



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

Dec 9, 2019 – 12:38 AM EST

PDB ID : 6R3A  
EMDB ID: : EMD-4716  
Title : BACTERIOPHAGE SPP1 MATURE CAPSID PROTEIN  
Authors : Ignatiou, A.; El Sadek Fadel, M.; Buerger, J.; Mielke, T.; Topf, M.; Tavares, P.  
Deposited on : 2019-03-19  
Resolution : 4.00 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](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.

---

MolProbity : 4.02b-467  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

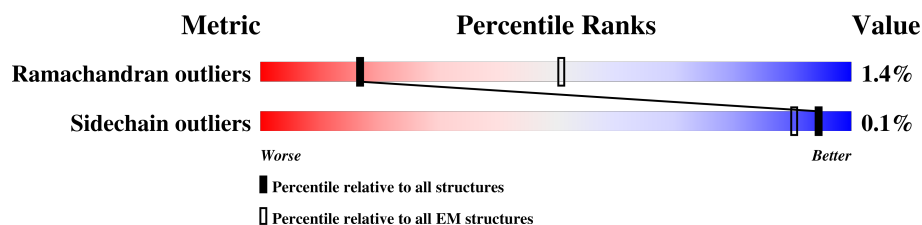
# 1 Overall quality at a glance

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)
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

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	323	96% .
1	B	323	98% .
1	C	323	96% .
1	D	323	97% .
1	E	323	96% .
1	F	323	96% .
1	G	323	96% .

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 17360 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.

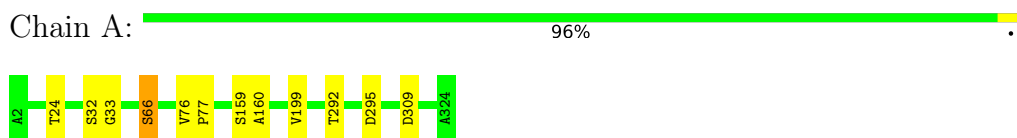
- Molecule 1 is a protein called Major capsid protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	323	Total 2480	C 1559	N 420	O 492	S 9	0	0
1	B	323	Total 2480	C 1559	N 420	O 492	S 9	0	0
1	C	323	Total 2480	C 1559	N 420	O 492	S 9	0	0
1	D	323	Total 2480	C 1559	N 420	O 492	S 9	0	0
1	E	323	Total 2480	C 1559	N 420	O 492	S 9	0	0
1	F	323	Total 2480	C 1559	N 420	O 492	S 9	0	0
1	G	323	Total 2480	C 1559	N 420	O 492	S 9	0	0

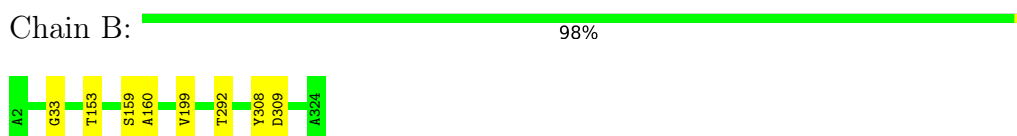
### 3 Residue-property plots [i](#)

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.

