



# Architecture and Self-Assembly of *Clostridium sporogenes* and *Clostridium botulinum* Spore Surfaces Illustrate a General Protective Strategy across Spore Formers

Thamarai K. Janganan,<sup>a,b\*</sup> Nic Mullin,<sup>a,c</sup> Ainhoa Dafis-Sagarmendi,<sup>a,b</sup> Jason Brunt,<sup>d\*</sup> Svetomir B. Tzokov,<sup>a,b</sup> Sandra Stringer,<sup>d</sup> Anne Moir,<sup>a,b</sup> Roy R. Chaudhuri,<sup>a,b</sup> Robert P. Fagan,<sup>a,b</sup> Jamie K. Hobbs,<sup>a,c</sup>  Per A. Bullough<sup>a,b</sup>

<sup>a</sup>Krebs Institute, University of Sheffield, Sheffield, United Kingdom

<sup>b</sup>Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, United Kingdom

<sup>c</sup>Department of Physics and Astronomy, University of Sheffield, Sheffield, United Kingdom

<sup>d</sup>Quadram Institute, Norwich, United Kingdom

**ABSTRACT** Spores, the infectious agents of many *Firmicutes*, are remarkably resilient cell forms. Even distant relatives can have similar spore architectures although some display unique features; they all incorporate protective proteinaceous envelopes. We previously found that *Bacillus* spores can achieve these protective properties through extensive disulfide cross-linking of self-assembled arrays of cysteine-rich proteins. We predicted that this could be a mechanism employed by spore formers in general, even those from other genera. Here, we tested this by revealing in nanometer detail how the outer envelope (exosporium) in *Clostridium sporogenes* (surrogate for *C. botulinum* group I), and in other clostridial relatives, forms a hexagonally symmetric semipermeable array. A cysteine-rich protein, CsxA, when expressed in *Escherichia coli*, self-assembles into a highly thermally stable structure identical to that of the native exosporium. Like the exosporium, CsxA arrays require harsh “reducing” conditions for disassembly. We conclude that *in vivo*, CsxA self-organizes into a highly resilient, disulfide cross-linked array decorated with additional protein appendages enveloping the forespore. This pattern is remarkably similar to that in *Bacillus* spores, despite a lack of protein homology. In both cases, intracellular disulfide formation is favored by the high lattice symmetry. We have identified cysteine-rich proteins in many distantly related spore formers and propose that they may adopt a similar strategy for intracellular assembly of robust protective structures.

**IMPORTANCE** Bacteria such as those causing botulism and anthrax survive harsh conditions and spread disease as spores. Distantly related species have similar spore architectures with protective proteinaceous layers aiding adhesion and targeting. The structures that confer these common properties are largely unstudied, and the proteins involved can be very dissimilar in sequence. We identify CsxA as a cysteine-rich protein that self-assembles in a two-dimensional lattice enveloping the spores of several *Clostridium* species. We show that apparently unrelated cysteine-rich proteins from very different species can self-assemble to form remarkably similar and robust structures. We propose that diverse cysteine-rich proteins identified in the genomes of a broad range of spore formers may adopt a similar strategy for assembly.

**KEYWORDS** *Bacillus anthracis*, *Bacillus cereus*, *Bacillus subtilis*, *Clostridium difficile*, anaerobes, atomic force microscopy, botulism, disulfide bonding, electron microscopy, nanomaterials, protein structure-function, sporulation


**Citation** Janganan TK, Mullin N, Dafis-Sagarmendi A, Brunt J, Tzokov SB, Stringer S, Moir A, Chaudhuri RR, Fagan RP, Hobbs JK, Bullough PA. 2020. Architecture and self-assembly of *Clostridium sporogenes* and *Clostridium botulinum* spore surfaces illustrate a general protective strategy across spore formers. *mSphere* 5:e00424-20. <https://doi.org/10.1128/mSphere.00424-20>.

**Editor** Craig D. Ellermeier, University of Iowa

**Copyright** © 2020 Janganan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Per A. Bullough, [p.bullough@sheffield.ac.uk](mailto:p.bullough@sheffield.ac.uk).

\* Present address: Thamarai K. Janganan, School of Life Sciences, University of Bedfordshire, Luton, United Kingdom; Jason Brunt, Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom.

 EM and AFM have revealed for the first time the detailed structure of a *Clostridium* spore surface. We show that it assembles through a mechanism of self-organization now emerging as a general mechanism in very distantly related spore formers. @PerBullough

