Supporting Information for

Architecture of the Sap S-layer of *Bacillus anthracis* revealed by integrative structural biology

Adrià Sogues#, 1, 2, Kendra Leigh#, 3, 4, 5, 6, Ethan V. Halingstad#, 7, Sander E. Van der Verren#, 1, 2, Adam J. Cecil7, Antonella Fioravanti1, 2, Alexander J. Pak7, 8, 9, Misha Kudryashev 3, 4, 5, 6, Han Remaut1, 2

1. Structural and Molecular Microbiology, VIB-VUB Center for Structural Biology, VIB, Pleinlaan 2, 1050, Brussels, Belgium

2. Structural Biology Brussels, Vrije Universiteit Brussel, VUB, Pleinlaan 2, 1050, Brussels, Belgium

3. Max Planck Institute of Biophysics, Frankfurt on Main, Germany

4. Buchmann Institute for Molecular Life Sciences, Goethe University of Frankfurt on Main, Frankfurt, Germany

5. In Situ Structural Biology, Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Berlin, Germany

6. Institute of Medical Physics and Biophysics, Charité-Universitätsmedizin Berlin, Berlin, Germany

7. Chemical and Biological Engineering, Colorado School of Mines, Golden, Colorado, USA

8. Quantitative Biosciences and Engineering, Colorado School of Mines, Golden, Colorado, USA

9. Materials Science Program, Colorado School of Mines, Golden, Colorado, USA

# contributed equally
Corresponding authors: han.remaut@vub.be, mikhail.kudryashev@mdc-berlin.de

This PDF file includes:

Supplementary Methods
Figures S1 to S3
Tables S1 to S5
Legend for Movie S1

Supplementary Methods

Construction Sap mutants.

Sap mutant T33D-S334A-T335V was derived from the Sap$_{FL}$ plasmid using site-directed mutagenesis with primers 543 and 544. Similarly, the cysteine mutant S334C plasmid was obtained using primers 470 and 471 with the Sap$_{FL}$ plasmid as a template. Individual Sap domains were also cloned into a linearized pASK-IBA3plus vector using Gibson assembly. Domain 1 (from E125 to S207) was amplified using the primers 373 and 549. Domain 1-2 (from E125 to G384) was amplified using the pair of primers 373 and 374. Domain 3 (from A385 to Q491) was amplified using the primers 375 and 376. Domain 4-5 (from K492 to K709) was amplified using the primers 379 and 550. Finally, Domain 6 (from G710 to P811) was amplified with primers 377 and 378. For cloning purposes, all the cultures were grown in LB at 37°C and supplemented with 100 μg/mL of Ampicillin when required. All plasmids were sequence-verified (Eurofins) using primers 305, 306, 385, 386 and 387.

Production and purification of Sap^{AD}

E. coli BL21 cells transformed with pAFSLP1 were grown at 37°C in LB medium in the presence of 0.1% glucose and 100 μg/ml ampicillin, induced with 10um IPTG at OD 0.6-0.8 and harvested by centrifugation after overnight induction. For cell lysis, 1-mL B-PER (ThermoFisher) was added to a 20-mL pellet resuspended in buffer A (20 mM Tris pH 8, 300 mM NaCl, 20 mM imidazole) supplemented with 2 mM MgCl$_2$, 0.1% Triton X-100, 0.1 mg/mL lysozyme, 0.5 mg/mL DNase and protease inhibitors (0.1 mg/mL 4-(2-aminoethyl) benzenesulfonyl fluoride and 1 μg/mL leupeptin). The lysate was incubated for 15 minutes at room temperature after which it was centrifuged at 4000 RPM at 4°C. Two times 600 μL of supernatant was loaded onto a His SpinTrap

column (Cytiva) pre-equilibrated with Buffer A by centrifugation for 30 s at 100 x g, followed by a single wash with buffer A centrifuged at 100 x g for 30 s. Next, 6xHis-tagged Sap was eluted once using 400 µL of Buffer B (50 mM Tris pH 8, 300 mM NaCl, 500 mM imidazole) after which it was buffer exchanged into 1x PBS using a Zeba™ Spin Desalting Column (7K MWCO, 0.5 mL, ThermoScientific) and normalized to 2 mg/mL for *in vitro* S-layer formation.

Production and purification of Sap^{FL,} Sap^{FL} mutants and Sap domains.

