



Full wwPDB/EMDataBank EM Map/Model Validation Report ⓘ

Feb 21, 2018 – 04:52 PM EST

PDB ID : 6B23
EMDB ID: : EMD-7035
Title : Capsid protein and C-terminal part of CpmB protein in the Staphylococcus aureus pathogenicity island 1 80alpha-derived procapsid
Authors : Kizziah, J.L.; Dearborn, A.D.; Dokland, T.
Deposited on : 2017-09-19
Resolution : 3.76 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

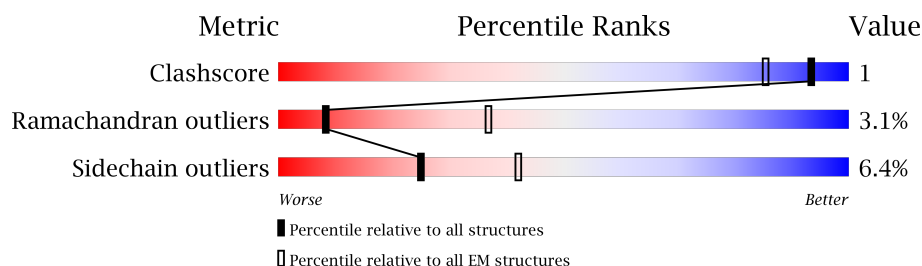
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.76 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	77% 10% 12%
1	B	324	78% 9% • 12%
1	C	324	74% 12% • 12%
1	D	324	80% 7% • 12%
2	a	72	17% • 79%
2	b	72	19% • 79%
2	c	72	43% • 54%
2	d	72	46% 54%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9854 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	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	B	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	C	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		
1	D	284	Total	C	N	O	S	0	0
			2270	1446	370	443	11		

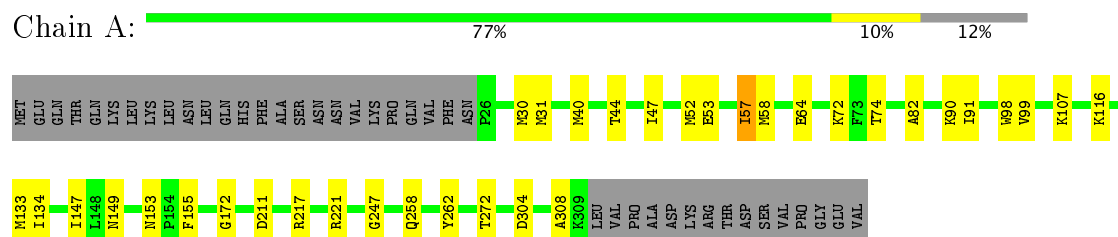
- Molecule 2 is a protein called Capsid morphogenesis B protein.

Mol	Chain	Residues	Atoms				AltConf	Trace
2	a	15	Total	C	N	O	0	0
			118	73	24	21		
2	b	15	Total	C	N	O	0	0
			118	73	24	21		
2	c	33	Total	C	N	O	0	0
			269	163	50	56		
2	d	33	Total	C	N	O	0	0
			269	163	50	56		

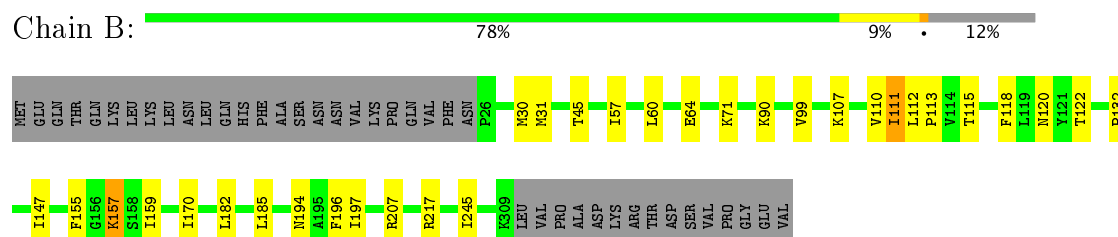
3 Residue-property plots [i](#)

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.

