



## wwPDB EM Validation Summary Report ⓘ

Dec 11, 2024 – 06:03 pm GMT

PDB ID : 9G93  
EMDB ID : EMD-45459  
Title : CryoET structure of the in vitro grown Bacillus anthracis Sap S-layer  
Authors : Sogues, A.; Leigh, K.; Van der Verren, S.; Kudryashev, M.; Pak, A.; Hal-  
ingstad, E.V.; Cecil, A.J.; Fioravanti, A.; Remaut, H.  
Deposited on : 2024-07-24  
Resolution : 7.20 Å(reported)

This is a wwPDB EM 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.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

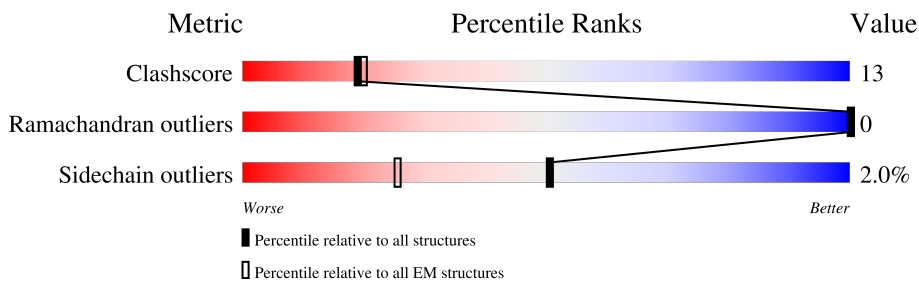
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 7.20 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	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	814	
1	B	814	
1	C	814	
1	D	814	
1	E	814	
1	F	814	
1	G	814	
1	I	814	
1	J	814	

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	K	814	 39% 7% 54%
1	L	814	 10% 89%

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 37178 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.

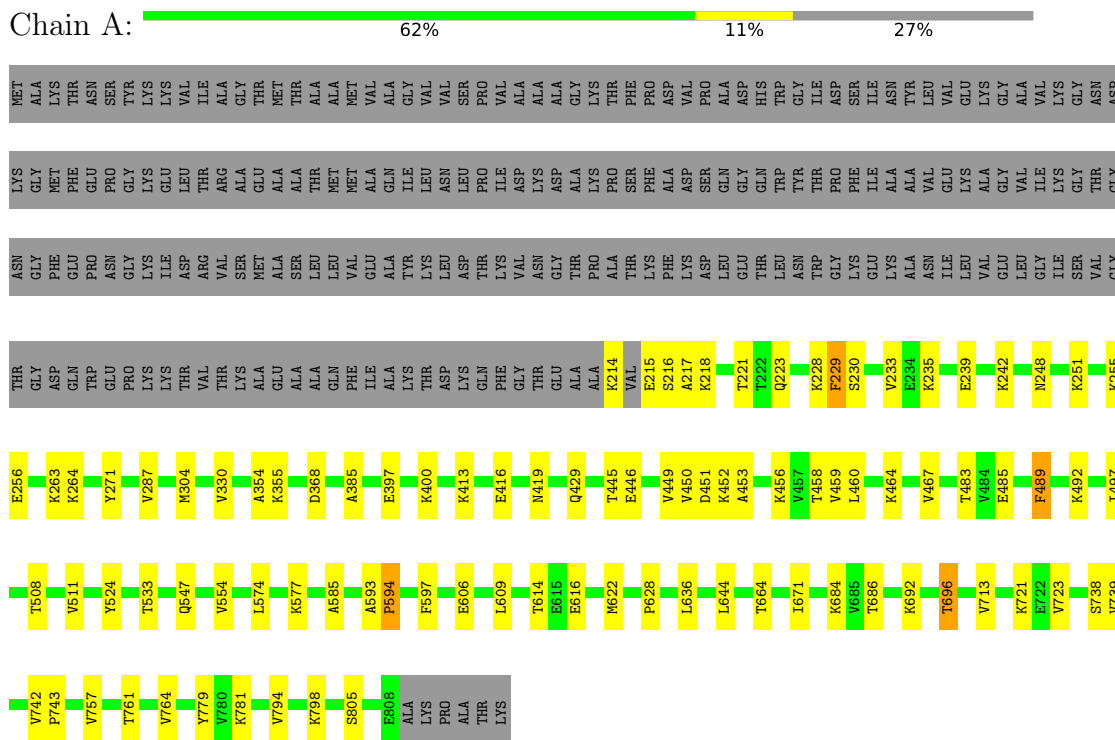
- Molecule 1 is a protein called S-layer protein sap.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	595	4456	2803	732	918	3	0	0
1	B	595	4456	2803	732	918	3	0	0
1	C	595	4456	2803	732	918	3	0	0
1	D	595	4456	2803	732	918	3	0	0
1	E	595	4456	2803	732	918	3	0	0
1	F	595	4456	2803	732	918	3	0	0
1	G	278	2097	1315	343	438	1	0	0
1	I	376	2804	1761	458	584	1	0	0
1	J	277	2088	1310	341	436	1	0	0
1	L	88	649	406	103	139	1	0	0
1	K	376	2804	1761	458	584	1	0	0

