



# wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jul 24, 2017 – 03:30 AM EDT

PDB ID : 5MUU  
EMDB ID: : EMD-3571  
Title : dsRNA bacteriophage phi6 nucleocapsid  
Authors : Sun, Z.; El Omari, K.; Sun, X.; Ilca, S.L.; Kotecha, A.; Stuart, D.I.; Poranen, M.M.; Huiskonen, J.T.  
Deposited on : unknown  
Resolution : 4.00 Å(reported)  
Based on PDB ID : 4K7H, ?

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](mailto: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.

---

MolProbity : 4.02b-467  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

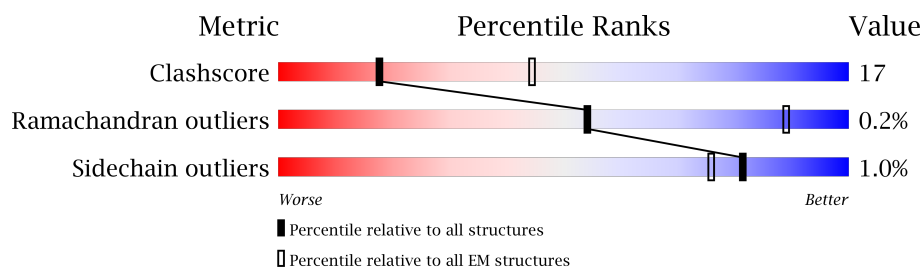
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.00 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	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  $\geq 3$ , 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	769	
1	B	769	
2	C	332	
3	D	149	
3	E	149	
3	F	149	
3	G	149	
3	H	149	
3	I	149	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	J	149	<div><div></div><div>80%</div><div>19%</div><div>.</div></div>
3	K	149	<div><div></div><div>73%</div><div>25%</div><div>..</div></div>
3	L	149	<div><div></div><div>72%</div><div>27%</div><div>.</div></div>
3	M	149	<div><div></div><div>74%</div><div>25%</div><div>.</div></div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 23386 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 inner protein P1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	766	Total	C	N	O	S	0	0
			5964	3767	1053	1121	23		
1	B	766	Total	C	N	O	S	0	0
			5964	3767	1053	1121	23		

- Molecule 2 is a protein called Packaging enzyme P4.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	39	Total	C	N	O	S	0	0
			288	171	56	60	1		

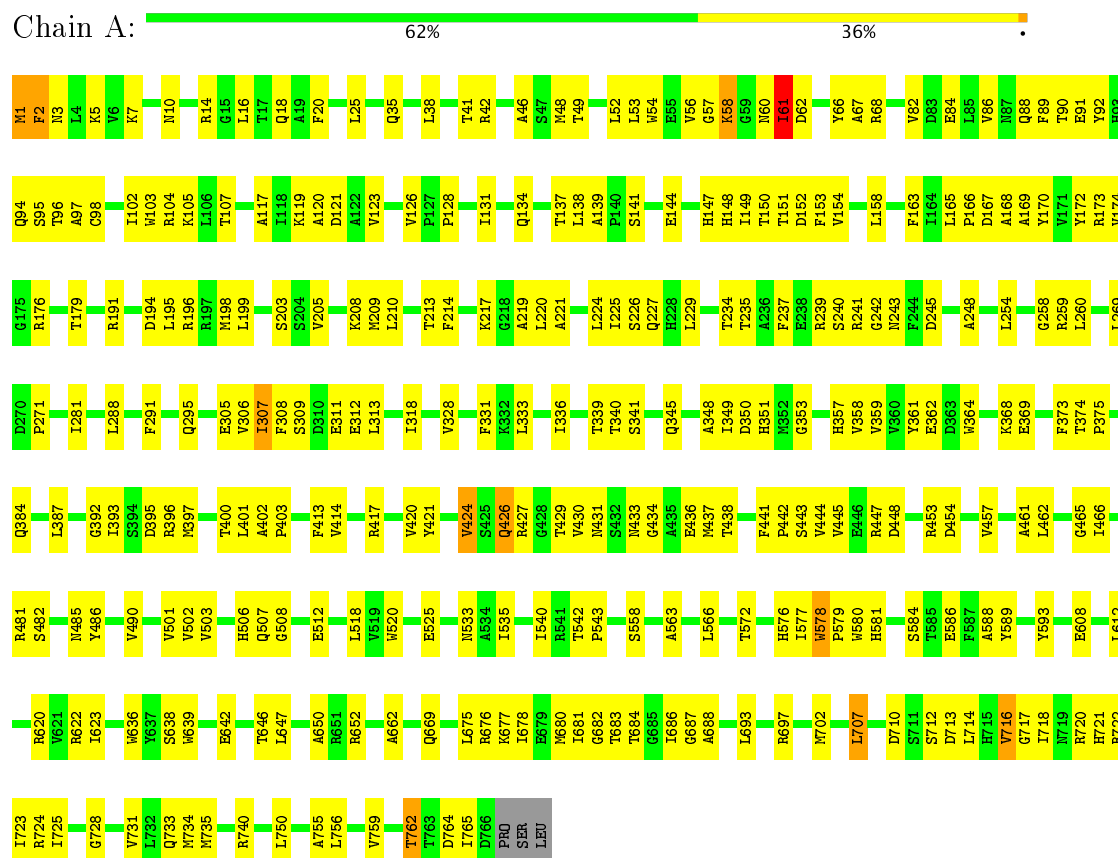
- Molecule 3 is a protein called Major outer capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	E	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	F	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	G	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	H	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	I	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	J	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	K	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	L	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		
3	M	148	Total	C	N	O	S	0	0
			1117	711	189	212	5		