**Received** 10 June 2020

**Accepted** 12 June 2020

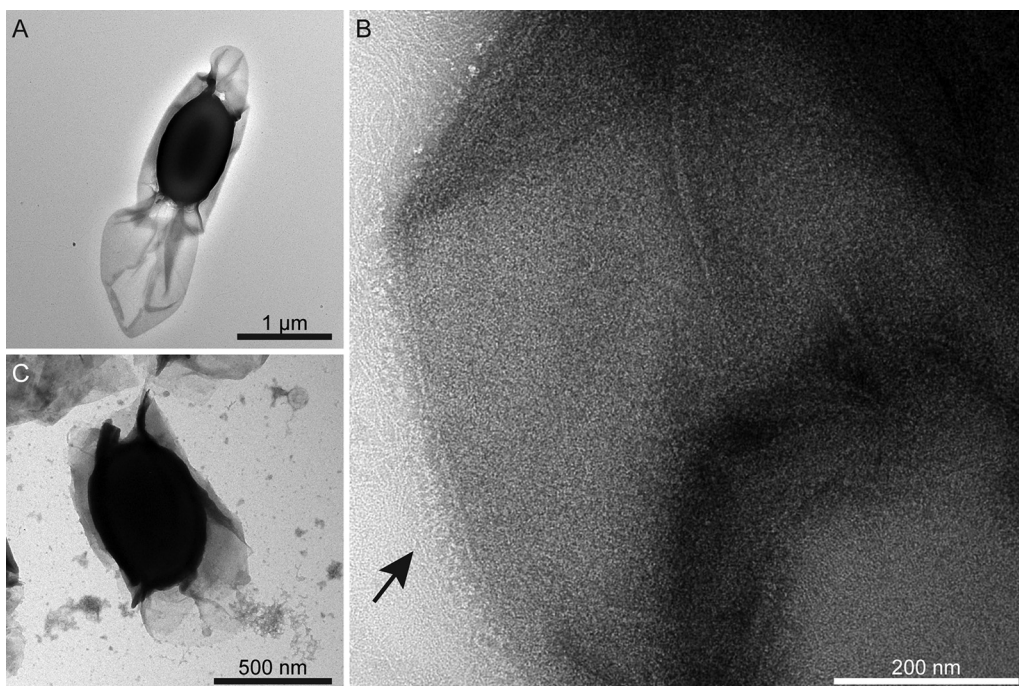
**Published** 1 July 2020

Spores formed by bacteria of the genera *Clostridium* and *Bacillus* provide a uniquely effective means of surviving environmental stress (1); they act as the infectious agent in pathogens such as *Bacillus anthracis*, *Clostridium botulinum*, and *Clostridium difficile*. In the anaerobic clostridia, they are essential for survival in air. Despite the early evolutionary divergence of the genera *Clostridium* and *Bacillus*, a number of the genes responsible for regulation and morphogenesis in sporulation are conserved. However, proteins making up the spore outer layers are much less extensively conserved (2). These layers include a complex protein coat and, in some species, such as the pathogens *B. anthracis* and *C. botulinum* (but not *B. subtilis*), a distinct and deformable outermost exosporium enveloping the spore. The outer protein layers confer much of the spore's resistance to chemical and enzymatic insult (1). The genetic control and the role of key morphogenetic proteins in spore outer layer assembly are well studied in *B. subtilis* (3) but far less so in other species, particularly the *Clostridia*.

The exosporium defines the interface between the spore and its environment. Where the spore acts as an infectious agent, it is the first point of contact between the spore and the host. In *B. anthracis*, it has roles in modulating spore germination, adhesion, protection (reviewed in reference 4), host cell uptake (5), and immune inhibition (6). The physical and structural properties of the exosporium have been best studied in the *B. cereus*-*B. anthracis* group, where it comprises a thin, continuous, and hexagonally crystalline proteinaceous layer (7) (known as the basal layer) whose lattice is formed by cysteine-rich proteins ExsY and CotY (8). Its external face is decorated by a "hairy nap" composed of BclA, which has an internal collagen-like repeat (CLR) domain (9) that is associated with the basal layer through the ExsFA/BxpB protein (10).

Much less is known of the corresponding properties in the *Clostridia*, with the exception of the medically important species *Clostridium difficile*, where several proteins important in spore coat and exosporium assembly have been identified (11–13). Now reclassified as *Clostridioides difficile*, this species is rather distant from the main group of *Clostridia*, however (14–16). *Clostridium botulinum* has an exosporium, but its composition and assembly are poorly understood. This species is significant as a potential bioterror agent; its toxin is responsible for botulism, a severe neuroparalytic disease that affects humans and other mammals and birds (17). Among the highly pathogenic proteolytic strains of group I *C. botulinum*, the closely related species *Clostridium sporogenes* is a useful nonpathogenic experimental surrogate (17, 18). This makes *C. sporogenes* an attractive target for probing clostridial spore structure and function. The *C. sporogenes* exosporium is morphologically similar to that of the *B. cereus* group and has been proposed to have a hexagonally symmetric crystalline basal layer (19) and a hairy nap (20), but the details of the molecular architecture of the exosporium have not been explored. Proteins extracted from purified *C. sporogenes* exosporium (20) include, among others, a clostridium-specific cysteine-rich protein, CsxA, that was detected in very-high-molecular-weight material, together with a BclA-like protein; the latter is a possible contributor to the hairy nap, by analogy with *B. cereus*.

We now reveal the three-dimensional (3D) molecular structure of a clostridial exosporium, using electron crystallography and atomic force microscopy (AFM). The novel cysteine-rich CsxA protein has been identified as the defining structural component of the basal layer array. This provides the first detailed view of the structure of the spore envelope of *C. botulinum* group I strains, and as CsxA homologues are encoded more widely, it provides insights for future spore coat and exosporium research in the genus *Clostridium*. We show that recombinant CsxA can self-assemble into crystalline arrays identical in structure to the exosporium. Thus, we show that apparently unrelated cysteine-rich proteins from different spore-forming species can self-assemble to form remarkably similar and robust structures. We propose that diverse cysteine-rich proteins identified in the genomes of a broad range of spore formers may adopt a similar strategy of molecular tiling to build up spore structures.