Sap^{FL}, Sap^{FL} mutants and individual Sap domains (all N-terminally tagged with 6xHis) were expressed in E. coli BL21 (DE3) grown in Terrific Broth (TB) supplemented with 100 µg/mL of Ampicillin at 37 °C and induced with 200 µg/L anhydrotetracycline when the OD_{600} reached 0.5. Following induction, the temperature was dropped to 23°C for Sap^{FL} (and mutants) and 28°C for the Sap domains. Cells were left to express the desired protein overnight and harvested the next day by centrifugation. Bacterial pellets were kept at -20°C until purification. Frozen pellet from Sap^{FL} was resuspended in 100 mL of lysis buffer (50 mM HEPES, pH 8, 400 mM NaCl, 1 mM $MgCl_2$, DNase, lysozyme and EDTA-free protease inhibitor cocktail (ROCHE)) at 4°C and lysed by sonication for 5 minutes. The lysate was centrifuged for 45 min at 30000 x g at 4°C and the pellet containing Sap^{FL} was resuspended in Buffer A-Urea (300 mM NaCl, 10 mM HEPES, pH 8, 10 mM imidazole and 8 M urea) followed by centrifugation for 1 h at 30000 x g at 4°C. At this point in the purification, Sap^{FL} is found in the soluble fraction, so the supernatant was mixed with 5 mL of 40 IDA^{high} agarose IMAC beads charged with Ni^{2+} (Bio-Works) pre-equilibrated with Buffer A-Urea and incubated at room temperature for 45 min with constant rotation. The beads containing Sap^{FL} were pelleted by centrifugation at 4000 x g for 10 minutes at 4°C and washed twice with 50

mL of Buffer A (30 mM NaCl, 10 mM HEPES, pH 8, 10 mM imidazole) to remove the urea. Finally, Sap^{FL} was eluted with 100% Buffer B (300 mM NaCl, 10 mM HEPES, pH 8, 1 M imidazole), filtered with a 0.45-µm filter (Acrodisc LC 13 mm, syringe filter, Life Science) and applied onto a Superdex 200 16/60 size-exclusion column (GE Life Sciences) that was equilibrated with 10 mM HEPES, pH 8 and 100 mM NaCl at 4°C. Fractions containing Sap^{FL} were pooled, immediately flash-frozen in liquid nitrogen and kept at −80 °C.

Sap^{FL} mutant T33D-S334A-T335V (non-polymerizing) and all individual Sap domains were purified as soluble proteins without denaturing agents. Frozen bacterial pellets were resuspended in 100 mL of lysis buffer and lysed by sonication for 10 minutes at 4°C. The lysate was centrifuged for 30 min at 30000 x g at 4°C. The clarified supernatant containing the 6xHis tagged protein was loaded onto a 5-mL Ni-NTA affinity chromatography column (HisTrap FF crude, GE Healthcare). The column was washed with 5 column volumes using Buffer A and proteins were eluted with a linear gradient of Buffer B. Eluted protein was concentrated to 7 mL and loaded onto a Superdex 200 16/60 size-exclusion column (GE Life Sciences) that was equilibrated with SEC buffer (10 mM HEPES, pH 8 and 150 mM NaCl) at 4°C. The peak corresponding to the protein of interest was concentrated, flash-frozen in small aliquots in liquid nitrogen and stored at −80°C. After purification, all samples were run on a sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gel to evaluate their purity.

### Sap^{AD}-gold labelling and cryo-EM analysis

The C-terminal His-tag of Sap^{AD} was functionalised with 1.8-nm Ni-NTA-nanogold beads (nanoprobes). 25 µM of matured Sap^{AD}-tubules was mixed with 30 µM Ni-NTA-

nanogold, incubated for 30 min at room temperature and assessed with cryo-EM. For grid preparation, 3 µL of this mixture was spotted on Quantifoil R2/1 mesh 400 grids (Quantifoil), back-blotted manually for 1.5 s on a CP3 cryoplunge (Gatan), maintained at 100% humidity and room temperature, and vitrified using liquid ethane. The sample was imaged on a 300-kV CryoARM300 (JEOL) equipped with an omega energy filter and K3 (Gatan) direct electron detector at the VIB-VUB facility for Bio Electron Cryogenic Microscopy (BECM), Brussels. Data were collected at a nominal magnification of 60.000x corresponding to a pixel size of 0.764 A, at -1.5 to -3 μm defocus, ~60 e$^-$/A^2 in 59 frames, and 1125 movies. Images were motion-corrected and dose-weighted using MotionCor2.1 (35) with 7x5 patches, and frame grouping by 3. Contrast transfer function fitting was performed using CTFFIND4 v4.1 (45). Lattice maps were calculated in real space using RELION v3.1 (40) for manual picking, particle extraction and two-dimensional classification. In total, 102 tubules/sheets were picked and spread over 47 images. Multiple rows of unit cells were extracted per object using the helical extraction function in Relion and oversampled by 60 (cropping 600A boxes every 10 A), resulting in 32254 particles. For best alignment, particles were extracted as 800-pixel box size, binned by 2 and aligned using a mask of 600A to visualize multiple unit cells at once.

