



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 13, 2024 – 03:33 pm GMT

PDB ID : 8RW9
Title : Domains 1 and 2 of Bacillus anthracis Sap S-layer in complex with Nb694
Authors : Sogues, A.; Remaut, H.
Deposited on : 2024-02-02
Resolution : 1.99 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

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
Xtriage (Phenix) : 1.13
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.003 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

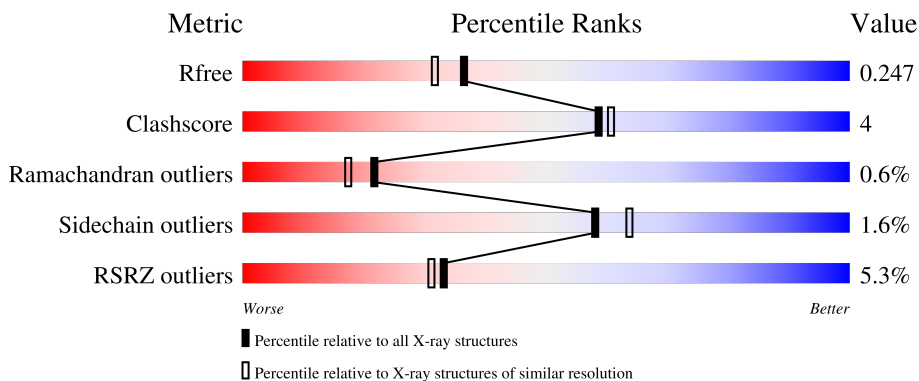
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.99 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.



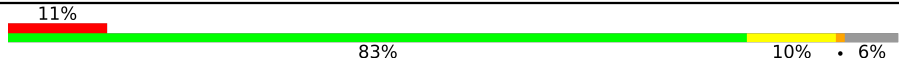


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	130	 3% 88% 8% .
1	B	130	 2% 86% 8% . 5%
1	C	130	 4% 89% 5% . 5%
1	D	130	 2% 92% . 5%
2	E	176	 9% 83% 11% . 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	176	 11% 83% 10% • 6%
2	G	176	 3% 88% 7% • 5%
2	H	176	 4% 88% 5% •• 5%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9666 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Nanobody694.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	125	971	613	164	190	4	0	0	0
1	B	124	961	607	161	189	4	0	0	0
1	C	123	956	604	161	187	4	0	0	0
1	D	123	952	602	160	186	4	0	0	0

- Molecule 2 is a protein called S-layer protein sap.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	E	167	1225	768	200	256	1	0	0	0
2	F	166	1233	774	200	258	1	0	0	0
2	G	168	1234	774	202	257	1	0	0	0
2	H	167	1241	779	203	258	1	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	208	MET	-	initiating methionine	UNP P49051
E	209	HIS	-	expression tag	UNP P49051
E	210	HIS	-	expression tag	UNP P49051
E	211	HIS	-	expression tag	UNP P49051
E	212	HIS	-	expression tag	UNP P49051
E	213	HIS	-	expression tag	UNP P49051
E	214	HIS	-	expression tag	UNP P49051
F	208	MET	-	initiating methionine	UNP P49051

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	209	HIS	-	expression tag	UNP P49051
F	210	HIS	-	expression tag	UNP P49051
F	211	HIS	-	expression tag	UNP P49051
F	212	HIS	-	expression tag	UNP P49051
F	213	HIS	-	expression tag	UNP P49051
F	214	HIS	-	expression tag	UNP P49051
G	208	MET	-	initiating methionine	UNP P49051
G	209	HIS	-	expression tag	UNP P49051
G	210	HIS	-	expression tag	UNP P49051
G	211	HIS	-	expression tag	UNP P49051
G	212	HIS	-	expression tag	UNP P49051
G	213	HIS	-	expression tag	UNP P49051
G	214	HIS	-	expression tag	UNP P49051
H	208	MET	-	initiating methionine	UNP P49051
H	209	HIS	-	expression tag	UNP P49051
H	210	HIS	-	expression tag	UNP P49051
H	211	HIS	-	expression tag	UNP P49051
H	212	HIS	-	expression tag	UNP P49051
H	213	HIS	-	expression tag	UNP P49051
H	214	HIS	-	expression tag	UNP P49051