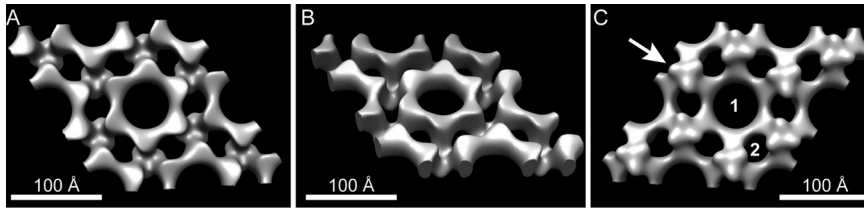
**FIG 1** A crystalline exosporium envelops the *C. sporogenes* spore. (A) Transmission electron micrograph of a spore negatively stained with uranyl formate. The thin exosporium layer surrounds the dense spore core and is often extended at one pole. (B) A high-magnification image from an area of exosporium displaying a “hairy nap” fringe (arrow). (C) *csxA* mutant spore showing the core wrapped in broken sheets of material, with sloughed-off fragments in the top left corner of the image.

## RESULTS AND DISCUSSION

**The exosporium of *C. sporogenes*, and of other *Clostridia*, is formed from a hexagonally symmetric two-dimensional lattice enveloping the spore.** Electron microscopy (EM) shows the exosporium enveloping the electron-dense spore core; it is generally more extensive at one pole (Fig. 1A). In all electron-transparent areas, the exosporium appears as a thin two-dimensional crystalline layer, mostly associated with a hairy nap and various other appendages (Fig. 1B) (20). Fourier amplitudes and phases were averaged from 5 high-magnification images of negatively stained exosporium. Unit cell parameters are  $a = b = 110 \pm 5 \text{ \AA}$  and  $\gamma = 120 \pm 3^\circ$ ; phases are consistent with  $p6$  symmetry. The projection map (see Fig. S1A in the supplemental material) reveals a densely stained core surrounded by a ring of 6 stain-excluding densities (black circle), separated by deeply stained pits (black rectangle). Each ring is connected to two adjacent rings by a trimeric linker (Fig. S1A; arrow). We also determined projection maps from exosporium of *C. acetobutylicum*, *C. tyrobutyricum*, *C. puniceum*, and *C. pasteurianum* (Fig. S2). These all display a density distribution nearly identical to that of *C. sporogenes* (Fig. S1A).

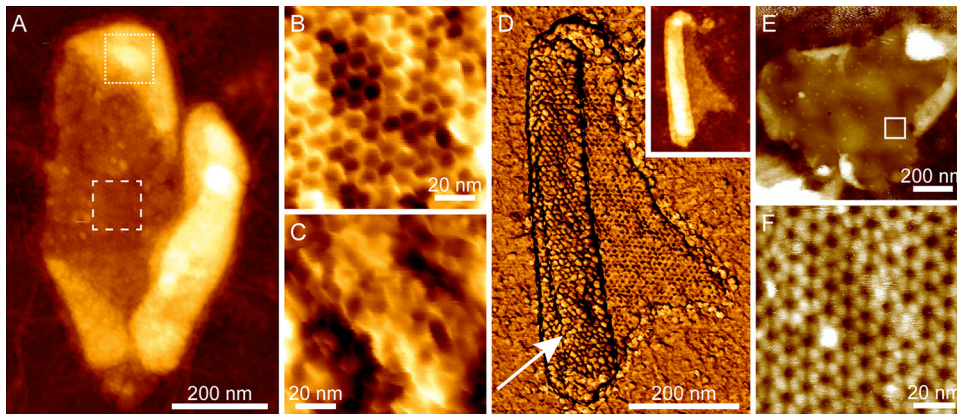
**Three-dimensional (3D) reconstruction of *Clostridium sporogenes* exosporium reveals a semipermeable protein lattice.** A total of 61 images of negatively stained exosporium from intact spores were recorded and processed in 10 separate tilt series of  $\pm 55^\circ$ . 3D merging statistics are given in Table S1 in the supplemental material. In 3D (Fig. 2), the basic repeating unit is revealed as a cog-like ring with 6-fold symmetry linked to adjacent rings through a small 3-fold symmetric bridge (arrow in Fig. 2C). The internal diameter of the ring is  $\sim 55 \text{ \AA}$ . The outer diameter of the ring at its widest point is about  $\sim 85 \text{ \AA}$  and the depth is  $\sim 30 \text{ \AA}$ , similar to the thickness of the basal layer determined by AFM (see below). Major and minor pores (labeled “1” and “2,” respectively, in Fig. 2C) fully permeate the layer. The density revealed in this reconstruction arises only from structures that have crystalline order—any disordered features, such as surface appendages, will have been suppressed by the image averaging.