Dynamic light scattering (DLS)

To assess the presence or absence of S-layer, we used dynamic light scattering (DLS). First, protein samples were centrifuged at 20000 x g for 20 min and the concentration was adjusted to 2 mg/mL. The experiments were collected at 25°C in 4-µL cyclic olefin copolymer disposable cuvettes at an angle of 90° using a Dynapro

NanoStar DLS machine (Wyatt Technology). The Dynamics software (version 7.1.9.3) was used to schedule data acquisition and data analysis. For each sample, 20 measurements of 10 s were averaged, and this operation was repeated 20 times for each condition. The plots represent the distribution of the intensity over the radius of the particles.
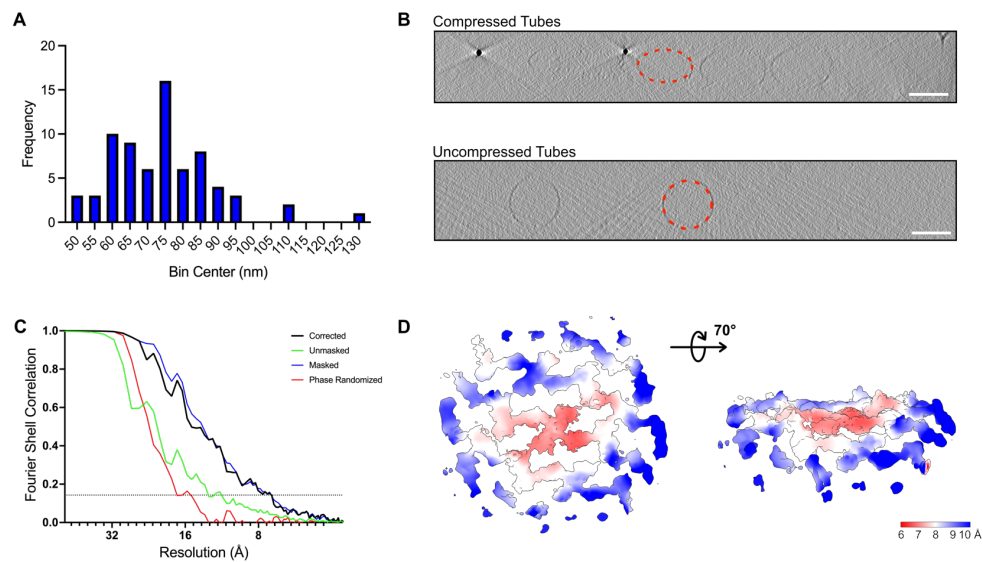
Supplementary Figure 1. A. Histogram of tomogram tube widths in nanometers. **B.** Representative slices through two different tomograms demonstrating how some tubes have compressed cross-sections and others have round cross-sections (dotted red lines). Scale bar: 50 nm **C.** Fourier shell correlation for the final cryo-ET density as generated by RELION 3.1 post-processing. **D.** Cryo-ET map colored by local resolution as calculated by RELION 3.1.