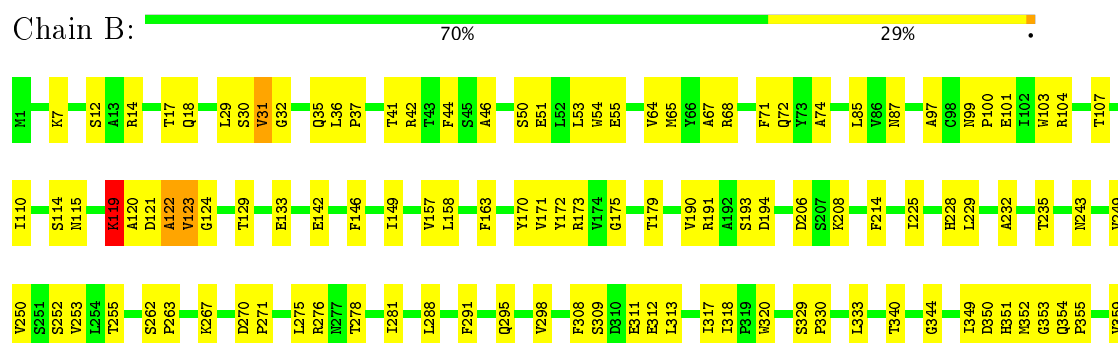
### 3 Residue-property plots

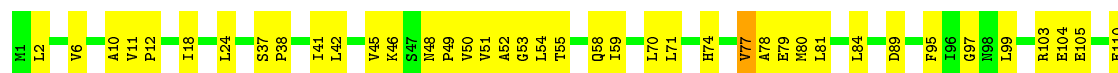
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according 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.

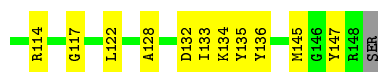
#### • Molecule 1: Major inner protein P1



#### • Molecule 1: Major inner protein P1







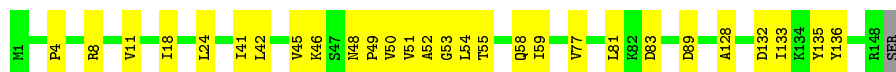
- Molecule 3: Major outer capsid protein

Chain F: 78% 21% .



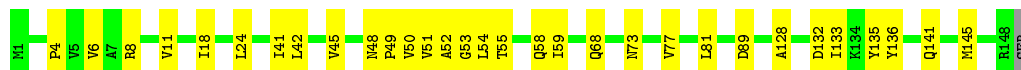
- Molecule 3: Major outer capsid protein

Chain G: 81% 19% .



- Molecule 3: Major outer capsid protein

Chain H: 79% 21% .



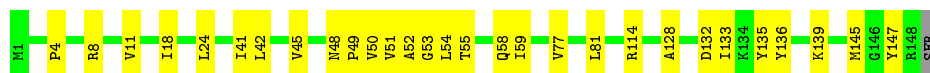
- Molecule 3: Major outer capsid protein

Chain I: 78% 21% .



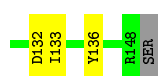
- Molecule 3: Major outer capsid protein

Chain J: 80% 19% .



