



Full wwPDB/EMDatabank EM Map/Model Validation Report ⓘ

Jan 10, 2018 – 06:04 PM EST

PDB ID : 6C21
EMDB ID: : EMD-7332
Title : Capsid protein in the Staphylococcus aureus phage 80alpha mature capsid
Authors : Kizziah, J.L.; Dearborn, A.D.; Dokland, T.
Deposited on : 2018-01-05
Resolution : 5.20 Å(reported)

This is a Full wwPDB/EMDatabank EM Map/Model Validation Report
for a publicly released PDB/EMDB entry.

We welcome your comments at 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

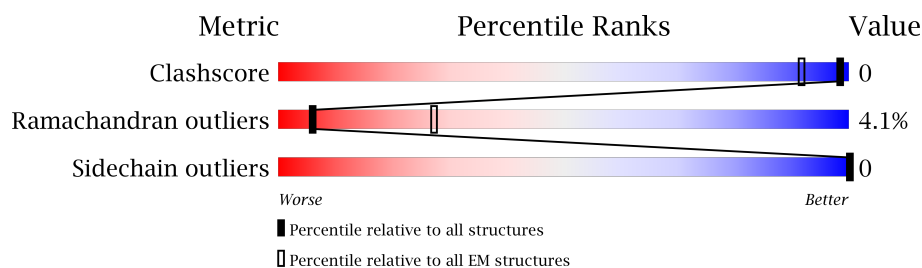
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 5.20 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 ≥ 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	81% 15%
1	B	324	80% 5% 15%
1	C	324	81% 15%
1	D	324	81% 15%
1	E	324	81% 15%
1	F	324	83% 15%
1	G	324	81% 15%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 9562 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
1	E	275	Total 1366	C 815	N 275	O 276	0	0
1	F	275	Total 1366	C 815	N 275	O 276	0	0
1	G	275	Total 1366	C 815	N 275	O 276	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.

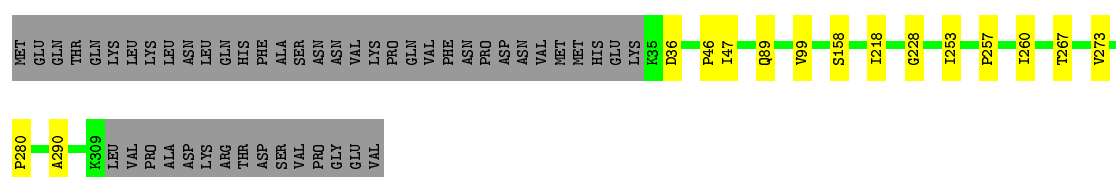
- Molecule 1: Major head protein

Chain A: 




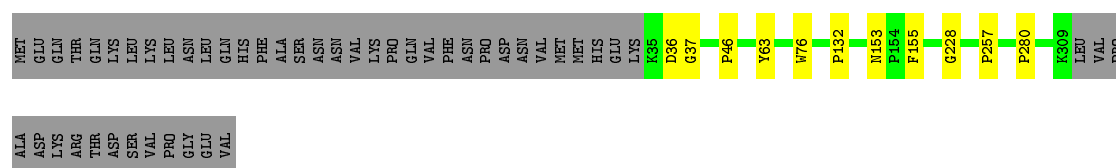
- Molecule 1: Major head protein

Chain B: 




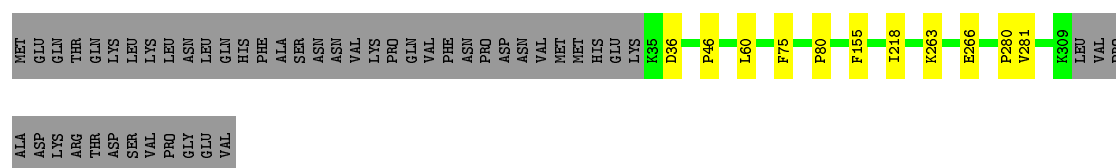
- Molecule 1: Major head protein

Chain C: 



- Molecule 1: Major head protein

Chain D: 



- Molecule 1: Major head protein

K309
LEU
VAL
PRO
ALA
ASP
LYS
ARG
THR
ASP
SER
VAL
PRO
GLY
GLU
VAL

- | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| MET | GLU | GLN | GLU | THR | GLN | LYS | LEU | LYS | LEU | ASN | ASN | GLN | HIS | PHI | SER | ASP | ASN | VAL | LYS | PRO | GLN | VAL | PHI | ASP | ASP | ASN | VAL | MET | MET | HIS | GLU | LYS | K35 | D36 | E42 | PRO | M167 | I170 | R217 | P280 | K309 | LEU | VAL | PRO | ALA | ASP | LYS | LYS | THR | ASP | SER |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