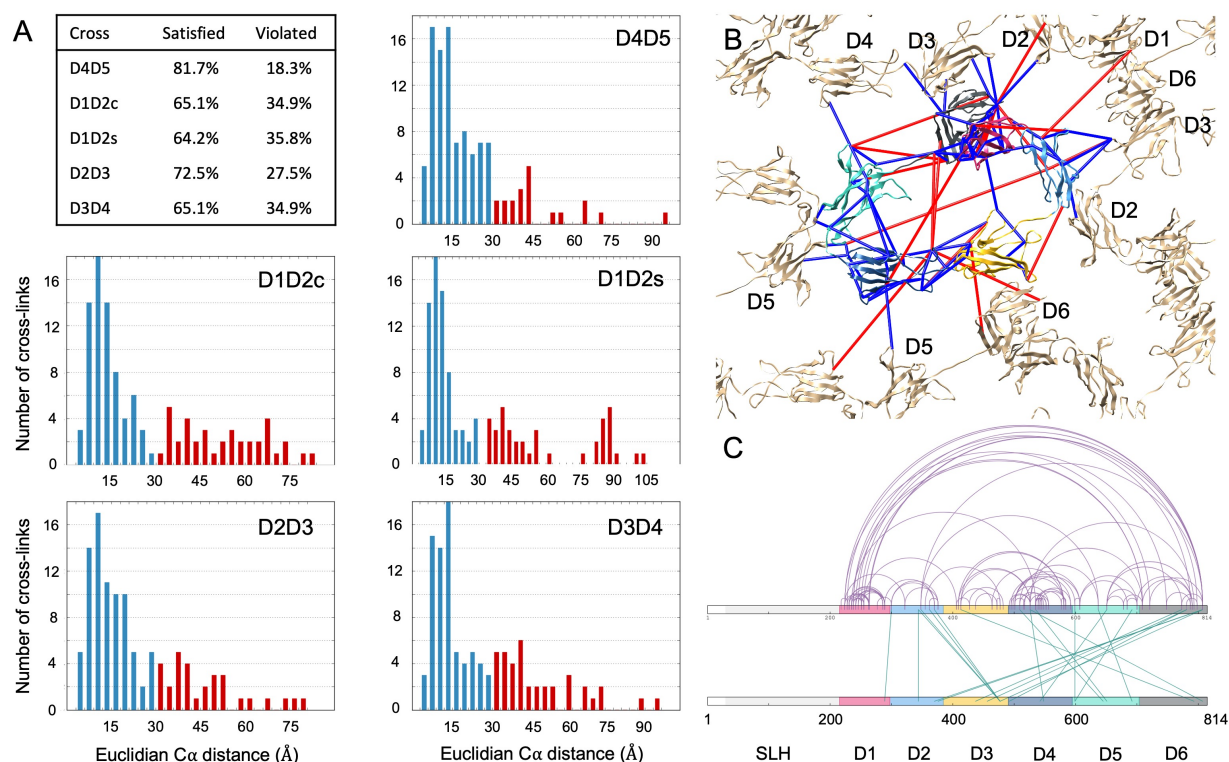
Supplementary Figure 2: Sap^AD lysine crosslinking mass spectrometry. A. Crosslinking pairs evaluated against plausible lattice models defined by the two domains forming the cross feature. Satisfied crosslinks are indicated in blue, while violations of the allowed distance (<30 Å) are noted in red. A top-right table summarizes the results. **B**. Ribbon representation of the model where the cross is formed by D4D5. Blue lines represent crosslinks satisfying the <30 Å distance constraint, whereas red lines indicate violations. **C**. Domain representation of the Sap protein. Intramolecular crosslinks are shown in purple, and intermolecular crosslinks are shown in green.
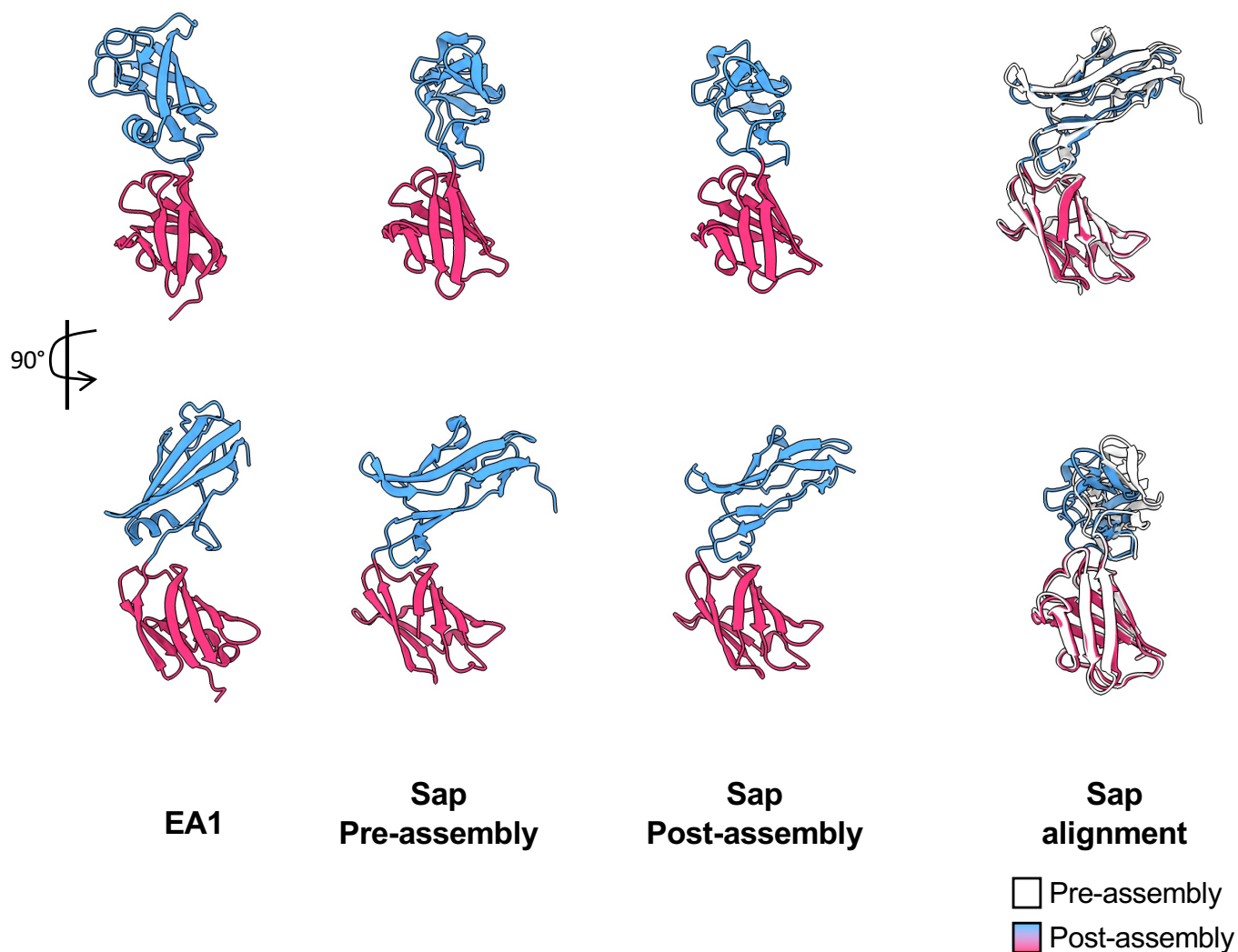
Supplementary Figure 3. **Structural comparison of Sap and EA1 D1D2 domains**. Ribbon representation of domains 1 and 2 in blue and pink respectively. Sap is shown in the pre-assembly (PDB 6HHU) and post-assembly state and also as a superposition aligned at domain two. EA1 (PDB 8OPR) is also shown.