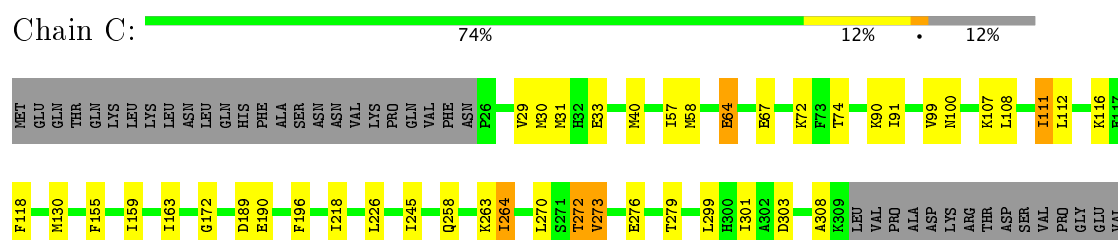
- Molecule 1: Major head protein



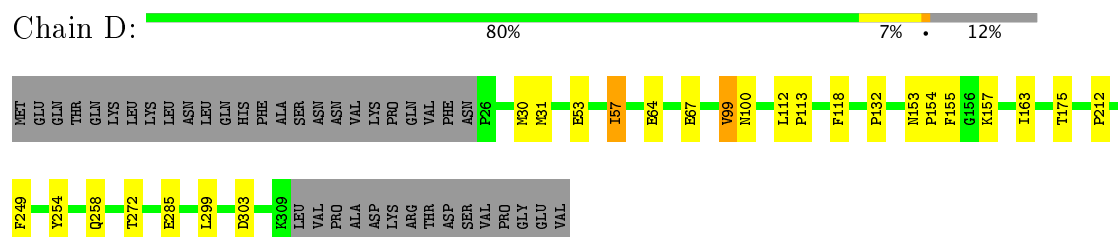
- Molecule 1: Major head protein



- Molecule 1: Major head protein



- Molecule 1: Major head protein



- Molecule 2: Capsid morphogenesis B protein

Chain a:  17% . 79%

MET GLU THR LYS TYR GLU LEU ASN ASN THR LYS LYS VAL ALA ALA ASN PHE GLY LEU ASN ASN GLU GLU ASP THR ASN LEU LEU ILE ILE ASN ALA VAL ASP LEU ASP ILE ILE LYS ASN ASN MET GLN GLU ILE ILE SER SER GLU LEU GLN GLN SER GLU GLN SER SER LYS LYS LYS GLN TYR G58 T59

L64 R69 K72

- Molecule 2: Capsid morphogenesis B protein

Chain b:  19% . 79%

MET GLU THR LYS TYR GLU LEU ASN ASN THR LYS LYS VAL ALA ALA ASN PHE GLY LEU ASN ASN GLU GLU ASP THR ASN LEU LEU ILE ILE ASN ALA VAL ASP LEU ASP ILE ILE LYS ASN ASN MET GLN GLU ILE ILE SER SER GLU LEU GLN GLN SER GLU GLN SER SER LYS LYS LYS GLN TYR G58 L64

K72

- Molecule 2: Capsid morphogenesis B protein

Chain c:  43% . 54%

MET GLU THR LYS TYR GLU LEU ASN ASN THR LYS LYS VAL ALA ALA ASN PHE GLY LEU ASN ASN GLU GLU ASP THR ASN LEU LEU ILE ILE ASN ALA VAL ASP LEU ASP ILE ILE LYS ASN ASN MET G40 Q51 K72

- Molecule 2: Capsid morphogenesis B protein

Chain d:  46% 54%

MET GLU THR LYS TYR GLU LEU ASN ASN THR LYS LYS VAL ALA ALA ASN PHE GLY LEU ASN ASN GLU GLU ASP THR ASN LEU LEU ILE ILE ASN ALA VAL ASP LEU ASP ILE ILE LYS ASN ASN MET G40 K72