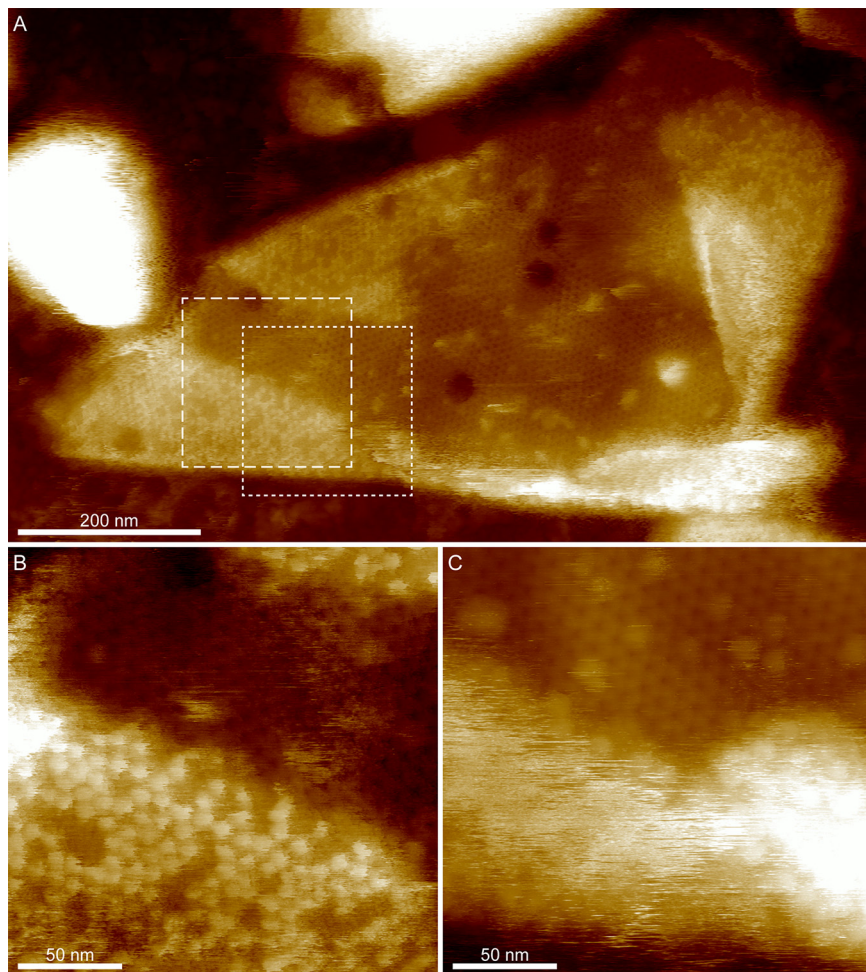


**FIG 2** Three-dimensional reconstruction of *C. sporogenes* exosporium revealing a single-layered hexagonal protein lattice permeated by pores. The surface is rendered to enclose a volume approximately matching that shown in Fig. 6A. (A and C) Views perpendicular to the plane of the exosporium layer. (B) View at 40° to the plane. The arrow indicates the 3-fold symmetric linker. Numerals 1 and 2 in panel C denote central and peripheral pores. The surface shown in panel A corresponds to the putative exterior surface. The resolution is  $\sim 25$  Å.

**Atomic force microscopy (AFM) reveals opposing inner and outer faces of the exosporium.** The external surface of the *C. sporogenes* exosporium is largely covered by a hairy nap (Fig. 1B) and decorated with several types of appendage (20–22). As a first step in assigning the exterior face (Fig. 2) that would bind the nap, we employed AFM to image the opposing exosporium surfaces. Exosporium fragments imaged in air displayed areas with thicknesses of  $\sim 70$  Å and  $\sim 140$  Å (Fig. 3A; see also Fig. S3A), which we interpret as the thicknesses of single layers and folded-over double layers of exosporium, respectively. We controlled the orientation of exosporium fragments by changing the binding conditions. Under conditions of binding at pH 4, the majority of fragments presented a honeycomb-like hexagonal lattice with unit cell axes of  $\sim 105$  Å (referred to here as the “honeycomb lattice”) (Fig. 3B), consistent with that measured by EM. We interpret these as single layers of exosporium, bound to the substrate nap side down, with the interior surface of the basal layer exposed to the imaging tip. Some fragments showed folded areas (2 layers thick) with a disordered surface (Fig. 3C), although occasionally a punctate hexagonal lattice with  $a = b = 98 \pm 5$  Å could be



**FIG 3** Atomic force microscopy reveals inner and outer exosporium surfaces. (A) Height image of a fragment of exosporium in air. The central area is a single layer thick (approximately 70 Å). The brighter (higher) peripheral areas are approximately twice the thickness of the central area, indicating that these are folded-over flaps of exosporium with the opposite surface exposed to the imaging tip. Color scale, 40 nm. (B) Higher-magnification view of the area denoted by the dashed box in panel A, showing a honeycomb-like hexagonal lattice with a unit cell measurement of  $105 \pm 3$  Å, proposed to represent the internal surface of the basal layer. Color scale, 3 nm. (C) Higher-magnification view of the area denoted by the dotted box in panel A. This view shows a disordered structure proposed to represent the hairy nap on the outer surface of the exosporium. Color scale, 5 nm. (D) (Inset) Height image of a different exosporium fragment in air, showing a similar folded structure. Color scale, 30 nm. (Main image) Simultaneously acquired phase image showing the same honeycomb lattice on the single-layered part of the fragment and a punctate hexagonal lattice with a similar unit cell size ( $98 \pm 5$  Å) on the double-layered area. Color scale, 10°. (E) Height image of an exosporium fragment imaged in water. The central area is one layer thick (approximately 30 nm in water, largely due to swelling of the hairy nap). Small folded regions are visible at the edges. Color scale, 80 nm. (F) Higher-magnification height image acquired in the area denoted by the box in panel E, showing a hexagonal lattice with a unit cell measurement of  $104 \pm 3$  Å, consistent with that observed in air and proposed to represent the internal surface of the basal layer. Color scale, 6 nm.