Data Collection	
Microscope	FEI Titan Krios G1
Voltage	300 kV
Camera	Gatan K2 Summit® Direct Electron Detector
Energy Filter	Gatan GIF Quantum® SE energy filter
Objective Aperture	70 µm
Nominal Magnification	81000x
Pixel Size	0.9 Å (super-resolution)
Tilt Scheme	Dose Symmetric
Set Defocus	-2, -3, or -3.5 µm
Tilt Range	-60° to 60°
Total Electron Dose	156.6 e⁻/Å² (except for two tomograms, which had doses of 178.2 e⁻/Å² and 130.5 e⁻/Å², respectively)
Number of Tomograms Collected	12 (3 with a high dose 0° tilt)
Data Processing	
No. of Particles in Final Reconstruction	10126
Final Resolution (0.143 FSC, masked)	7.2 Å
Imposed Symmetry	C1
Final Pixel size	1.8 Å/pix (bin 2 from super-resolution)
Number of Tomograms Contributing	7 (no high dose tomograms included)

Supplementary table 1. Cryo-EM statistics

	Sap domains D3+D6	Sap domains D1D2	Sap domains D4D5
Data collection			
Space group	P 1 21 1	P 1 21 1	P1
Cell dimensions			
a, b, c (Å)	40.15, 61.01, 97.65	48.44, 128.70, 55.12	59.304, 63.539, 73.659
α, β, γ (°)	90, 91.5, 90	90, 95.76, 90	68.312, 81.61, 82.741
Resolution (Å)*	51.74 - 2.069 (2.14 - 2.07)	38.57 - 2.014 (2.086 - 2.014)	37.32 - 2.48 (2.569 - 2.48)
R_{pim}	0.069 (1.039)	0.033 (0.268)	0.045 (0.304)
R_{merge}	0.169 (2.60)	0.082 (0.672)	0.073 (0.493)
I/σ(I)	8.1 (0.9)	13 (2.4)	9.6 (2.2)
Completeness (%)	100 (100)	99.91 (99.93)	98.2 (98.3)
Multiplicity	6.9 (7.1)	7.1 (7.2)	3.6 (3.6)
Refinement			
Resolution (Å)	2.07	2.01	2.48
No. reflections	28947 (2868)	44265 (4388)	34266 (3434)
R-work/R-free	0.20/0.25	0.201/0.236	0.219/0.265
Number of atoms	3424	5657	6493
protein	3115	5064	6304
ligands/ions	0	0	0
water	309	593	90
B-factors (Å^2)	46.04	35.16	73.67
protein	45.83	34.62	73.84
ligands/ions	/	/	/
water	48.14	39.79	61.43
RMS deviations			
Bond length (Å)	0.008	0.008	0.008
Bond angles (°)	0.96	0.90	1.01
Ramachandran favored (%)	98.30	99.85	98.47
Ramachandran outliers (%)	0.24	0.15	1.53
PDB code	**8S83**	**8RX2**	**8S80**

* Values in parenthesis refer to the highest recorded resolution shell.

