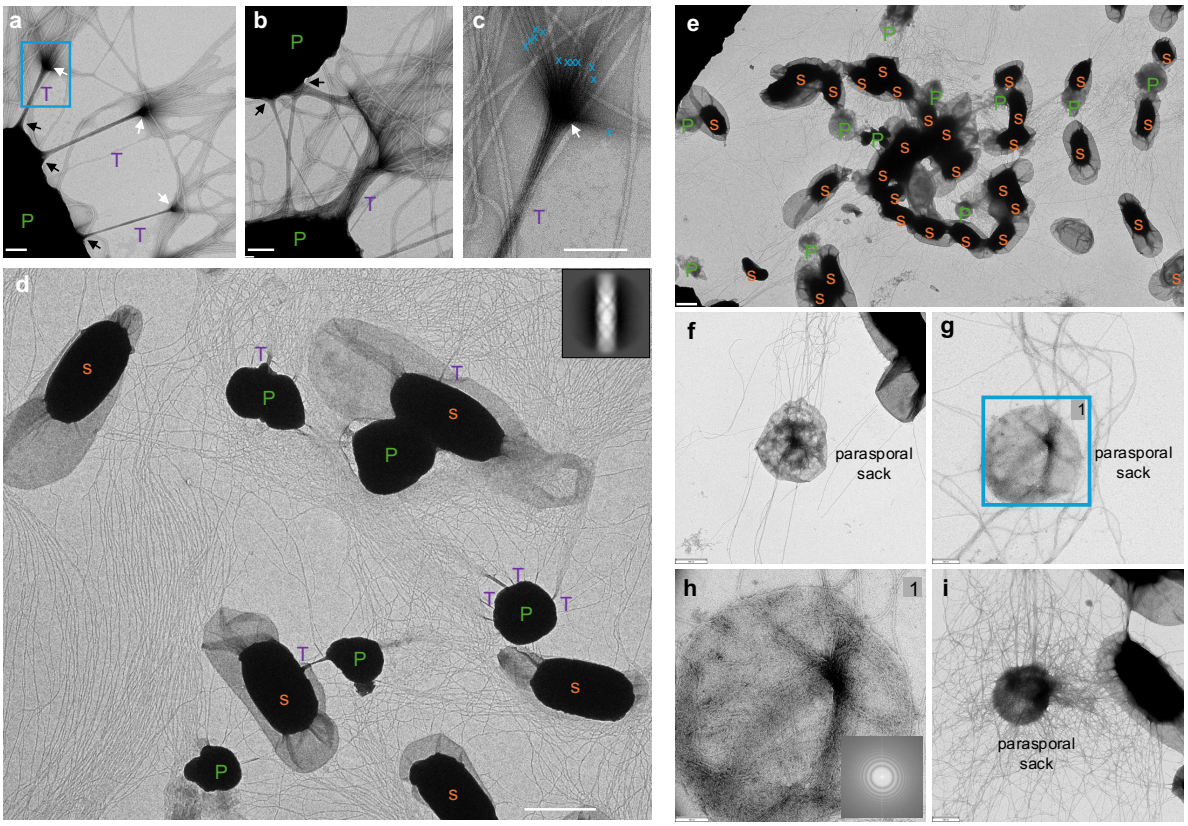
Supplemental information

Auto-crosslinking sporesilk fibers promote endospore and Cry toxin clustering

Authors

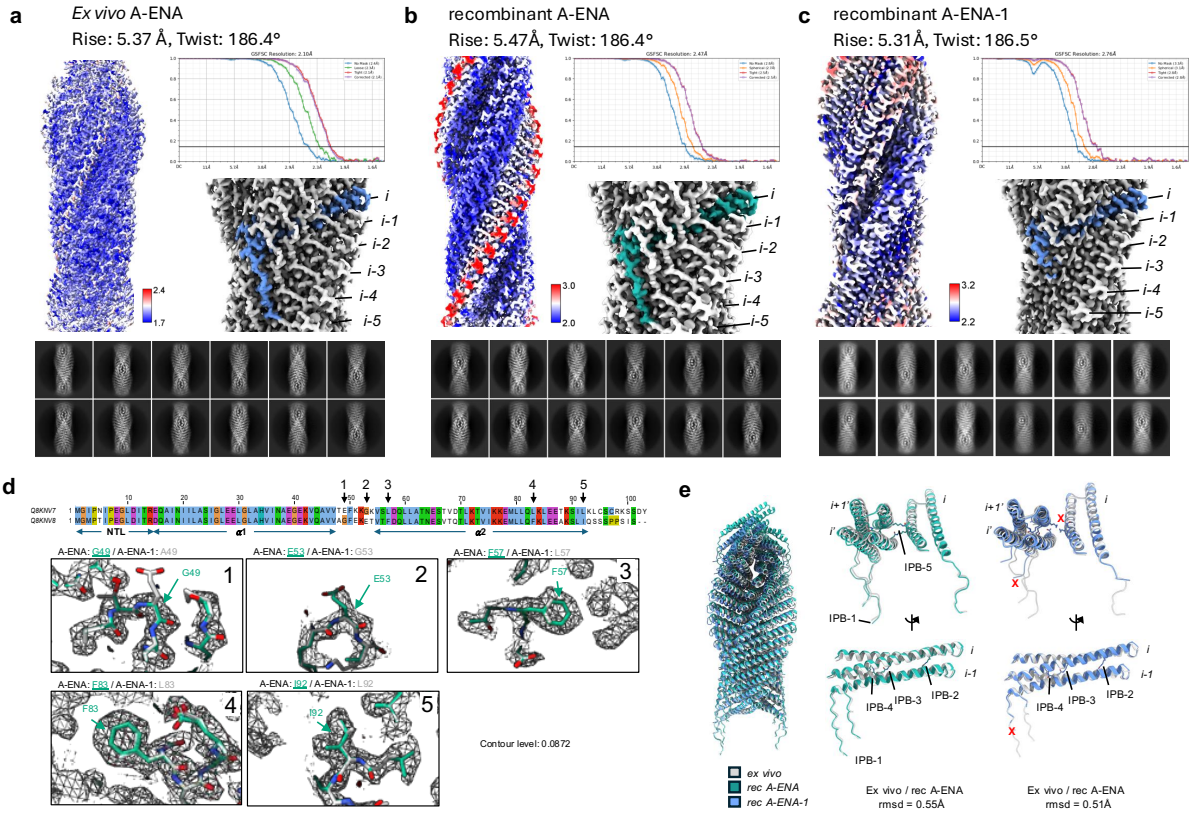
Mike Sleutel[*1,2,3], Adrià Sogues[1,2,3], Han Remaut[**1,2,4]