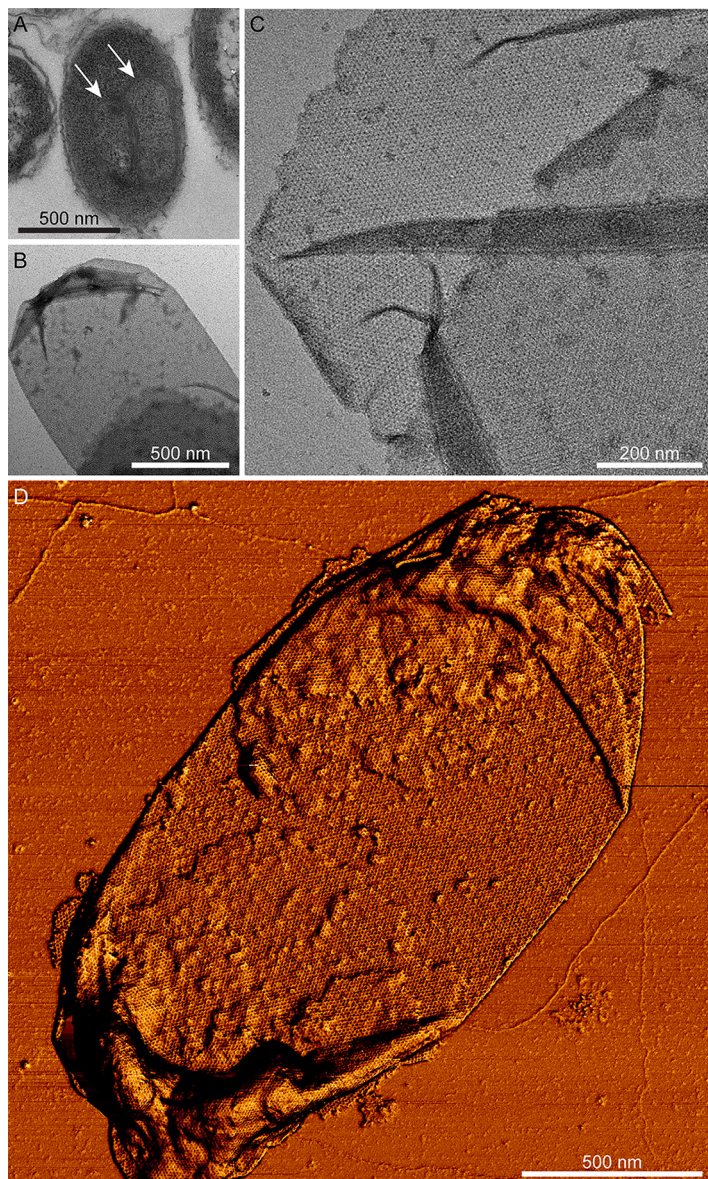


**FIG 4** AFM height images of hydrated exosporium reveal the underlying crystal lattice and a disordered surface. (A) Exosporium fragment imaged with a free amplitude of  $\sim 2$  nm and a relative set point of  $\sim 50\%$  (high force). The honeycomb lattice is visible as a single-layer area in the center of the fragment. (B) Magnified image of the area shown by the dashed box in panel A, also recorded under high-force conditions. The honeycomb lattice is visible in the upper half of the image, and a punctate lattice is visible on the folded-over area in the lower half of the image. The apparent height difference between the lower half and the upper half is 75 to 130 Å. (C) Magnified image of an area overlapping panel B and shown by the dotted box in panel A, imaged using low force ( $\sim 1$  nm free amplitude,  $\sim 90\%$  set point). The honeycomb lattice remains visible in the upper half, and the lower half, where the fragment has folded over the opposing surface, appears disordered. The height difference between lower half and the upper half is 200 to 450 Å.

seen (Fig. 3D; white arrow). These areas represent the external surface of the exosporium, with the basal layer crystal largely obscured by the hairy nap.

By the use of imaging in water, we again found a face with a honeycomb lattice consistent with that observed in air and by EM, though with pronounced protrusions at the 3-fold symmetric vertices of each hexagon (Fig. 3E and F). The apparent thickness of exosporium in water ranged from approximately 200 to 500 Å (Fig. S3B). The greater thickness of the hydrated exosporium is largely due to swelling of the hairy nap, which increases in thickness by a factor of  $5\times$  to  $10\times$  when measured in water, while the basal layer approximately doubles in thickness (Fig. S3). Disordered regions were again observed on folded fragments, but when the imaging force was increased, a punctate lattice with unit cell dimensions similar to those of the ordered areas in dry fragments became apparent (Fig. 4). Presumably at higher imaging forces, the AFM tip was able to penetrate to the more extensively ordered anchoring zone of the hairy nap filaments.





**FIG 5** Recombinant CsxA expressed in *E. coli* forms 2D crystals. (A) Thin-section electron micrograph showing formation of stacked layers within the cytoplasm (white arrows) of an *E. coli* BL21(DE3) pLysS cell expressing CsxA. (B) Electron micrograph of negatively stained sac-like 2D crystal of CsxA released from *E. coli* cell by sonication. (C) High-magnification image of a broken sac, exposing a single 2D crystalline layer. (D) AFM phase image of a purified intact sac showing the hexagonal arrays. The dark to bright variation in phase represents 7.06°.

**The 3D structure of self-assembled recombinant CsxA matches that of the native exosporium array.** *C. sporogenes* exosporium protein CsxA is found exclusively in the very-high-molecular-weight material extracted from the *C. sporogenes* exosporium, strongly suggesting that it has a structural role in the basal layer (20). The CsxA proteins of group I *C. botulinum* strains are highly similar, with amino acid identities ranging from 77% (strain A3 Loch Maree) to 100% (strain Prevot 594) (20), so data may be extrapolated more widely across the group. The *C. sporogenes* *csxA* gene was cloned and expressed in *Escherichia coli*, with a C-terminal His tag. *E. coli* cells expressing CsxA assembled stacks of sheet-like structures in the cytoplasm (Fig. 5A, white arrows). These were isolated from sonicated cells by nickel affinity purification. The sheets were mostly folded into closed sac-like two-dimensional crystals (Fig. 5B to D).