Supplementary table 2. X-ray crystallography statistics

Supplementary table 3. Crosslinking MS statistics

Crosslinked peptides	Residue 1	Residue 2	# Monomer	# Polymer	# total
SAKAVTTQK-VEVKFSK-a3-b4	218	228	4	4	8
AVTTQKVEVK-EAKPATK-a6-b3	224	810	5	1	6
VEVKFSK-EAKPATK-a4-b3	228	810	6	0	6
ANNDKVLVK-EDIKVTNK-a5-b4	251	242	11	11	22
ANNDKVLVK-LTKEDIK-a5-b3	251	238	9	6	15
VLVKEVTLSEDK-ADFTSKDFK-a4-b6	255	406	0	6	6
VLVKEVTLSEDK-AVTTQKVEVK-a4-b6	255	224	5	6	11
VLVKEVTLSEDK-LTKEDIK-a4-b3	255	238	9	12	21
EVTLSEDKK-AVEKLTK-a8-b4	263	235	4	4	8
EVTLSEDKK-FSKAVEK-a8-b3	263	231	5	5	10
KSATVELYSNLAAK-EAKPATK-a1-b3	264	810	4	0	4
KSATVELYSNLAAK-FSKAVEK-a1-b3	264	231	2	3	5
KSATVELYSNLAAK-PATKHHHHHH-a1-b4	264	814	0	8	8
KSATVELYSNLAAK-VEVKFSK-a1-b4	264	228	5	9	14
QTYTVDVNKVGK-AVEKLTK-a9-b4	286	235	8	11	19
QTYTVDVNKVGK-EDIKVTNK-a9-b4	286	242	8	9	17
QTYTVDVNKVGK-FSKAVEK-a9-b3	286	231	7	8	15
VGKTEVAVGSLEAK-ANNDKVLVK-a3-b5	289	251	10	9	19
VGKTEVAVGSLEAK-EDIKVTNK-a3-b4	289	242	7	9	16
VGKTEVAVGSLEAK-FSKAVEK-a3-b3	289	231	4	5	9
VGKTEVAVGSLEAK-VTNKANNDK-a3-b4	289	246	5	9	14
VGKTEVAVGSLEAKTIEMADQTVVADEPTALQFTVK-DGKVVAESK-a3-b3	289	370	0	5	5
TEVAVGSLEAKTIEMADQTVVADEPTALQFTVK-DGKVVAESK-a11-b3	300	370	8	117	125
TEVAVGSLEAKTIEMADQTVVADEPTALQFTVK-QTYTVDVNKVGKTEVAVGSLEAK-a11-b12	300	289	1	7	8
TEVAVGSLEAKTIEMADQTVVADEPTALQFTVK-SATVELYSNLAAKQTYTVDVNK-a11-b13	300	277	17	106	123
TEVAVGSLEAKTIEMADQTVVADEPTALQFTVK-VVAESKEVK-a11-b6	300	376	2	3	5
TIEMADQTVVADEPTALQFTVKDENGTEVVSPEGIEFVTPAAEK-AVTTQKVEVK-a22-b6	322	224	10	7	17
TIEMADQTVVADEPTALQFTVKDENGTEVVSPEGIEFVTPAAEK-VVAESKEVK-a22-b6	322	376	3	5	8
DENGTEVVSPEGIEFVTPAAEKINAK-DENGTEVVSPEGIEFVTPAAEKINAK-a22-b22	344	344	0	4	4
DENGTEVVSPEGIEFVTPAAEKINAK-GKELVSK-a22-b2	344	477	3	2	5
INAKGEITLAK-GKELVSK-a4-b2	348	477	2	2	4
GEITLAKGTSTTVK-GKELVSK-a7-b2	355	477	3	6	9
GTSTTVKAVYK-GKELVSK-a7-b2	362	477	0	7	7
GTSTTVKAVYK-VVAESKEVK-a7-b6	362	376	2	6	8
DGKVVAESK-GKELVSK-a3-b2	370	477	0	6	6
DFKQNNK-SGEKEAK-a3-b4	409	581	2	4	6
QNNKVYEGDNAYVQVELK-APVLDQYGKEFTAPVTVK-a4-b9	413	526	2	15	17
QNNKVYEGDNAYVQVELK-EQKVEFDK-a4-b3	413	793	1	7	8
QNNKVYEGDNAYVQVELK-SGEKEAK-a4-b4	413	581	5	3	8
QNNKVYEGDNAYVQVELK-VTVLSAGKAPVK-a4-b8	413	464	0	6	6
VYEGDNAYVQVELKDQFNAVTTGK-ADFTSKDFK-a14-b6	427	406	1	6	7
VYEGDNAYVQVELKDQFNAVTTGK-ELVSKTVEIEAFAQK-a14-b5	427	482	11	306	317
DQFNAVTTGKVEYESLNTEVAVVDK-VTVKDSK-a10-b4	437	472	2	2	4
VTVLSAGKAPVK-SGEKEAK-a8-b4	464	581	5	6	11
VTVKDSK-GKELVSK-a4-b2	472	477	3	3	6
LEKTNVALSTK-GLEKELDK-a3-b4	501	607	4	5	9
TNVALSTKDVTDLK-EQKLEAK-a8-b3	509	548	5	8	13
TNVALSTKDVTDLK-GAEAAELKVTTTNK-a8-b8	509	645	0	19	19
VKAPVLDQYGK-GLEKELDK-a2-b4	517	607	2	2	4
VKAPVLDQYGK-LEKTNVALSTK-a2-b3	517	501	0	6	6
APVLDQYGKEFTAPVTVK-AMKEIK-a9-b3	526	495	2	2	4
APVLDQYGKEFTAPVTVK-DFKQNNK-a9-b3	526	409	2	5	7
APVLDQYGKEFTAPVTVK-EAKPATKHHHHHH-a9-b7	526	814	0	7	7
APVLDQYGKEFTAPVTVK-EQKLEAK-a9-b3	526	548	0	5	5
APVLDQYGKEFTAPVTVK-EQKVEFDK-a9-b3	526	793	2	3	5
APVLDQYGKEFTAPVTVK-TVEIEAFAQKAMK-a9-b10	526	492	4	14	18
EFTAPVTVKVLDK-AVEKLTKEDIK-a9-b4	535	235	0	5	5
EFTAPVTVKVLDK-DGKELK-a9-b3	535	542	6	6	12