Content:



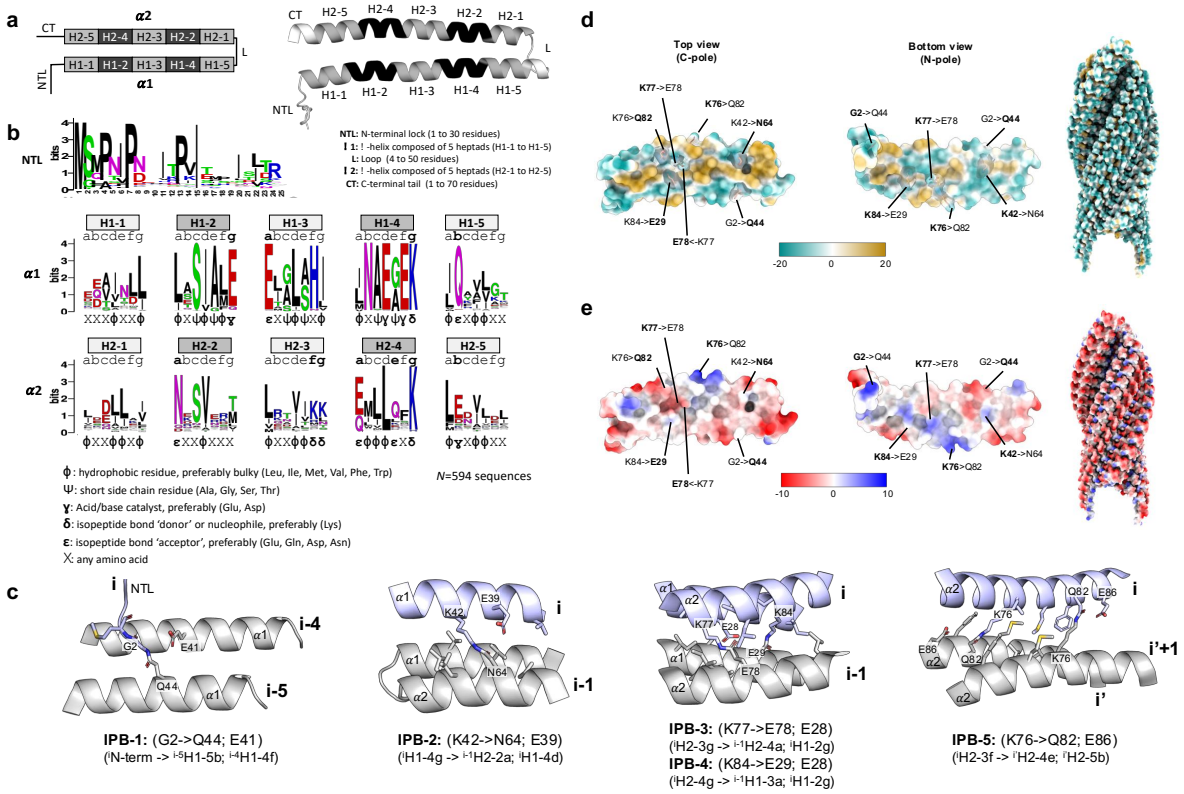
Extended Data Figure 1. **Extracellular matrix of Bti functioning as a molecular net entrapping spores and PSBs.** nsTEM images of a resuspended spore biofilm deposited on a non-glow discharged **(a-c)** Cu-mesh formvar grid: spores (S), residual mother cells (M) and parasporal bodies (P) form micron sized clusters that are held together by a proteinaceous extracellular matrix which collapses in certain regions into bundled tethers (T). Magnified view of boxed area in panel a, showing the bundled A-ENA tethers coupled to the surface of an IPB feather out into 12 individual nanofibers. **(d)** nsTEM images of a resuspended spore biofilm deposited on a glow discharged Cu-mesh formvar grid. Contrary to (a-c) where the ECM is in a more collapsed state, here the fiber network spreads out into large, parallel arrays of A-ENA fibers. We attribute that difference to the change in affinity between A-ENA and the formvar substrate after glow discharging, which is expected to render the surface more hydrophilic. Given the polar, hydrophilic character of the surface exposed residues of A-ENA, an increased affinity after glow discharging is expected. As such, the expanded state of fibers can be considered as more representative of the hydrated state of the ECM when embedded in a liquid. Scale bars: 100 nm (a,b,c), 1μm (d). **(f-i)** nsTEM image of a spore sample that was incubated in 0.5M CAPS pH 11.0 for 18h, supplemented afterwards with 8M urea and heated to 99°C for 5min, centrifuged at 22.000 rcf for 15 min at RT and resuspended in miliQ. Most PSBs have been dissolved as a result of this treatment. (f, g) zoomed-in view of the parasporal sack. (h) Magnified view of a parasporal sack with its corresponding power spectrum (PS) as