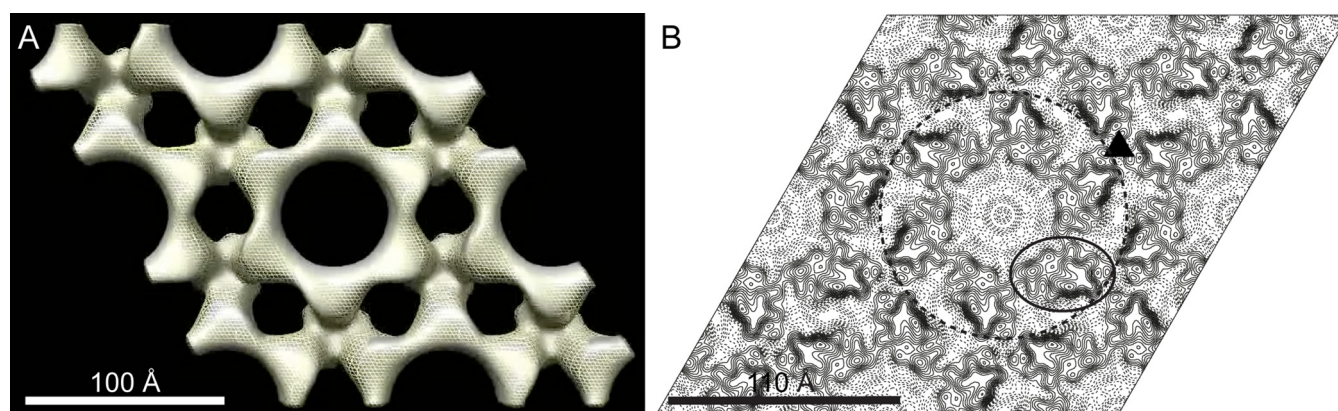
**TABLE 1** CsxA crystals adopt 6-fold symmetry<sup>a</sup>

Two-sided plane group	Phase residual vs other spots (90° random)	No. of comparisons	Target residual based on statistics taking Friedel wt into account
<i>p</i> 2	31.8*	57	37.5
<i>p</i> 3	13.9*	82	25.7
<i>p</i> 312	49.1	173	26.3
<i>p</i> 321	40.0	178	26.6
<i>p</i> 6	17.5*	221	28.7
<i>p</i> 622	46.2	408	27.3

<sup>a</sup>Data represent the internal phase residuals determined after the imposition of all possible two-sided plane groups, calculated from one of the micrographs of frozen hydrated CsxA crystals. Internal phase residuals were determined from spots of IQ1 to IQ5 to 7-Å resolution. The values marked with an asterisk (\*) are good candidates for the symmetry, as the experimental phase residual was close to or better than that expected based on the signal-to-noise ratio.

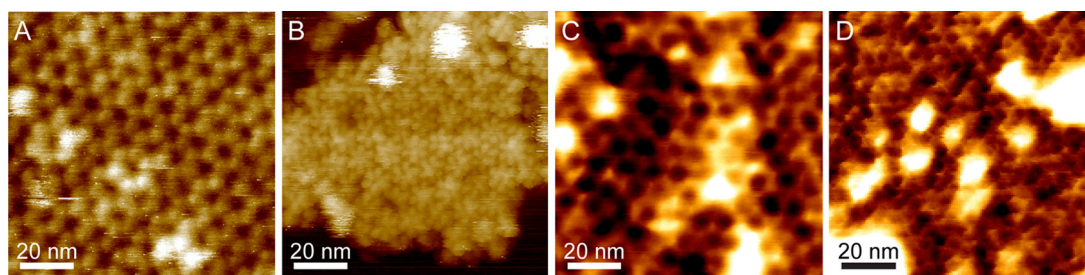
A projection map of negatively stained CsxA crystals was calculated by averaging 5 images (Fig. S1B). The unit cell parameters ( $a = b = 111 \pm 2 \text{ \AA}$  and  $\gamma = 120.3^\circ \pm 0.4^\circ$ , *p*6 symmetry [inferred from electron cryomicroscopy [CryoEM]; Table 1]) and projection structure are identical to those of the native exosporium (Fig. S1). We recorded 66 images from 11 tilt series of negatively stained two-dimensional (2D) crystals of CsxA and calculated a 3D reconstruction using the method described for analysis of the native exosporium (shown superimposed in Fig. 6A; see also Table S1 and Movie S1 in the supplemental material). The structures were identical, confirming that CsxA is the core protein component of the exosporium basal layer lattice.

**Electron cryomicroscopy (CryoEM) of recombinant CsxA crystals reveals a ring of six protein subunits at 9-Å resolution.** The CsxA 2D crystals were better ordered than those seen with the native exosporium and were more amenable to higher-resolution analysis. The unit cell dimensions in vitreous ice were  $a = b = 103 \pm 2 \text{ \AA}$  and  $\gamma = 120 \pm 2^\circ$ . Analysis of high-resolution Fourier phases indicated *p*6 symmetry, thus unambiguously determining the hexameric nature of the protein assembly (Table 1). We calculated a projection map from averaged amplitudes and phases from 14 images. The phase measurements were significant to  $\sim 9\text{-\AA}$  resolution (Table S2). The projection map (Fig. 6B) reveals a 6-fold symmetric ring of protein density with an outer diameter of  $\sim 115 \text{ \AA}$ . The ring encloses a less dense core (center of Fig. 6B). The possible approximate envelope of one subunit is outlined, and it is notable that the closest points of contact between subunits are within the hexameric ring and in the vicinity of



**FIG 6** The 3D structure of the recombinant CsxA crystal is superimposable on that of native exosporium. (A) CsxA density superimposed on that of native exosporium. Surfaces are rendered at a density threshold roughly equivalent to that expected for a 33-kDa monomer. Solid gray represents native exosporium, and yellow mesh represents recombinant CsxA. (B) Projection map of vitreous ice-embedded CsxA crystal at 9-Å resolution. The map is contoured at  $0.2 \times$  root mean square (RMS) deviation from the mean density. Solid contours represent density above the mean density; dashed contours represent density below the mean density. The repeating unit is a hexameric ring (dashed circle) of protein density enclosing a less dense cavity. The triangle marks the axis of 3-fold symmetry where rings appear connected. The density for one potential subunit is outlined.