EFTAPVTVKVLDK-EAKATLALELK-a9-b3	535	584	46	336	382
EFTAPVTVKVLDK-ELKEQK-a9-b3	535	545	5	6	11
EFTAPVTVKVLDK-LEAKYVNK-a9-b4	535	552	2	5	7
EFTAPVTVKVLDK-VTTTNKEGK-a9-b6	535	651	0	8	8
ELKEQK-VLDKDGK-a3-b4	545	539	3	7	10
LEAKYVNK-DVTDLKVK-a4-b6	552	515	0	6	6
YVNKELVLNAAGQEAGNYTVVLTAK-APVLDQYGKEFTAPVTVK-a4-b9	556	526	1	41	42
YVNKELVLNAAGQEAGNYTVVLTAK-EIKLEK-a4-b3	556	498	1	16	17
ELVLNAAGQEAGNYTVVLTAKSGEK-EAKATLALELK-a21-b3	577	584	0	7	7
ELVLNAAGQEAGNYTVVLTAKSGEK-EFTAPVTVKVLDK-a21-b9	577	535	4	222	226
SGEKEAK-AMKEIK-a4-b3	581	495	6	17	23
EAKATLALELK-DGKELK-a3-b3	584	542	2	6	8
ATLALELKAPGAFSK-LEKTNVALST-a8-b3	592	501	3	12	15
ATLALELKAPGAFSK-VLDKDGK-a8-b4	592	539	4	5	9
APGAFSKFEVR-APGAFSKFEVR-a7-b7	599	599	0	9	9
APGAFSKFEVR-GLEKELDK-a7-b4	599	607	0	6	6
,APGAFSKFEVR-LEKTNVALSTK-a7-b3	599	501	4	6	10
FEVRGLEKELDK-LEKTNVALSTK-a8-b3	607	501	6	0	6
NAMTVSVLPVDANGLVLKGAEAAELK-TNVALSTKDVTDLK-a18-b8	637	509	49	18	67
NAMTVSVLPVDANGLVLKGAEAAELK-VTVVLDGKLITTHSFK-a18-b8	637	692	0	338	338
NAMTVSVLPVDANGLVLKGAEAAELKVTTTNK-INAKGEITLAK-a18-b4	637	348	0	6	6
VTTTNKEGK-DGKEQK-a6-b3	651	790	6	6	12
EGKEVDATDAQVTVQNNSVITVGQGAK-DGKELK-a3-b3	654	542	0	5	5
EVDATDAQVTVQNNSVITVGQGAKAGETYK-VTTTNKEGK-a24-b6	678	651	13	60	73
EVDATDAQVTVQNNSVITVGQGAKAGETYK-VVDTAPTAKGLAVEFTSTSLK-a24-b9	678	709	0	4	4
AGETYKVTVVLDGK-GLEKELDK-a6-b4	684	607	4	7	11
VTVVLDGKLITTHSFK-APGAFSKFEVR-a8-b7	692	599	0	27	27
LITTHSFKVVDTAPTAK-AGETYKVTVVLDGK-a8-b6	700	684	3	16	19
LITTHSFKVVDTAPTAK-DGKEQK-a8-b3	700	790	4	4	8
LITTHSFKVVDTAPTAK-FEVRGLEKELDK-a8-b8	700	607	10	8	18
LITTHSFKVVDTAPTAK-VTTTNKEGK-a8-b6	700	651	6	5	11
VVDTAPTAKGLAVEFTSTSLK-APVLDQYGKEFTAPVTVK-a9-b9	709	526	2	6	8
VVDTAPTAKGLAVEFTSTSLK-DGKEQK-a9-b3	709	790	4	5	9
VVDTAPTAKGLAVEFTSTSLK-EDIKVTNK-a9-b4	709	242	0	4	4
VVDTAPTAKGLAVEFTSTSLK-LTKEDIK-a9-b3	709	238	0	4	4
GLAVEFTSTSLKEVAPNADLK-AVQVAVSIKEAK-a12-b9	721	807	1	5	6
GLAVEFTSTSLKEVAPNADLK-EAKPATK-a12-b3	721	810	4	4	8
GLAVEFTSTSLKEVAPNADLK-MSAKAVTTQK-a12-b4	721	218	4	0	4
GLAVEFTSTSLKEVAPNADLK-PATKHHHHHH-a12-b4	721	814	0	7	7
AALLNILSVDGVPATTAKATVSNVEFVSADTNVVAENGTVGAK-NLTVVKDGK-a18-b6	748	787	4	175	179
AALLNILSVDGVPATTAKATVSNVEFVSADTNVVAENGTVGAK-VLVKEVTLSEDK-a18-b4	748	255	0	21	21
ATVSNVEFVSADTNVVAENGTVGAKGATSIYVK-DQFNAVTTGKVEYESLNTEVAVVDK-a25-b10	773	437	3	148	151
ATVSNVEFVSADTNVVAENGTVGAKGATSIYVK-EGKEVDATDAQVTVQNNSVITVGQGAK-a25-b3	773	654	0	4	4
ATVSNVEFVSADTNVVAENGTVGAKGATSIYVK-PATKHHHHHH-a25-b4	773	814	0	7	7
GATSIYVKNLTVVK-ATGKVTVLSAGK-a8-b4	781	456	0	5	5
GATSIYVKNLTVVK-TVEIEAFAQKAMK-a8-b10	781	492	0	6	6
NLTVVKDGK-LTKEDIK-a6-b3	787	238	0	6	6
VEFDKAVQVAVSIK-ANNDKVLVK-a5-b5	798	251	0	5	5
VEFDKAVQVAVSIK-GATSIYVKNLTVVK-a5-b8	798	781	0	132	132
AVQVAVSIKEAK-PATKHHHHHH-a9-b4	807	814	0	9	9
PATKHHHHHH-AVTTQKVEVK-a4-b6	814	224	6	5	11
PATKHHHHHH-DGKVVAESK-a4-b3	814	370	0	8	8
PATKHHHHHH-EVTLSEDKK-a4-b8	814	263	0	8	8
PATKHHHHHH-GKELVSK-a4-b2	814	477	0	6	6
PATKHHHHHH-INAKGEITLAK-a4-b4	814	348	0	4	4
PATKHHHHHH-VTNKANNDK-a4-b4	814	246	0	5	5
PATKHHHHHH-VVAESKEVK-a4-b6	814	376	3	5	8