inset. The PS shows no distinct maxima suggesting that the sack structure is not crystalline in nature following this treatment. (i) what remains is a parasporal sack-like structure that are decorated with A-ENA fibers making connections to neighboring spores. Scale bars: 1μm (e), 500 nm (f,i), 200 nm (g), 100 nm (h).



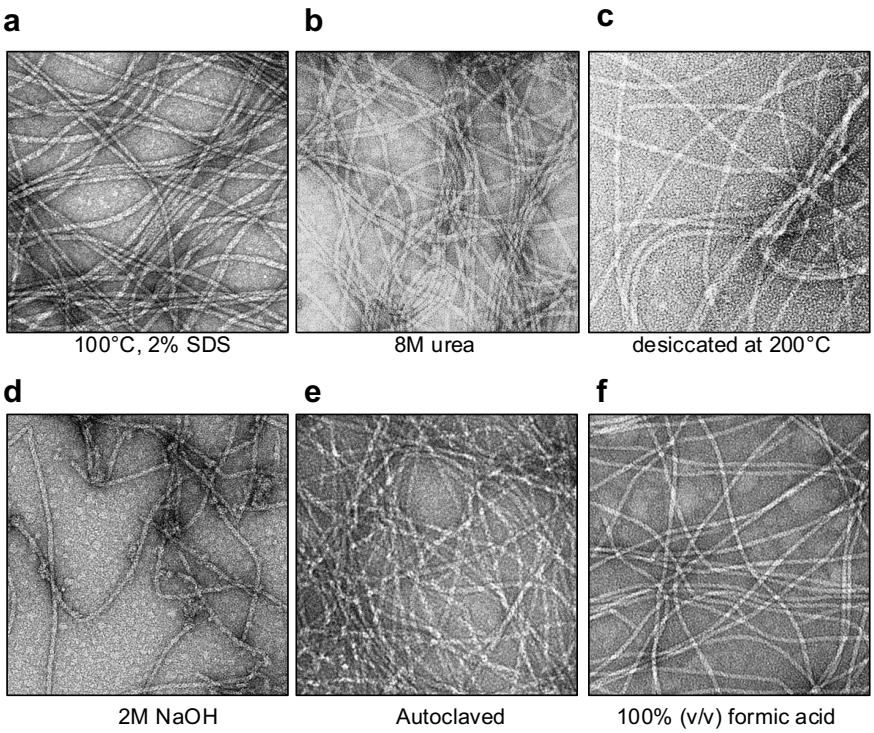
Extended Data Figure 2. CryoEM structures of *ex vivo* A-ENA fibers and recombinant A-ENA and A-ENA-1 fibers. (a-c) 3D cryoEM maps of *ex vivo* A-ENA fibers (a), recombinant A-ENA fibers (b) and recombinant A-ENA1 fibers (c) coloured according to local resolution; upper right panels show Fourier shell correlation curve; mid right panels show zoom-in of the cryoEM volume with one protomer highlighted in blue; lower panels show 2D class averages of boxed segments of the respective fibers (box-size is 210x210Å). **(d)** Pairwise sequence alignment between Q8KNV7 (A-ENA-1) and Q8KNV8 (A-ENA): residue positions with notable side-chain identity type differences are demarked with an arrow. Numbers correspond to the figure panels below, showing docked models (stick) for A-ENA (cyan) and A-ENA-1 (grey) with 3D cryoEM map (wireframe) of *ex vivo* A-ENA fibers contoured at 0.0872 used for the identification of the most likely major subunit of the extracellular matrix fibers. At all five sites the cryoEM maps are compatible with the local A-ENA sequence and show no clear signs of density compatible with the A-ENA1 sequence, thus suggesting ex vivo A-ENA fibers are predominantly composed of A-ENA, and comprise little to no A-ENA1 (see also Figure S6). **(e)** left: structural overlay of the atomic models of *ex vivo A-ENA* (grey), recombinant A-ENA

(cyan) and recombinant A-ENA-1(light blue) nanofibers (left) and individual subunits (middle and right).

