14 (8%)	1 (0%)	25	64
1	Q	185/187 (99%)	175 (95%)	8 (4%)	2 (1%)	12	47
1	V	185/187 (99%)	168 (91%)	13 (7%)	4 (2%)	5	29
1	X	185/187 (99%)	169 (91%)	13 (7%)	3 (2%)	8	38
1	a	185/187 (99%)	175 (95%)	8 (4%)	2 (1%)	12	47
1	d	185/187 (99%)	175 (95%)	9 (5%)	1 (0%)	25	64
All	All	1665/1683 (99%)	1546 (93%)	103 (6%)	16 (1%)	16	49

All (16) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	d	73	SER
1	V	73	SER
1	G	10	PRO
1	G	109	SER
1	V	90	ARG
1	X	90	ARG
1	V	10	PRO
1	X	41	PRO
1	L	10	PRO
1	Q	167	LEU
1	V	44	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	X	7	ASN
1	a	10	PRO
1	G	149	PRO
1	a	44	VAL
1	Q	10	PRO

### 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	167/167 (100%)	165 (99%)	2 (1%)	67	78
1	F	167/167 (100%)	166 (99%)	1 (1%)	84	88
1	G	167/167 (100%)	165 (99%)	2 (1%)	67	78
1	L	167/167 (100%)	167 (100%)	0	100	100
1	Q	167/167 (100%)	167 (100%)	0	100	100
1	V	167/167 (100%)	166 (99%)	1 (1%)	84	88
1	X	167/167 (100%)	167 (100%)	0	100	100
1	a	167/167 (100%)	167 (100%)	0	100	100
1	d	167/167 (100%)	167 (100%)	0	100	100
All	All	1503/1503 (100%)	1497 (100%)	6 (0%)	88	91

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	90	ARG
1	A	100	ARG
1	F	9	LYS
1	G	9	LYS
1	G	34	LYS
1	V	90	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	F	166	ASN

### 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 oligosaccharides 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.