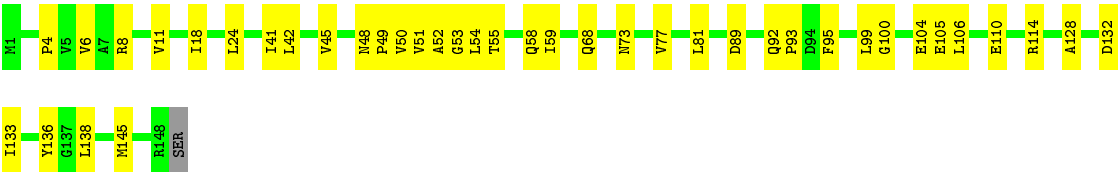
- Molecule 3: Major outer capsid protein

Chain K: 73% 25% ..



- Molecule 3: Major outer capsid protein

Chain L: 72% 27% .



• Molecule 3: Major outer capsid protein

Chain M: 

74%

25%





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	13291	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	0.73	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	37037	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	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	0.48	0/6084	0.56	2/8267 (0.0%)
1	B	0.57	0/6084	0.59	0/8267
2	C	0.45	0/292	0.57	0/394
3	D	0.38	0/1133	0.42	0/1535
3	E	0.38	0/1133	0.46	0/1535
3	F	0.38	0/1133	0.42	0/1535
3	G	0.38	0/1133	0.42	0/1535
3	H	0.38	0/1133	0.42	0/1535
3	I	0.38	0/1133	0.42	0/1535
3	J	0.39	0/1133	0.43	0/1535
3	K	0.34	0/1133	0.50	0/1535
3	L	0.38	0/1133	0.43	0/1535
3	M	0.38	0/1133	0.42	0/1535
All	All	0.46	0/23790	0.51	2/32278 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	121	ASP	CB-CG-OD2	5.18	122.96	118.30
1	A	707	LEU	CA-CB-CG	5.15	127.15	115.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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	5964	0	5953	300	0
1	B	5964	0	5953	217	0
2	C	288	0	278	26	0
3	D	1117	0	1149	55	0
3	E	1117	0	1148	69	0
3	F	1117	0	1149	21	0
3	G	1117	0	1149	19	0
3	H	1117	0	1149	23	0
3	I	1117	0	1149	24	0
3	J	1117	0	1149	27	0
3	K	1117	0	1149	52	0
3	L	1117	0	1149	46	0
3	M	1117	0	1149	41	0
All	All	23386	0	23673	796	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 17.

The worst 5 of 796 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:B:712:SER:CB	1:B:716:VAL:HG23	1.26	1.54
1:B:712:SER:HB2	1:B:716:VAL:CG2	1.07	1.51
1:B:710:ASP:OD2	1:B:715:HIS:CG	1.70	1.44
3:K:92:GLN:OE1	3:K:97:GLY:HA2	1.31	1.30
1:B:387:LEU:HD11	1:B:576:HIS:CE1	1.67	1.28

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	764/769 (99%)	724 (95%)	36 (5%)	4 (0%)	32	73
1	B	764/769 (99%)	735 (96%)	26 (3%)	3 (0%)	38	77
2	C	37/332 (11%)	37 (100%)	0	0	100	100
3	D	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
3	E	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
3	F	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
3	G	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
3	H	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
3	I	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
3	J	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
3	K	146/149 (98%)	141 (97%)	5 (3%)	0	100	100
3	L	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
3	M	146/149 (98%)	145 (99%)	1 (1%)	0	100	100
All	All	3025/3360 (90%)	2942 (97%)	76 (2%)	7 (0%)	54	84

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	578	TRP
1	B	119	LYS
1	A	61	ILE
1	A	716	VAL
1	B	122	ALA

### 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	635/638 (100%)	626 (99%)	9 (1%)	71	87
1	B	635/638 (100%)	626 (99%)	9 (1%)	71	87
2	C	32/269 (12%)	32 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	D	116/117 (99%)	116 (100%)	0	100	100
3	E	116/117 (99%)	114 (98%)	2 (2%)	66	86
3	F	116/117 (99%)	116 (100%)	0	100	100
3	G	116/117 (99%)	116 (100%)	0	100	100
3	H	116/117 (99%)	116 (100%)	0	100	100
3	I	116/117 (99%)	115 (99%)	1 (1%)	82	92
3	J	116/117 (99%)	116 (100%)	0	100	100
3	K	116/117 (99%)	114 (98%)	2 (2%)	66	86
3	L	116/117 (99%)	116 (100%)	0	100	100
3	M	116/117 (99%)	115 (99%)	1 (1%)	82	92
All	All	2462/2715 (91%)	2438 (99%)	24 (1%)	81	90

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	123	VAL
1	B	580	TRP
3	K	95	PHE
1	B	512	GLU
1	B	513	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 54 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	497	HIS
1	B	715	HIS
3	K	73	ASN
1	B	498	ASN
1	B	556	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.