



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

Jan 10, 2018 – 06:02 PM EST

PDB ID : 6C22  
EMDB ID: : EMD-7333  
Title : Capsid protein in the Staphylococcus aureus phage 80alpha-derived SaPI1 mature capsid  
Authors : Kizziah, J.L.; Dearborn, A.D.; Dokland, T.  
Deposited on : 2018-01-05  
Resolution : 8.00 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  
<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-20030736

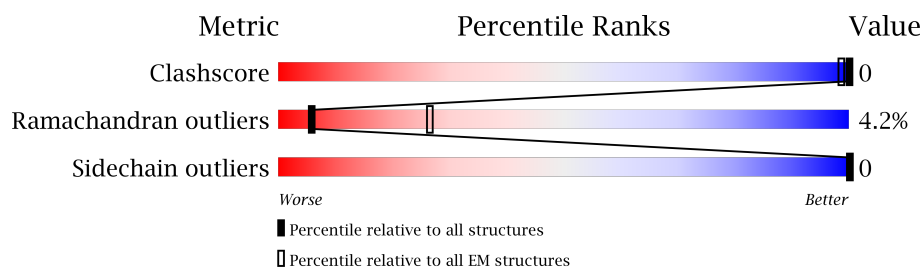
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.00 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	324	82% 15%
1	B	324	81% 15%
1	C	324	81% 15%
1	D	324	81% 15%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 5464 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 head protein.

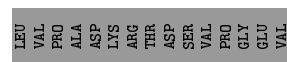
Mol	Chain	Residues	Atoms				AltConf	Trace
1	A	275	Total 1366	C 815	N 275	O 276	0	0
1	B	275	Total 1366	C 815	N 275	O 276	0	0
1	C	275	Total 1366	C 815	N 275	O 276	0	0
1	D	275	Total 1366	C 815	N 275	O 276	0	0



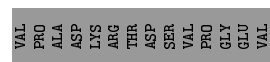
- Molecule 1: Major head protein



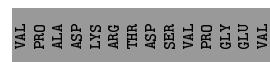
- Molecule 1: Major head protein



- Molecule 1: Major head protein



- Molecule 1: Major head protein



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	6471	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	25	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	62000	Depositor
Image detector	KODAK SO-163 FILM	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	0.30	0/1366	0.51	0/1901
1	B	0.30	0/1366	0.52	0/1901
1	C	0.30	0/1366	0.51	0/1901
1	D	0.31	0/1366	0.52	0/1901
All	All	0.30	0/5464	0.51	0/7604

There are no bond length outliers.

There are no bond angle outliers.

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	1366	0	619	0	0
1	B	1366	0	619	0	0
1	C	1366	0	619	0	0
1	D	1366	0	619	0	0
All	All	5464	0	2476	0	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 0.

There are no clashes within the asymmetric unit.

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	273/324 (84%)	246 (90%)	17 (6%)	10 (4%)	4	33
1	B	273/324 (84%)	241 (88%)	20 (7%)	12 (4%)	3	29
1	C	273/324 (84%)	240 (88%)	21 (8%)	12 (4%)	3	29
1	D	273/324 (84%)	239 (88%)	22 (8%)	12 (4%)	3	29
All	All	1092/1296 (84%)	966 (88%)	80 (7%)	46 (4%)	6	30

5 of 46 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	257	PRO
1	A	280	PRO
1	B	273	VAL
1	B	280	PRO
1	C	280	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	1/286 (0%)	1 (100%)	0	100	100
1	B	1/286 (0%)	1 (100%)	0	100	100
1	C	1/286 (0%)	1 (100%)	0	100	100
1	D	1/286 (0%)	1 (100%)	0	100	100
All	All	4/1144 (0%)	4 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

### 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 ⓘ

There are no ligands in this entry.

## 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.