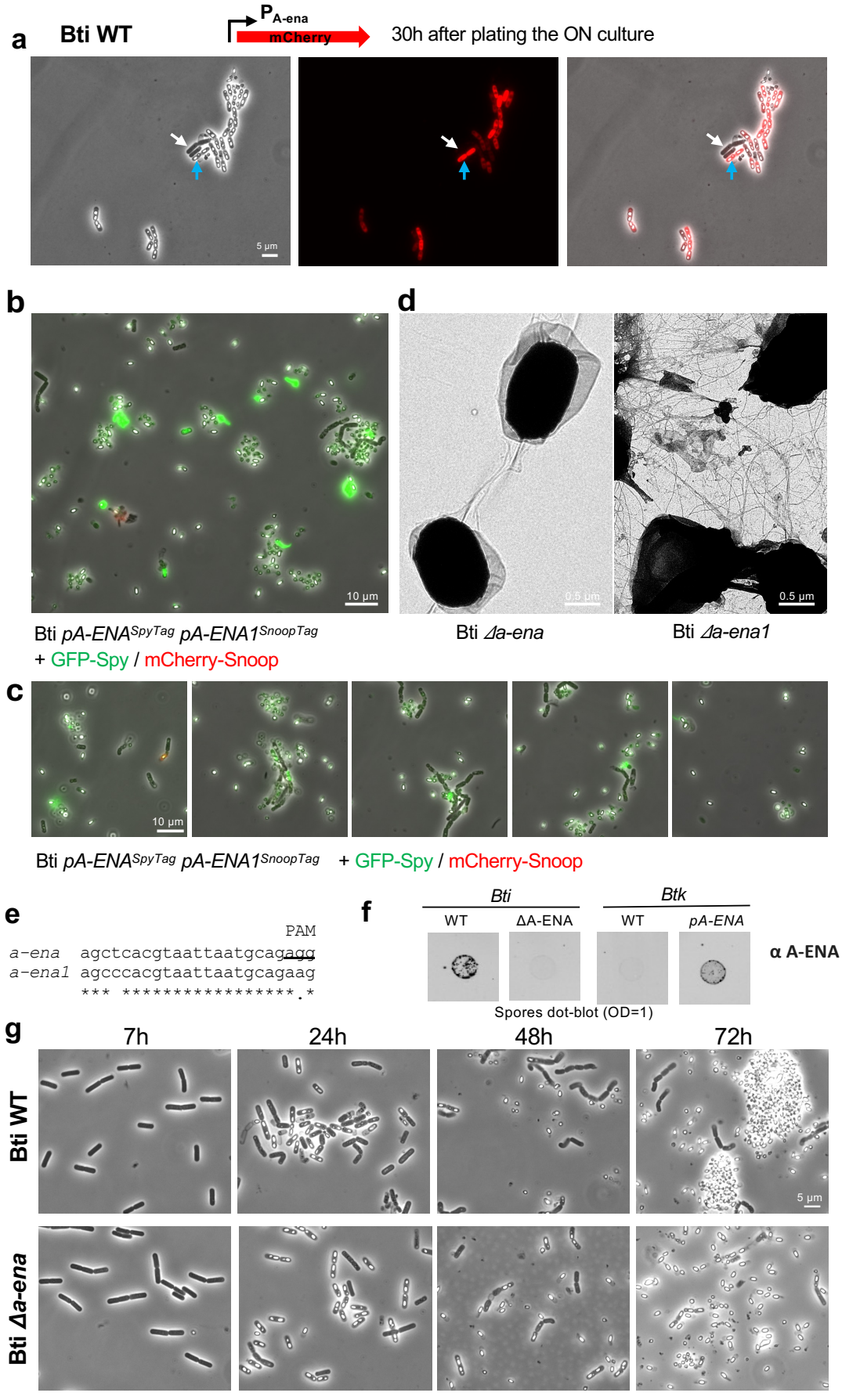


Extended Data Figure 3. Conservation and assembly principles of cross-helix isopeptide bonds in A-ENA-like proteins. (a) Schematic and ribbon representation of the generic structure of A-ENA-like monomers, comprising an N-terminal lock (NTL) of variable length, followed by two anti-parallel α-helices separated by a short loop (L) and each formed of five consecutive 7-residue (heptad, positions a-g)) α-helical modules (H1-1, H1-2, H1-3, H1-4 and H1-5 in helix 1; H2-1, H2-2, H2-3, H2-4 and H2-5 in helix 2), followed by a variable length C-terminal tail (CT). **(b)** Consensus weblogo derived from a multiple sequence alignment (MSA) of A-ENA-like sequences comprising the top 250 PSI-BLAST hits using Q8KNV8 as a query sequence, and filtered down to 95 % sequence identity redundancy. **(c)** Ribbon representation of the five isopeptide bond (IPB1-5) modules present in Q8KNV8. The IPB modules each comprise a nucleophile (N-terminal amine or Lys), an electrophile (Glutamate, Glutamine, Aspartate or Asparagine), and an acid-base catalyst (Glutamate or Aspartate) located in close proximity across intermolecular NTL-helix (IPB1) or helix-helix interfaces within (IPB2-4) or across (IPB5) the A-ENA protofilaments f and f'. Subunit labelling as in Figure 3, and residue numbering according to Q8KNV8. (b.c) In the sequence logo, residues involved in isopeptide bond (IPB) formation have been highlighted with ε (electrophile or IPB 'acceptor'), δ (nucleophile or IPB 'donor'), and γ (acid-base catalyst). The MSA shows