**FIG 7** AFM of recombinant CsxA crystals revealing the external surface of the basal layer and a conformational change depending on hydration. (A) Height image of recombinant CsxA 2D crystal fragment taken in water, showing a honeycomb-like lattice with a unit cell measurement of  $103 \pm 8 \text{ \AA}$ . Color scale, 6 nm. (B) Height image of recombinant CsxA 2D crystal fragment taken in water, showing a hexagonal lattice of flower-like structures with a unit cell measurement of  $100 \pm 6 \text{ \AA}$ . Color scale, 15 nm. (C) Height image of a recombinant CsxA 2D crystal fragment taken in air, showing a honeycomb-like lattice with a unit cell measurement of  $102 \pm 2 \text{ \AA}$ . Color scale, 2.5 nm. (D) Height image of a recombinant CsxA 2D crystal fragment taken in air, showing a lattice with an apparent spacing of  $50 \pm 4 \text{ \AA}$ . Color scale, 2.5 nm.

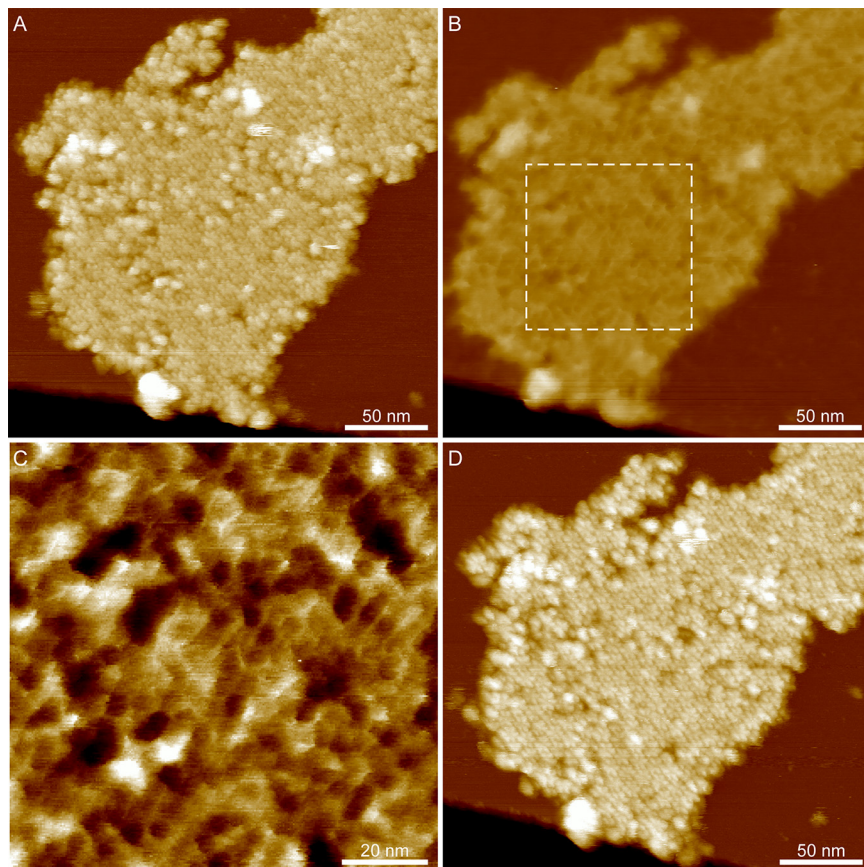
the 3-fold symmetry axes that connect rings; this is consistent with the lower-resolution 3D reconstruction (Fig. 2).

#### AFM of CsxA crystals reveals a reversible conformational change on hydration.

2D crystals of CsxA lack the external decoration present on native exosporium fragments, allowing both surfaces of the basal layer to be imaged with AFM. Two distinct surface structures, both with a unit cell parameters of  $\sim 110 \text{ \AA}$ , were observed in AFM images of hydrated CsxA crystals. One surface showed a honeycomb lattice of pits (Fig. 7A) similar to that seen in native exosporium (Fig. 3F), meaning that we can confidently assign it to the internal surface of the basal layer. The other (external) surface displayed a lattice of hexameric assemblies with petal-like lobes (Fig. 7B). When samples were dried and imaged in air, we observed little difference in the overall arrangement of the  $110\text{-\AA}$  lattice of pits on the honeycomb face except that the 3-fold linkers were less pronounced, as seen with dehydrated native exosporium (Fig. 3B; see also Fig. 7C). On the “petal” face, instead of an  $\sim 110\text{-\AA}$  lattice, we observed an array of pits with apparent  $\sim 50\text{-\AA}$  spacing (Fig. 7D). However, the Fourier transform showed weak first-order spots, indicating that the true unit cell parameter was still  $\sim 110 \text{ \AA}$ . The overall sheet thickness of crystals decreased from  $\sim 65 \text{ \AA}$  in water to  $\sim 40 \text{ \AA}$  in air, regardless of which surface was exposed to the AFM tip (Fig. S3C to F). A cycle of dehydration followed by rehydration on crystals displaying the petal face showed the structural change to be reversible (Fig. 8). The mechanism and functional implications of this change remain to be determined. However, extrapolating to the native exosporium, both hydrated and dehydrated basal layer structures are likely to represent *in vivo* states, reflecting the different environments which spores would experience, such as dry to water-saturated soils or surfaces and the environments inside infected hosts or predators.