VAL
PRO
GLY
GLU
VAL

- | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|
| MET | GLN | GLU | THR | GLN | LYS | LEU | LEU | ASN | ASN | VAL | LYS | PRO | GLN | VAL | ASP | ASN | VAL | MET | MET | HIS | GLU | LYS | K35 | D36 | M41 | T45 | P80 | E190 | R241 | G242 | E243 | K263 | D265 | E266 | V273 | E276 | N282 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|

K309 LEU VAL PRO ALA ASP LYS ARG THR ASP SER VAL PRO GLY GLU VAL

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, 1	Depositor
Number of particles used	11843	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY; Estimated with ACE software	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	120	Depositor
Electron dose ($e^-/\text{\AA}^2$)	15	Depositor
Minimum defocus (nm)	1250	Depositor
Maximum defocus (nm)	3000	Depositor
Magnification	96000	Depositor
Image detector	GATAN ULTRASCAN 4000 (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	0.31	0/1366	0.52	0/1901
1	B	0.31	0/1366	0.53	0/1901
1	C	0.31	0/1366	0.53	0/1901
1	D	0.32	0/1366	0.54	0/1901
1	E	0.32	0/1366	0.56	0/1901
1	F	0.33	0/1366	0.54	0/1901
1	G	0.32	0/1366	0.54	0/1901
All	All	0.32	0/9562	0.54	0/13307

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	2	0
1	B	1366	0	619	1	0
1	C	1366	0	619	0	0
1	D	1366	0	619	0	0
1	E	1366	0	619	0	0
1	F	1366	0	619	0	0
1	G	1366	0	619	0	0
All	All	9562	0	4333	3	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.

All (3) 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:257:PRO:HA	1:B:290:ALA:H	1.67	0.58
1:A:101:ALA:H	1:A:292:ARG:HA	1.82	0.43
1:A:101:ALA:H	1:A:293:ALA:H	1.68	0.42

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%)	238 (87%)	26 (10%)	9 (3%)	4	36
1	B	273/324 (84%)	239 (88%)	21 (8%)	13 (5%)	2	28
1	C	273/324 (84%)	234 (86%)	28 (10%)	11 (4%)	3	31
1	D	273/324 (84%)	231 (85%)	31 (11%)	11 (4%)	3	31
1	E	273/324 (84%)	234 (86%)	25 (9%)	14 (5%)	2	26
1	F	273/324 (84%)	232 (85%)	34 (12%)	7 (3%)	6	41
1	G	273/324 (84%)	234 (86%)	25 (9%)	14 (5%)	2	26
All	All	1911/2268 (84%)	1642 (86%)	190 (10%)	79 (4%)	6	31

All (79) 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

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	46	PRO
1	C	132	PRO
1	C	280	PRO
1	D	46	PRO
1	D	80	PRO
1	D	280	PRO
1	D	281	VAL
1	E	46	PRO
1	E	279	THR
1	E	280	PRO
1	F	280	PRO
1	G	80	PRO
1	A	99	VAL
1	A	158	SER
1	A	172	GLY
1	A	218	ILE
1	B	158	SER
1	C	257	PRO
1	D	155	PHE
1	E	216	GLU
1	E	281	VAL
1	A	216	GLU
1	A	228	GLY
1	B	36	ASP
1	B	267	THR
1	C	63	TYR
1	C	76	TRP
1	C	153	ASN
1	C	155	PHE
1	D	36	ASP
1	D	75	PHE
1	D	218	ILE
1	D	263	LYS
1	D	266	GLU
1	E	153	ASN
1	E	219	TYR
1	E	228	GLY
1	E	239	LEU
1	E	263	LYS
1	F	167	ASN
1	G	41	ASN
1	G	45	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	G	190	GLU
1	G	242	GLY
1	G	266	GLU
1	G	276	GLU
1	G	282	ASN
1	B	89	GLN
1	B	228	GLY
1	C	36	ASP
1	C	37	GLY
1	C	228	GLY
1	D	60	LEU
1	E	224	ASP
1	E	266	GLU
1	F	36	ASP
1	F	42	GLU
1	G	36	ASP
1	G	243	GLU
1	B	99	VAL
1	B	218	ILE
1	E	76	TRP
1	E	218	ILE
1	F	80	PRO
1	F	217	ARG
1	G	263	LYS
1	A	76	TRP
1	B	260	ILE
1	G	241	ARG
1	B	253	ILE
1	G	273	VAL
1	B	47	ILE
1	F	170	ILE
1	B	46	PRO
1	G	264	ILE

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
1	E	1/286 (0%)	1 (100%)	0	100	100
1	F	1/286 (0%)	1 (100%)	0	100	100
1	G	1/286 (0%)	1 (100%)	0	100	100
All	All	7/2002 (0%)	7 (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 [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.