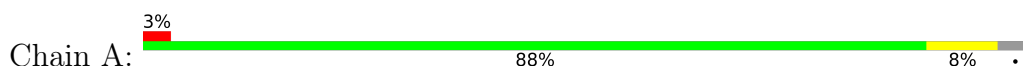
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	96	Total O 96 96	0	0
3	B	79	Total O 79 79	0	0
3	C	88	Total O 88 88	0	0
3	D	116	Total O 116 116	0	0
3	E	137	Total O 137 137	0	0
3	F	118	Total O 118 118	0	0
3	G	136	Total O 136 136	0	0
3	H	123	Total O 123 123	0	0

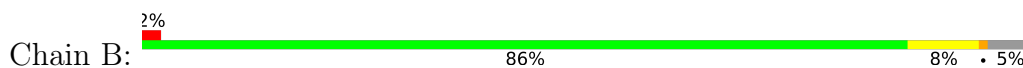
3 Residue-property plots [i](#)

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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.

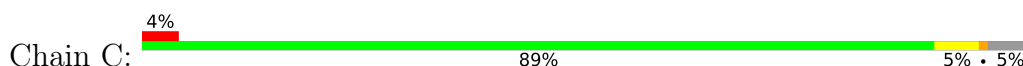
- Molecule 1: Nanobody694



- Molecule 1: Nanobody694



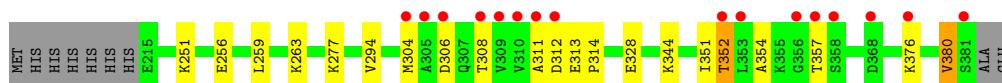
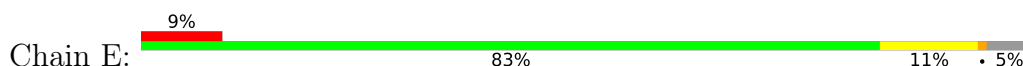
- Molecule 1: Nanobody694




- Molecule 1: Nanobody694

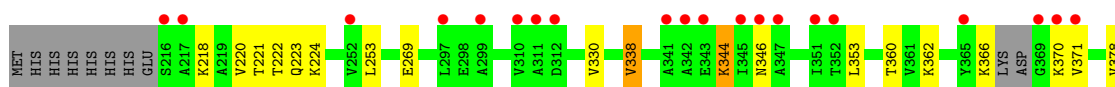


- Molecule 2: S-layer protein sap




- Molecule 2: S-layer protein sap

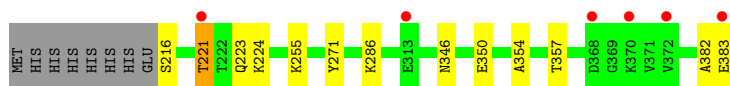
Chain F:  11% 83% 10% • 6%




E383

- Molecule 2: S-layer protein sap

Chain G:  3% 88% 7% • 5%



- Molecule 2: S-layer protein sap

Chain H:  4% 88% 5% • 5%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	33.41Å 113.97Å 122.70Å 90.71° 96.01° 97.05°	Depositor
Resolution (Å)	42.00 – 1.99 42.00 – 1.99	Depositor EDS
% Data completeness (in resolution range)	96.9 (42.00-1.99) 97.0 (42.00-1.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.00 (at 1.98Å)	Xtrriage
Refinement program	PHENIX 1.21_5207	Depositor
R, R_{free}	0.209 , 0.249 0.207 , 0.247	Depositor DCC
R_{free} test set	8875 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	48.4	Xtrriage
Anisotropy	0.304	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 52.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.055 for h,-h-k,-h-l 0.000 for -h,h+k,-l 0.000 for -h,-k,h+l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	9666	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.73% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.42	0/997	0.63	0/1354
1	B	0.40	0/986	0.65	1/1339 (0.1%)
1	C	0.39	0/981	0.65	1/1332 (0.1%)
1	D	0.47	0/977	0.66	0/1327
2	E	0.43	0/1233	0.68	0/1671
2	F	0.39	0/1240	0.61	0/1675
2	G	0.41	0/1242	0.59	0/1682
2	H	0.48	1/1248 (0.1%)	0.69	2/1685 (0.1%)
All	All	0.43	1/8904 (0.0%)	0.65	4/12065 (0.0%)