the IPB partners to be amongst the strongest conserved residues in A-ENA-like sequences, and to be located in fixed positions in the helical heptads (labelled a-g) in accordance with their positioning within in the helix-helix interfaces. **(d,e)** Top (C-pole) and bottom (N-pole) view of a single A-ENA monomer in surface representation, coloured according to hydrophobicity (d; palette lipophilicity in ChimeraX v1.7) or to electrostatic potential (e; palette red-white-blue in ChimeraX v1.7). Both surfaces represent the contact interface in the longitudinal stacking of subunits, thus resulting in polar protofibrils. These contact surfaces are primarily hydrophobic, with the exception of residues involved in IPB formation, highlighted in stick representation. As such, hydrophobic knobs-in-hole stacking of the subunits results in the close positioning of the IPB donor – acceptors and their acid-base catalysts, activating the nucleophilic residues for attack and nucleophilic substitution (i.e. dehydration) with the electrophilic residues. The defined positioning of nucleophilic and electrophilic residues on adjacent helices results in spontaneous, autocatalytic cross-linking of A-ENA subunits under physiological conditions.

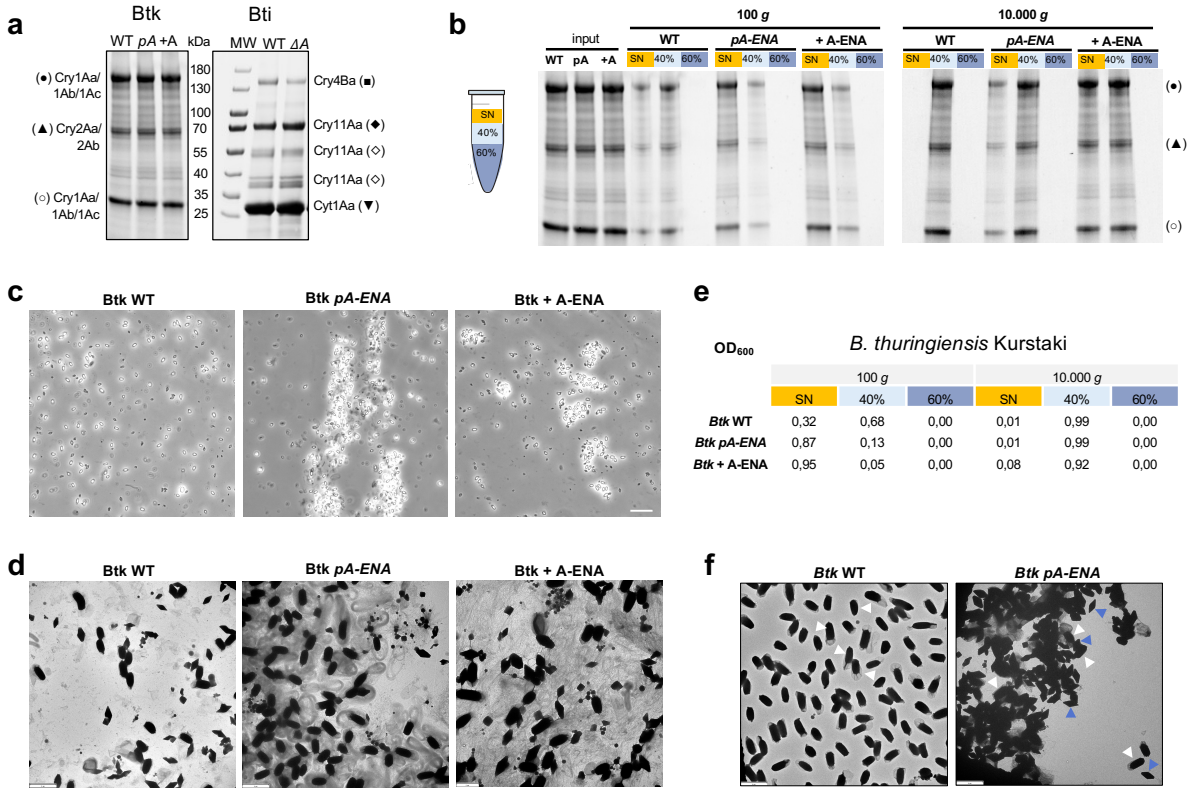


Extended Data Figure 4. Stability analysis of recA-ENA fibers. (a) Incubated for 1h at 99°C in 2% (w/v) SDS, **(b)** Incubated for 1h at room temperature in 8M urea, **(c)** A-ENA fibers suspended in miliQ were desiccated at 200°C for 15min, and rehydrated in miliQ for imaging, **(d)** Incubated for 1h at room temperature in 2M NaOH, **(e)** Autoclaved for 20min at 121 °C, **(f)** Incubated for 30min at room temperature in 100% (v/v) formic acid.