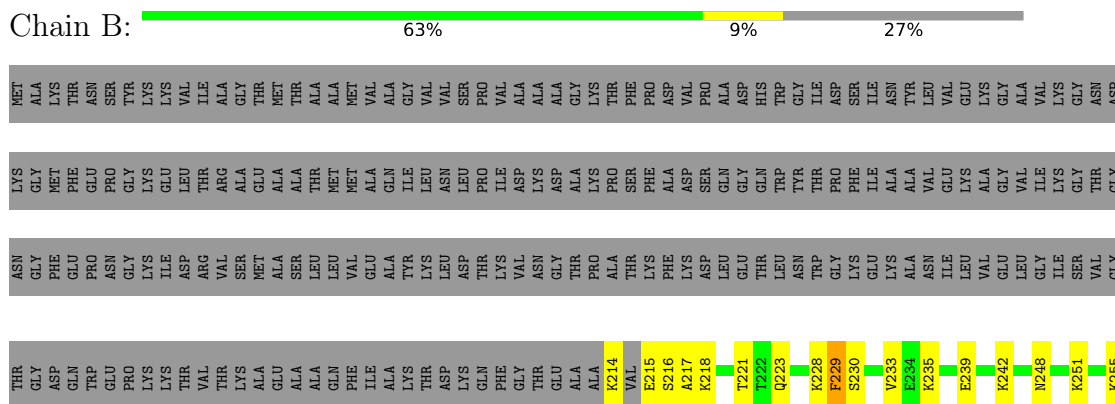
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: S-layer protein sap



- Molecule 1: S-layer protein sap

















## 4 Experimental information

Property	Value	Source
EM reconstruction method	SUBTOMOGRAM AVERAGING	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of subtomograms used	10126	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 ( $e^-/\text{\AA}^2$ )	156	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	81000	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  >5	RMSZ	# Z  >5
1	A	0.44	1/4495 (0.0%)	0.63	3/6079 (0.0%)
1	B	0.44	1/4495 (0.0%)	0.63	3/6079 (0.0%)
1	C	0.44	1/4495 (0.0%)	0.63	3/6079 (0.0%)
1	D	0.44	2/4495 (0.0%)	0.62	2/6079 (0.0%)
1	E	0.45	2/4495 (0.0%)	0.62	2/6079 (0.0%)
1	F	0.44	1/4495 (0.0%)	0.62	2/6079 (0.0%)
1	G	0.49	3/2115 (0.1%)	0.65	3/2857 (0.1%)
1	I	0.48	3/2827 (0.1%)	0.63	2/3825 (0.1%)
1	J	0.49	3/2106 (0.1%)	0.66	4/2845 (0.1%)
1	K	0.46	2/2827 (0.1%)	0.62	1/3825 (0.0%)
1	L	0.48	0/654	0.58	0/884
All	All	0.45	19/37499 (0.1%)	0.63	25/50710 (0.0%)