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, I	Depositor
Number of particles used	14087	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	40	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	DIRECT ELECTRON DE-20 (5k x 3k)	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	0.38	0/2312	0.51	0/3116
1	B	0.37	0/2312	0.51	0/3116
1	C	0.37	0/2312	0.52	0/3116
1	D	0.37	0/2312	0.53	0/3116
2	a	0.43	0/117	0.63	0/155
2	b	0.40	0/117	0.55	0/155
2	c	0.37	0/269	0.45	0/357
2	d	0.37	0/269	0.47	0/357
All	All	0.37	0/10020	0.52	0/13488

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	2270	0	2261	4	0
1	B	2270	0	2261	9	0
1	C	2270	0	2261	10	0
1	D	2270	0	2261	5	0
2	a	118	0	132	0	0
2	b	118	0	132	0	0
2	c	269	0	275	0	0
2	d	269	0	275	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	9854	0	9858	28	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 1.

All (28) 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:D:112:LEU:HD22	1:D:118:PHE:CZ	2.35	0.61
1:C:159:ILE:HG23	1:C:301:ILE:HG21	1.84	0.60
1:C:159:ILE:CG2	1:C:301:ILE:HG21	2.38	0.53
1:D:57:ILE:HD11	1:D:249:PHE:CE2	2.44	0.53
1:B:170:ILE:HD12	1:B:182:LEU:HD23	1.92	0.52
1:A:47:ILE:CG2	1:A:134:ILE:HD11	2.40	0.52
1:C:196:PHE:CE2	1:C:245:ILE:HD11	2.45	0.50
1:D:163:ILE:HD13	1:D:303:ASP:HB2	1.94	0.50
1:C:218:ILE:HG22	1:C:226:LEU:O	2.12	0.49
1:A:57:ILE:HB	1:A:147:ILE:HD11	1.94	0.48
1:C:263:LYS:HA	1:C:279:THR:HG21	1.95	0.47
1:B:170:ILE:HD12	1:B:182:LEU:CD2	2.44	0.46
1:C:112:LEU:HD22	1:C:118:PHE:CZ	2.51	0.45
1:B:60:LEU:HD22	1:B:197:ILE:HD11	1.99	0.45
1:B:196:PHE:CE2	1:B:245:ILE:HD11	2.52	0.45
1:C:163:ILE:HD13	1:C:303:ASP:HB2	1.99	0.45
1:B:170:ILE:HD11	1:B:185:LEU:HD12	1.99	0.44
1:D:112:LEU:HD22	1:D:118:PHE:CE1	2.53	0.44
1:B:147:ILE:HA	1:B:159:ILE:HD12	1.99	0.44
1:C:108:LEU:HD21	1:C:130:MET:HG2	2.00	0.43
1:B:111:ILE:HG22	1:B:111:ILE:O	2.19	0.43
1:A:52:MET:HG3	1:A:58:MET:HE2	2.02	0.42
1:D:153:ASN:N	1:D:154:PRO:HD3	2.35	0.42
1:B:111:ILE:HG21	1:B:118:PHE:CE1	2.55	0.41
1:C:111:ILE:HG22	1:C:112:LEU:HD23	2.03	0.41
1:B:112:LEU:N	1:B:113:PRO:CD	2.84	0.41
1:C:272:THR:HG22	1:C:273:VAL:HG23	2.02	0.41
1:A:40:MET:O	1:A:44:THR:HG23	2.21	0.40