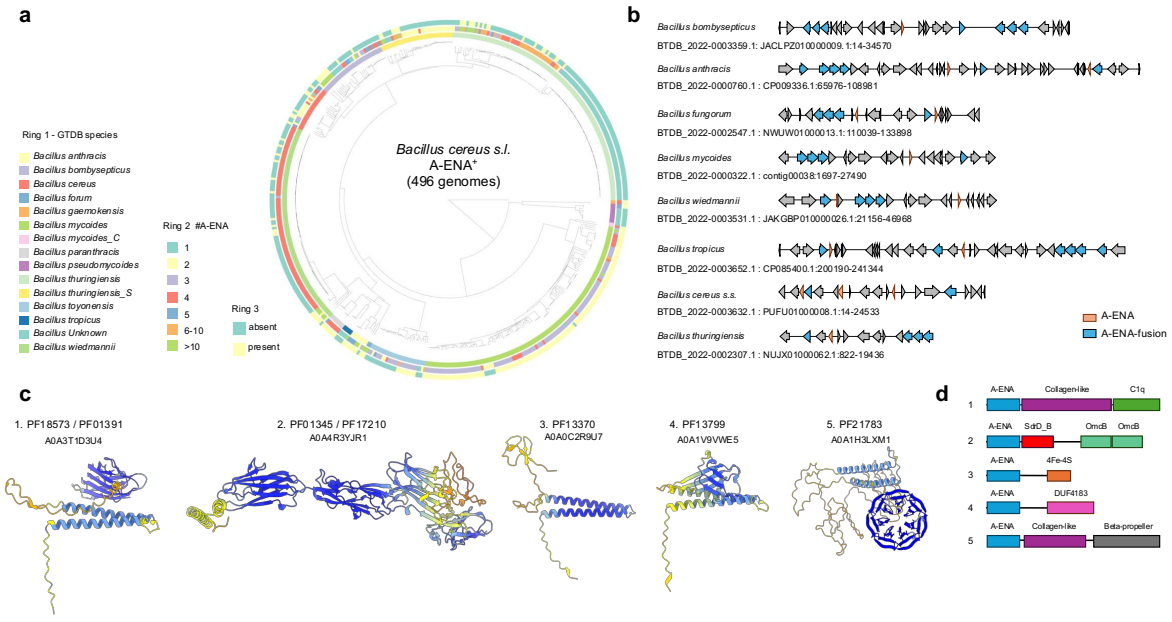
Extended Data Figure 5. Phenotypic characterization of *A-ENA* and *A-ENA-1* knockout

mutants. (**a**) phase-contrast (left) and fluorescence (middle; overlay right) microscopy of sporulating Bti WT cells expressing mCherry under the control of the *a-ena* promoter. Images were taken 30 h after plating. White arrows indicate vegetative cells (absence of fluorescence), while blue arrows highlight sporulating cells with mCherry expression in the mother cell compartment. (**b-c**) Bti WT strain carrying a plasmid expressing *a-ena*^{SpyTag} and *a-ena1*^{SnoopTag} under their native promoters. Mature spores were incubated with SpyCatcher-GFP and SnoopCatcher-mCherry, allowing the visualization of A-ENA (green) and A-ENA1 (red). (**d**) (*Left*) Negative-stain TEM (nsTEM) image of the *Δa-ena* mutant, showing the absence of extracellular fibers. (*Right*) nsTEM image of the *Δa-ena1* mutant, showing the presence of fibers. (**e**) Sequence alignment of the gRNA used to knock out *a-ena* with the homology sequence in the *a-ena1* gene. Despite their sequence similarity, *a-ena1* lacks a protospacer adjacent motif (PAM, NGG; underlined), preventing cleavage and ensuring that only a-ena is disrupted. (**f**) Dot blot analysis of spore suspensions (OD_{600} = 1) using an anti-A-ENA antibody. Strains are indicated above the blot. (**g**) Time series microscopy of sporulating Bti WT and *Δa-ena* mutants at 7, 24, 48, and 72 hours. Both strains show similar progression through sporulation stages, but clustering is observed in WT at 72 h.