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
2	H	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	234	GLU	CB-CG	7.38	1.66	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	234	GLU	C-N-CA	7.67	140.87	121.70
1	C	11	LEU	CA-CB-CG	5.91	128.90	115.30
2	H	235	LYS	CD-CE-NZ	5.24	123.75	111.70
1	B	42	GLY	N-CA-C	-5.17	100.18	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	234	GLU	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	971	0	913	9	0
1	B	961	0	906	6	0
1	C	956	0	906	8	0
1	D	952	0	900	4	0
2	E	1225	0	1254	16	0
2	F	1233	0	1280	12	0
2	G	1234	0	1270	15	0
2	H	1241	0	1293	11	0
3	A	96	0	0	2	0
3	B	79	0	0	1	0
3	C	88	0	0	1	0
3	D	116	0	0	1	0
3	E	137	0	0	3	0
3	F	118	0	0	2	0
3	G	136	0	0	4	0
3	H	123	0	0	3	0
All	All	9666	0	8722	76	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 4.

The worst 5 of 76 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)
2:F:221:THR:HG22	2:F:223:GLN:H	1.19	1.02
2:G:221:THR:HG22	2:G:223:GLN:H	1.31	0.95
2:H:234:GLU:HB2	2:H:235:LYS:HE3	1.51	0.92
2:H:216:SER:N	3:H:401:HOH:O	2.22	0.70
2:F:338:VAL:HG22	2:F:362:LYS:HB3	1.75	0.69

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 X-ray entries followed by that with respect to entries of similar resolution.

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	123/130 (95%)	122 (99%)	1 (1%)	0	100	100
1	B	122/130 (94%)	120 (98%)	2 (2%)	0	100	100
1	C	121/130 (93%)	120 (99%)	1 (1%)	0	100	100
1	D	121/130 (93%)	120 (99%)	1 (1%)	0	100	100
2	E	165/176 (94%)	157 (95%)	6 (4%)	2 (1%)	11	6
2	F	162/176 (92%)	156 (96%)	4 (2%)	2 (1%)	11	6
2	G	166/176 (94%)	163 (98%)	3 (2%)	0	100	100
2	H	163/176 (93%)	158 (97%)	2 (1%)	3 (2%)	7	3
All	All	1143/1224 (93%)	1116 (98%)	20 (2%)	7 (1%)	22	17

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	352	THR
2	H	235	LYS
2	H	262	ASP
2	F	253	LEU
2	E	312	ASP

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 X-ray entries followed by that with respect to entries of similar resolution.

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	102/108 (94%)	101 (99%)	1 (1%)	73	78
1	B	101/108 (94%)	98 (97%)	3 (3%)	36	37
1	C	101/108 (94%)	101 (100%)	0	100	100
1	D	100/108 (93%)	100 (100%)	0	100	100
2	E	134/150 (89%)	133 (99%)	1 (1%)	81	86
2	F	137/150 (91%)	132 (96%)	5 (4%)	30	30
2	G	135/150 (90%)	134 (99%)	1 (1%)	81	86
2	H	138/150 (92%)	134 (97%)	4 (3%)	37	39
All	All	948/1032 (92%)	933 (98%)	15 (2%)	58	64

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	360	THR
2	H	315	THR
2	F	370	LYS
2	H	327	THR
2	H	233	VAL

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

Mol	Chain	Res	Type
1	A	1	GLN
1	C	1	GLN
1	C	39	GLN
1	C	83	ASN
1	D	13	GLN

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](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	125/130 (96%)	0.27	4 (3%) 50 48	42, 50, 67, 94	0
1	B	124/130 (95%)	0.35	2 (1%) 70 69	41, 51, 66, 97	0
1	C	123/130 (94%)	0.67	5 (4%) 42 40	43, 55, 67, 86	0
1	D	123/130 (94%)	0.19	2 (1%) 70 69	39, 47, 61, 74	0
2	E	167/176 (94%)	0.65	16 (9%) 15 14	39, 53, 75, 85	0
2	F	166/176 (94%)	0.79	20 (12%) 10 9	44, 60, 78, 118	0
2	G	168/176 (95%)	0.50	6 (3%) 46 44	43, 55, 73, 92	0
2	H	167/176 (94%)	0.45	7 (4%) 41 39	41, 54, 71, 89	0
All	All	1163/1224 (95%)	0.50	62 (5%) 33 31	39, 53, 74, 118	0

The worst 5 of 62 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	E	305	ALA	8.9
2	E	311	ALA	5.7
2	E	306	ASP	5.6
2	E	352	THR	5.2
2	F	370	LYS	4.3

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.