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	282/324 (87%)	234 (83%)	39 (14%)	9 (3%)	5	39
1	B	282/324 (87%)	233 (83%)	41 (14%)	8 (3%)	6	42
1	C	282/324 (87%)	237 (84%)	33 (12%)	12 (4%)	3	32
1	D	282/324 (87%)	230 (82%)	45 (16%)	7 (2%)	6	44
2	a	13/72 (18%)	11 (85%)	0	2 (15%)	0	4
2	b	13/72 (18%)	13 (100%)	0	0	100	100
2	c	31/72 (43%)	31 (100%)	0	0	100	100
2	d	31/72 (43%)	31 (100%)	0	0	100	100
All	All	1216/1584 (77%)	1020 (84%)	158 (13%)	38 (3%)	8	40

All (38) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	30	MET
1	B	30	MET
1	B	107	LYS
1	B	157	LYS
1	D	99	VAL
1	A	91	ILE
1	A	99	VAL
1	A	172	GLY
1	A	247	GLY
1	B	99	VAL
1	C	91	ILE
1	D	30	MET
1	D	132	PRO
1	A	308	ALA
2	a	59	THR
1	B	122	THR
1	A	82	ALA

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	a	69	ARG
1	B	132	PRO
1	B	155	PHE
1	C	30	MET
1	C	64	GLU
1	C	272	THR
1	C	308	ALA
1	D	272	THR
1	A	272	THR
1	C	67	GLU
1	C	190	GLU
1	C	276	GLU
1	D	212	PRO
1	D	258	GLN
1	A	258	GLN
1	D	113	PRO
1	C	29	VAL
1	C	172	GLY
1	B	111	ILE
1	C	264	ILE
1	C	273	VAL

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	249/286 (87%)	230 (92%)	19 (8%)	15	52
1	B	249/286 (87%)	236 (95%)	13 (5%)	27	64
1	C	249/286 (87%)	229 (92%)	20 (8%)	14	50
1	D	249/286 (87%)	236 (95%)	13 (5%)	27	64
2	a	13/66 (20%)	12 (92%)	1 (8%)	15	51
2	b	13/66 (20%)	12 (92%)	1 (8%)	15	51
2	c	31/66 (47%)	29 (94%)	2 (6%)	20	57
2	d	31/66 (47%)	31 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	1084/1408 (77%)	1015 (94%)	69 (6%)	25 57

All (69) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	31	MET
1	A	53	GLU
1	A	57	ILE
1	A	64	GLU
1	A	72	LYS
1	A	74	THR
1	A	90	LYS
1	A	98	TRP
1	A	107	LYS
1	A	116	LYS
1	A	133	MET
1	A	149	ASN
1	A	153	ASN
1	A	155	PHE
1	A	211	ASP
1	A	217	ARG
1	A	221	ARG
1	A	262	TYR
1	A	304	ASP
2	a	64	LEU
2	b	64	LEU
1	B	31	MET
1	B	45	THR
1	B	57	ILE
1	B	64	GLU
1	B	71	LYS
1	B	90	LYS
1	B	110	VAL
1	B	115	THR
1	B	120	ASN
1	B	157	LYS
1	B	194	ASN
1	B	207	ARG
1	B	217	ARG
1	C	31	MET
1	C	33	GLU
1	C	40	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	57	ILE
1	C	58	MET
1	C	64	GLU
1	C	72	LYS
1	C	74	THR
1	C	90	LYS
1	C	99	VAL
1	C	100	ASN
1	C	107	LYS
1	C	111	ILE
1	C	116	LYS
1	C	155	PHE
1	C	189	ASP
1	C	258	GLN
1	C	264	ILE
1	C	270	LEU
1	C	299	LEU
2	c	51	GLN
2	c	72	LYS
1	D	31	MET
1	D	53	GLU
1	D	57	ILE
1	D	64	GLU
1	D	67	GLU
1	D	99	VAL
1	D	100	ASN
1	D	155	PHE
1	D	157	LYS
1	D	175	THR
1	D	254	TYR
1	D	285	GLU
1	D	299	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	194	ASN
1	C	153	ASN
1	C	258	GLN
1	D	286	GLN
2	d	62	GLN

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.