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	488	ALA	C-N	7.92	1.52	1.34
1	J	488	ALA	C-N	7.92	1.52	1.34
1	G	488	ALA	C-N	7.90	1.52	1.34
1	K	487	GLU	C-N	-7.16	1.17	1.34
1	E	487	GLU	C-N	-6.16	1.19	1.34

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	594	PRO	O-C-N	-7.91	109.75	123.20
1	A	594	PRO	O-C-N	-7.89	109.79	123.20
1	B	594	PRO	O-C-N	-7.82	109.91	123.20
1	F	594	PRO	O-C-N	-7.05	111.21	123.20
1	D	594	PRO	O-C-N	-6.83	111.60	123.20

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	4456	0	4624	180	0
1	B	4456	0	4624	110	0
1	C	4456	0	4624	134	0
1	D	4456	0	4624	88	0
1	E	4456	0	4624	210	0
1	F	4456	0	4624	157	0
1	G	2097	0	2157	37	0
1	I	2804	0	2891	86	0
1	J	2088	0	2149	54	0
1	K	2804	0	2891	81	0
1	L	649	0	663	4	0
All	All	37178	0	38495	950	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 13.

The worst 5 of 950 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:F:239:GLU:CB	1:F:742:VAL:CG1	1.77	1.62
1:B:239:GLU:CB	1:B:742:VAL:CG1	1.77	1.61
1:K:239:GLU:CB	1:K:742:VAL:CG1	1.77	1.60
1:I:239:GLU:CB	1:I:742:VAL:CG1	1.77	1.59
1:C:239:GLU:CB	1:C:742:VAL:CG1	1.77	1.58

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	593/814 (73%)	572 (96%)	21 (4%)	0	100	100
1	B	593/814 (73%)	572 (96%)	21 (4%)	0	100	100
1	C	593/814 (73%)	572 (96%)	21 (4%)	0	100	100
1	D	593/814 (73%)	572 (96%)	21 (4%)	0	100	100
1	E	593/814 (73%)	572 (96%)	21 (4%)	0	100	100
1	F	593/814 (73%)	572 (96%)	21 (4%)	0	100	100
1	G	276/814 (34%)	269 (98%)	7 (2%)	0	100	100
1	I	372/814 (46%)	364 (98%)	8 (2%)	0	100	100
1	J	275/814 (34%)	268 (98%)	7 (2%)	0	100	100
1	K	372/814 (46%)	364 (98%)	8 (2%)	0	100	100
1	L	86/814 (11%)	81 (94%)	5 (6%)	0	100	100
All	All	4939/8954 (55%)	4778 (97%)	161 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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	491/659 (74%)	481 (98%)	10 (2%)	50	68
1	B	491/659 (74%)	481 (98%)	10 (2%)	50	68
1	C	491/659 (74%)	481 (98%)	10 (2%)	50	68
1	D	491/659 (74%)	481 (98%)	10 (2%)	50	68
1	E	491/659 (74%)	481 (98%)	10 (2%)	50	68
1	F	491/659 (74%)	481 (98%)	10 (2%)	50	68
1	G	233/659 (35%)	227 (97%)	6 (3%)	41	59
1	I	311/659 (47%)	306 (98%)	5 (2%)	58	73

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	J	232/659 (35%)	227 (98%)	5 (2%)	47	65
1	K	311/659 (47%)	306 (98%)	5 (2%)	58	73
1	L	71/659 (11%)	70 (99%)	1 (1%)	62	75
All	All	4104/7249 (57%)	4022 (98%)	82 (2%)	50	68

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	664	THR
1	J	215	GLU
1	F	696	THR
1	G	491	GLN
1	J	489	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	E	762	ASN

### 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 oligosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	G	1
1	I	1
1	J	1
1	E	1
1	K	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	G	489:PHE	C	490:ALA	N	1.20
1	I	489:PHE	C	490:ALA	N	1.20
1	J	489:PHE	C	490:ALA	N	1.20
1	E	487:GLU	C	488:ALA	N	1.19
1	K	487:GLU	C	488:ALA	N	1.17