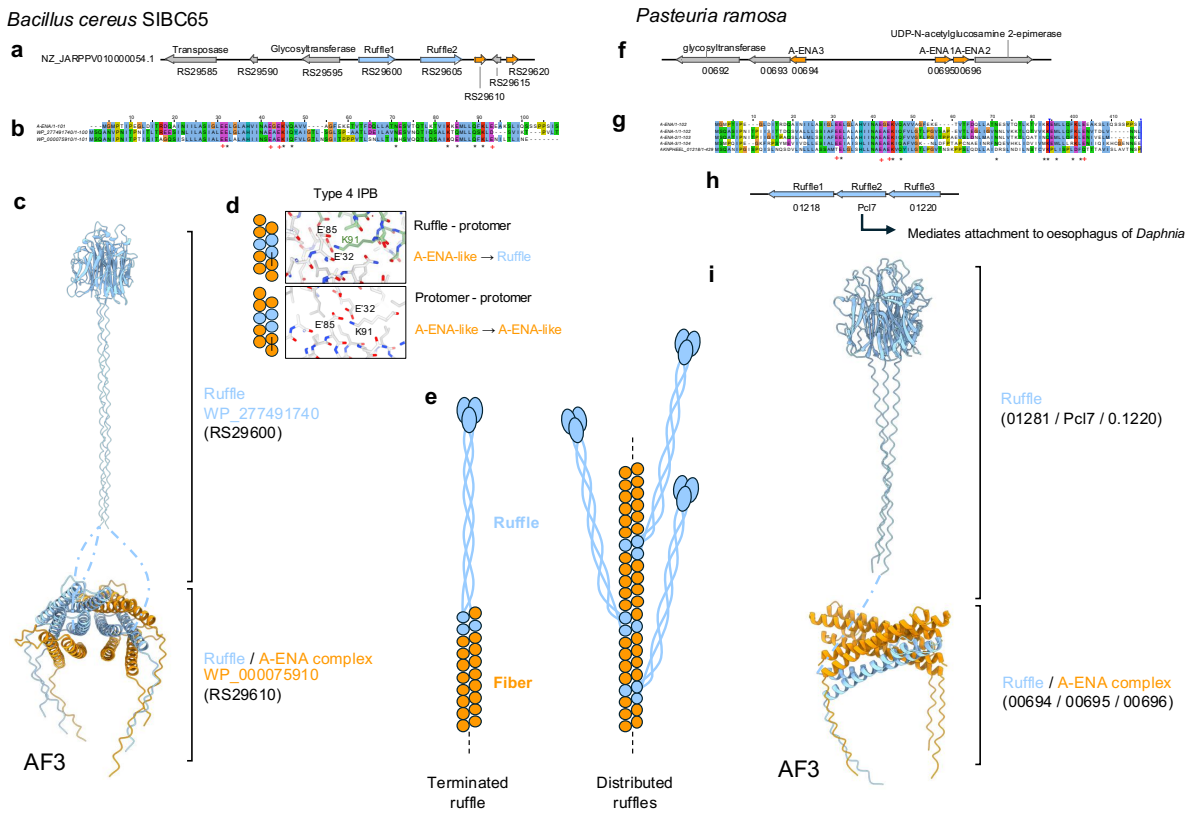


Extended Data Figure 6. Recombinant expression or exogenous addition of A-ENA results in endospore – PSB clustering in *Bacillus thuringiensis* Kurstaki. (**a**) SDS-PAGE analysis of dissolved crystal proteins from *Btk* (left) and *Bti* (right). For *Btk*, samples include

WT, Btk *pAS_A-ENA* (pA, expressing *a-ena* from a plasmid under its native promoter), and recombinant A-ENA (4.18 mg/ml final concentration) mixed with *Btk* WT ('+A'), all normalized by OD_{600}. For *Bti*, WT and *Δa-ena* (*ΔA*) were similarly normalized. Bands were analyzed by mass spectrometry (MS) fingerprinting, with identified proteins indicated next to each band. (**b**) SDS-PAGE analysis of different fractions after Histodenz cushion separation. The input (before fractionation) confirms equal toxin loading across conditions. The fraction type, strain, and centrifugation speed are indicated above the gel. (**c**) Light microscopy images of the indicated strains reveal clustering induced by the presence of A-ENA. scale bar = 10 μm. (**d**) nsTEM images of the analyzed strains show that A-ENA, whether expressed from a plasmid (*pA-ENA*) or recombinantly added to WT cells, induces phenotypes similar to those observed in *Bti*. Scale bar = 3 μm. (**e**) Normalized OD_{600} values of different fractions and strains after fractionation. (**f**) nsTEM of purified *Btk* spores obtained using a sucrose gradient. In the 80% sucrose fraction, pure spores from WT remain unclustered, whereas *pA-ENA* spores are associated with PSCs, forming clusters. White arrowheads indicate spores, while blue arrowheads indicate PSCs. Scale bar = 3 μm.



Extended Data Figure 7. (a) Phylogenetic tree of A-ENA carrying genomes in the *B.c.s.l.* group: inner ring: GTDB species, second ring: number of *a-ena* genes, outer ring: absence/presence (green/yellow) of an additional Pfam domain within genes carrying an A-ENA domain; **(b)** Examples of A-ENA gene cluster organization; **(c)** Pfam analysis of A-ENA orthologues with globular domain(s) in C-terminal extension to the A-ENA domain.



Extended Data Figure 8. (a) Genetic organization of ruffle and A-ENA genes in the genome of *Bacillus cereus* SIBC65; **(b)** multiple sequence alignment of A-ENA from Bti and ORFs RS29610 and RS29620 from *B.cereus* SIBC65: catalytic residues marked with +, IPB residues marked with *; **(c)** AF3 multimer prediction of a putative *B.cereus* ruffle/A-ENA-like complex: N-terminal A-ENA-like domains of the ruffle protein interlock with the A-ENA protomers; **(d)** Zoom-in of the type 4 IPB site in the protomer-protomer and protomer-ruffle interfaces; **(e)** Schematic representation of two putative modes of A-ENA ruffle decoration: *Terminated ruffle*: a single trimeric ruffle caps the A-ENA fiber terminus, *Distributed ruffles*: multiple trimeric ruffles are integrated in random positions into the A-ENA fiber. **(f)** Multiple sequence alignment of A-ENA from Bti and ORFs 00694, 00695, 00696 from *Pasteuria ramosa*: catalytic residues marked with +, IPB residues marked with *; **(g, h)** Genetic organization of ruffle and A-ENA genes in the genome of *Pasteuria ramosa*; **(i)** AF3 multimer prediction of a putative *P. ramose* ruffle/A-ENA-like complex: N-terminal A-ENA-like domains of the ruffle protein interlock with the A-ENA protomers.