		Clonning of Sap FL	
321	F_open_pASK_3	TAATAAGCTTGACCTGTGAAG	
322	R_open_pASK_3	CATTTGTATATCTCCTTCTTAAAG	
383	F-pASK-His-Sap	GGAGATATACAAATGCATCACCATCATCATCACAAAACATTCCCAGACGTTCC	
384	R-pASK-Sap	CAGGTCAAGCTTAttaTTTTGTTGCAGGTTTTGCTTC	
385	Sap_seq1	ACTGTAACTAAAGCAGAAGCTG	
386	Sap_seq2	AGACGGTAAAGTAGTAGCTGAAAG	
387	Sap_seq3	AAGCTCCAGTACTAGATCAATACG	
305	F_seq_pASK	GAGTTATTTTACCACTCCCT	
306	R_seq_pASK	CGCAGTAGCGGTAAACG	
		Clonning of SapFL triple mutant T33D-S334A-T335V	
456	F_Sap-A148	GTGATGCAGTTACTGTAAAAGCTGTTTATAAAAA	
457	R_Sap-A148	AACTGCATCACCTTTTGCTAAAGTGATTT	
		Clonning of Cys mutant (S334C)	
470	Sap_FL_S334C	GTACTTGCACTACTGTAAAAGCTGTTTATAAAA	
471	Sap_FL_S334C	GTAGTGCAAGTACCTTTTGCTAAAGTGATTT	
		Clonning of Domain 1	
373	F_D1-D2	GGAGATATACAAATGCATCACCATCATCATCACGAGTCCGCGAAGGCCGTGAC	
549	R_SapD1	CAGGTCAAGCTTAttaAGAACCCACTGCTACCTCAGTCTT	
		Clonning of Domain 1-2	
373	F_D1-D2	GGAGATATACAAATGCATCACCATCATCATCACGAGTCCGCGAAGGCCGTGAC	
374	R_D1-D2	CAGGTCAAGCTTAttaACCTTCCGCGCTTACCTTTAC	
		Clonning of Domain 3	
375	F_D3	GGAGATATACAAATGCATCACCATCATCATCACGCGGCAGTTGCGTCTATTTCTAA	
376	R_D3	CAGGTCAAGCTTAttaCTGTGCGAATGCTTCGATCTCAAC	
		Clonning of Domain 4-5	
379	F_D4	GGAGATATACAAATGCATCACCATCATCATCACAAAGCGATGAAGGAAATTAAGCTGGAG	
550	R_SapD45	CAGGTCAAGCTTAttaCGTGTCCACGACTTTGAAGCTATGCG	
		Clonning of Domain 6	
377	F_D6	GGAGATATACAAATGCATCACCATCATCATCACGGCCTGGCAGTAGAATTTACCAG	
378	R_D6	CAGGTCAAGCTTAttaTGGTTTGGCCTCCTTAATGGAAAC	

Supplementary table 4. Oligonucleotides table used in this study

Name	Vector family	Inducer	Resistance	Overexpressing gene	Reference
pAFSLP1	pET300	IPTG	Ampi	Sap-AD	Fioravanti et al. 2019
A110	pASK-IBA3plus	anhydrotetracycline	Ampi	Sap-FL	This study
A148	pASK-IBA3plus	anhydrotetracycline	Ampi	Sap FL T33D-S334A-T335V	This study
A133	pASK-IBA3plus	anhydrotetracycline	Ampi	Sap FL S334C	This study
A152	pASK-IBA3plus	anhydrotetracycline	Ampi	Sap D1	This study
A94	pASK-IBA3plus	anhydrotetracycline	Ampi	Sap D1-2	This study
A95	pASK-IBA3plus	anhydrotetracycline	Ampi	Sap D3	This study
A153	pASK-IBA3plus	anhydrotetracycline	Ampi	Sap D4-5	This study
A96	pASK-IBA3plus	anhydrotetracycline	Ampi	Sap D6	This study

Supplementary table 5. Plasmids table used in this study

Movie S1. Sap conformational change. The video illustrates the SapAA^D monomer (domains D1-D6) going from the pre-assembly to the post-assembly state. In the transition, a 20° pivot around the D4-D5 linker and a 180° rotational isomerization of D6 and D3D2D1 occur, pivoting around the D5-D6 and D4-D3 linkers. An intramolecular D3-D6 contact in the pre-assembly state is replaced by two isomorphous intermolecular D3-D6 contacts in the post-assembly S-layer, essential for S-layer formation and stability.