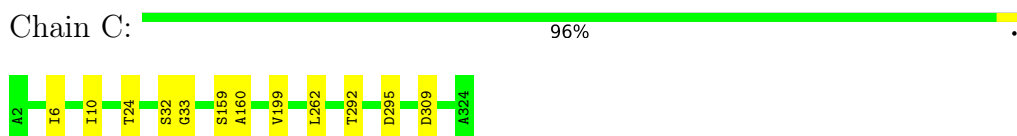
- Molecule 1: Major capsid protein



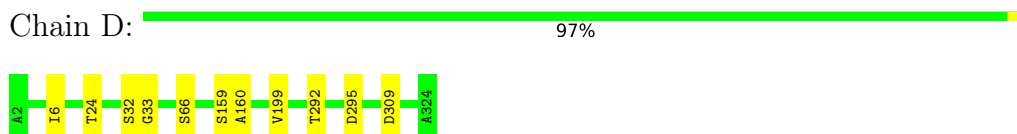
- Molecule 1: Major capsid protein



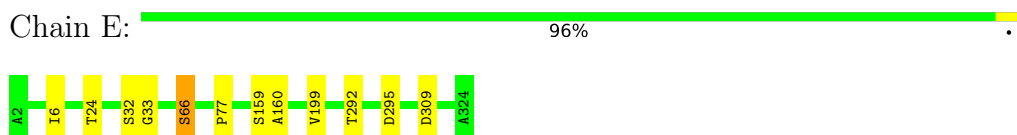
- Molecule 1: Major capsid protein



- Molecule 1: Major capsid protein



- Molecule 1: Major capsid protein



- Molecule 1: Major capsid protein



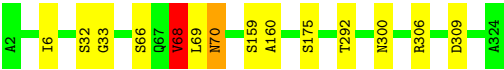


● Molecule 1: Major capsid protein

Chain G: 

96%

.



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	4500	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY; Software used Imagic 5	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	25	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	31000	Depositor
Image detector	GATAN K2 BASE (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	$\# Z  > 2$	RMSZ	$\# Z  > 2$
1	A	0.48	0/2528	0.71	4/3429 (0.1%)
1	B	0.50	0/2528	0.71	2/3429 (0.1%)
1	C	0.48	0/2528	0.69	2/3429 (0.1%)
1	D	0.48	0/2528	0.67	2/3429 (0.1%)
1	E	0.48	0/2528	0.69	2/3429 (0.1%)
1	F	0.48	0/2528	0.69	3/3429 (0.1%)
1	G	0.46	0/2528	0.69	4/3429 (0.1%)
All	All	0.48	0/17696	0.69	19/24003 (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	6
1	B	0	3
1	C	0	4
1	D	0	5
1	E	0	5
1	F	0	6
1	G	0	8
All	All	0	37

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	66	SER	CA-C-N	-6.68	102.50	117.20
1	B	33	GLY	N-CA-C	-6.53	96.78	113.10
1	G	292	THR	C-N-CD	-6.32	106.69	120.60
1	C	33	GLY	N-CA-C	-6.21	97.57	113.10
1	E	33	GLY	N-CA-C	-6.19	97.62	113.10

There are no chirality outliers.

5 of 37 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	159	SER	Peptide
1	A	24	THR	Peptide
1	A	32	SER	Peptide
1	A	66	SER	Mainchain
1	A	76	VAL	Peptide

## 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	2480	0	2420	0	0
1	B	2480	0	2420	0	0
1	C	2480	0	2420	0	0
1	D	2480	0	2420	0	0
1	E	2480	0	2420	0	0
1	F	2480	0	2420	0	0
1	G	2480	0	2420	0	0
All	All	17360	0	16940	0	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). Clashscore could not be calculated for this entry.

There are no clashes within the asymmetric unit.

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	321/323 (99%)	267 (83%)	50 (16%)	4 (1%)	14	55
1	B	321/323 (99%)	271 (84%)	47 (15%)	3 (1%)	19	60
1	C	321/323 (99%)	265 (83%)	51 (16%)	5 (2%)	11	50
1	D	321/323 (99%)	265 (83%)	52 (16%)	4 (1%)	14	55
1	E	321/323 (99%)	263 (82%)	52 (16%)	6 (2%)	9	46
1	F	321/323 (99%)	262 (82%)	54 (17%)	5 (2%)	11	50
1	G	321/323 (99%)	267 (83%)	49 (15%)	5 (2%)	11	50
All	All	2247/2261 (99%)	1860 (83%)	355 (16%)	32 (1%)	17	51

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	77	PRO
1	F	77	PRO
1	A	160	ALA
1	B	160	ALA
1	C	160	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	264/264 (100%)	264 (100%)	0	100	100
1	B	264/264 (100%)	264 (100%)	0	100	100
1	C	264/264 (100%)	263 (100%)	1 (0%)	92	96
1	D	264/264 (100%)	264 (100%)	0	100	100
1	E	264/264 (100%)	264 (100%)	0	100	100
1	F	264/264 (100%)	264 (100%)	0	100	100
1	G	264/264 (100%)	264 (100%)	0	100	100
All	All	1848/1848 (100%)	1847 (100%)	1 (0%)	94	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	10	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 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.