Extended Data Table 1: Plasmids and strains

Plasmid name	Plasmid #	Description	Reference
Plasmid for Bacillus thuringiensis knock out generation			
pJOE8999		CRISPR-Cas9 vector; KanR	BGSC
pJOE-Δa-ena	A266	pJOE899 derivative for A-ENA STOP insertions	this work
pJOE-Δa-ena-1	A190	pJOE899 derivative for A-ENA-1 deletion	this work
Plasmid for recombinant expression in E. coli			
pASK-IBA3plus		AmpR; Tet inducible	AddGene
pASK-A-ENA		pASK derivate. AmpR; Bti A-ENA Tet inducible	This work
pASK-A-ENA-1		pASK derivate. AmpR; Bti A-ENA-1 Tet inducible	This work
pASK-sfGFP:SpyCatcher	A142	pASK derivate. AmpR; sfGFP:SpyCatcher Tet inducible	This work
pASK-mCherry:SnopCatcher	A185	pASK derivate. AmpR; mCherry:SnopCatcher Tet inducible	This work
Plasmid for recombinant expression in Bacillus thuringiensis			
pAS		pJOE-derivative. KanaR; ΔCas9	This work
pAS-Pa-ena-A-ENA	A176	pAS derivative. KanaR. Bti A-ENA under A-ENA promoter	this work
pAS-Pa-ena1-A-ENA1	A186	pAS derivative. KanaR. Bti A-ENA-1under A-ENA-1 promoter	this work
pAS-Pa-ena1-A-ENA-1:snoopTAG	A187	pAS derivative. KanaR. Bti A-ENA-1:SnoopTag under A-ENA-1 promoter	this work
pAS-Pa-ena-A-ENA:SpyTAG	A180	pAS derivative. KanaR. Bti A-ENA:SpyTag under A-ENA promoter	this work
pAS-Pa-ena-A-ENA:SpyTAG+Pa-ena1-A-ENA-1:snoopTAG	A188	pAS derivative. KanaR. Bti A-ENA:SpyTag under A-ENA promoter+ENA-1:SnoopTag under A-ENA-1 promoter	this work
pAS-Pa-ena-mcherry	A197	pAS derivative. KanaR. mCherry under A-ENA promoter	this work
Bacterial strains			
E.coli BL21 (DE3)		fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5 ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10) TetS endA1	New England Biolabs
E. coli Dam-/Dcm-		rspL136 (StrR)dam13::Tn9 (CamR) xylA-5 mtl-1 thi-1 mcrB1 hsdR2	New England Biolabs
Bacillus thuringiensis serovar Israelensis AM65-52	Bti WT		(Bolotin et al., 2017)
Bacillus thuringiensis serovar Kurstaki HD 1	Btk WT		(Day et al., 2014)

Extended Data Table 2: CryoEM data and model statistics.

Parameter	*Ex vivo* A-ENA	*Rec* A-ENA	*Rec* A-ENA-1
Data collection and processing			
Voltage (kV)	300	300	300
Electron exposure (e^{-} Å^{-2})	60	60	60
Pixel size (Å)	0.766	0.766	0.766
Particle images (n)	532,909	940,701	121,101
Shift (pixel)	20	20	20
Helical symmetry			
Point group	C1	C1	C1
Helical rise (Å)	5.37	5.47	5.31
Helical twist (°)	186.4	186.4	186.5
Map resolution (Å)			
Map:map FSC (0.143)	2.10	2.38	2.64
Model:map FSC (0.38)	2.49	2.48	3.00
d_{99}	2.40	3.10	3.50
Refinement and Model validation			
Ramachandran Favored (%)	100.0	97.83	98.82
Ramachandran Outliers (%)	0.0	0.0	0.0
$CC_{main\ chain}$	0.84	0.88	0.83
Clash score	4.16	6.46	7.65
Bonds RMSD, length (Å)	0.003	0.005	0.004
Bonds RMSD, angles (°)	0.506	0.926	0.656
Deposition ID			
PDB (model)	9IHH	9IHJ	9Q82
EMDB (map)	EMD-52871	EMD-52872	EMD-52882

Extended Data References:

Bolotin, A., Gillis, A., Sanchis, V., Nielsen-LeRoux, C., Mahillon, J., Lereclus, D., and Sorokin, A. (2017). Comparative genomics of extrachromosomal elements in Bacillus thuringiensis subsp. israelensis. Res Microbiol *168*, 331-344.

Day, M., Ibrahim, M., Dyer, D., and Bulla, L., Jr. (2014). Genome Sequence of Bacillus thuringiensis subsp. kurstaki Strain HD-1. Genome Announc *2*.