**Assignment of internal and external surfaces to the 3D EM reconstruction.** By comparing the EM reconstructions of native exosporium and CsxA crystals (Fig. 2 and 6A) with the respective AFM images (Fig. 3 and 7), we can tentatively assign the surface of the *C. sporogenes* EM reconstruction of native exosporium (Fig. 2) as follows: the  $\sim 105\text{-\AA}$  honeycomb lattice observed on the internal surface of dehydrated native exosporium fragments by AFM (Fig. 3B) corresponds to the face showing the raised trimeric linkers in the EM reconstruction (Fig. 2C). In some cases, raised features were observed by AFM at the expected linker locations (Fig. 7C). These became more prominent under conditions of imaging in water (Fig. 3F and 7A). We propose that the apparent  $\sim 50\text{-\AA}$  lattice of pits observed on the other face of dehydrated CsxA crystals by AFM (Fig. 7D) corresponds to the larger central pores and smaller pores around the periphery of the raised hexamer surface (Fig. 2A and B and 6A). We suggest that, when hydrated, the raised hexamer (Fig. 2A and B) swells and dominates the topography, giving rise to the lattice of petals observed by AFM in liquid (Fig. 7B).





**FIG 8** AFM height images of CsxA showing a conformational change between wet and dry conditions. (A) A recombinant CsxA 2D crystal imaged in water showing a hexagonal lattice of flower-like arrays with a lattice parameter of  $\sim 110$  Å. Color scale, 22 nm. (B) The same fragment of CsxA shown in panel A after drying, imaged in air. The flower-like arrays are no longer visible, and a hexagonal lattice of pores is present, with an apparent spacing of  $\sim 50$  Å. Color scale, 22 nm. (C) Magnified image of the area indicated by the dashed box in panel B. Color scale, 2 nm. (D) The same fragment imaged in water after rehydration. The flower-like arrays became visible again, indicating that the conformational change was reversible. Color scale, 22 nm.

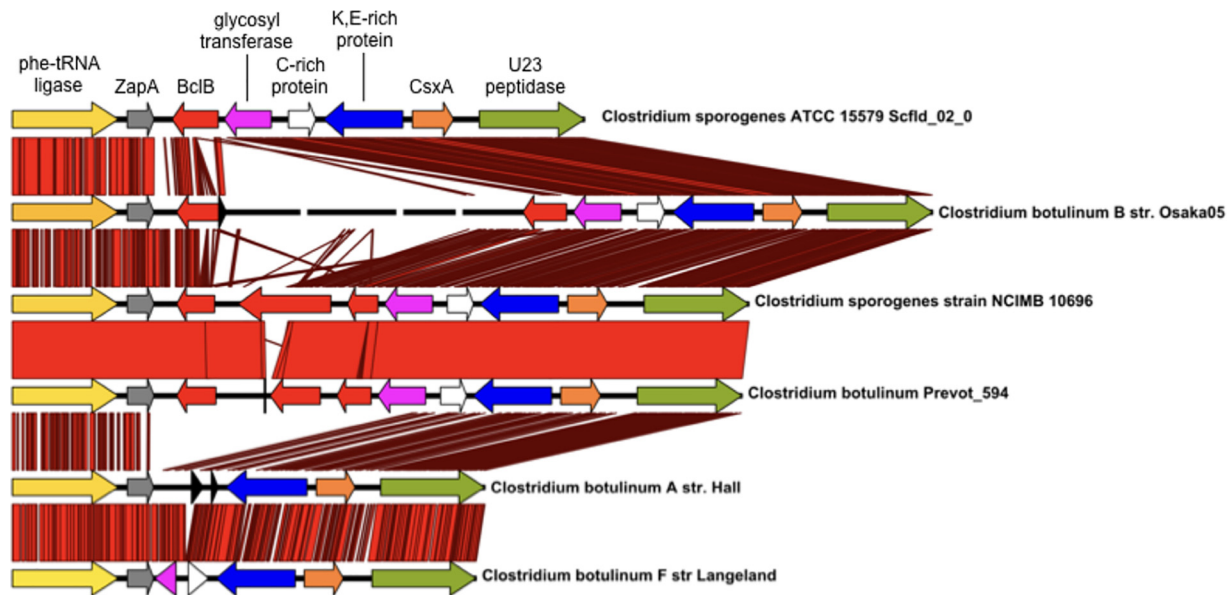
**CsxA arrays are thermally stable except under harsh “reducing” conditions.**

CsxA crystals were exposed to a variety of denaturing conditions (Table 2). As complete disassembly of CsxA crystals requires boiling in the presence of 2 M dithiothreitol (DTT), it is likely that disulfide bonding between some or all of the 25 cysteines of CsxA plays a critical role in holding together the crystal lattice. The projection map of ice-embedded CsxA (Fig. 6B) suggests that the six monomers within a single hexameric ring are closely packed and may be connected by multiple disulfide bonds. However, the packing between rings appears less tight and it is likely that cross links occur only in the vicinity of the 3-fold symmetric bridge (Fig. 6B, triangle). This could explain why

**TABLE 2** Effect of incubation of CsxA crystals under different conditions<sup>a</sup>

Treatment	Crystals observed	Crystalline order	Crystal abundance
None	Yes	High	High
8 M urea	Yes	Low-high	Low
95°C, 60 min	Yes	High	High
2 M DTT, 25°C	Yes	Low	Low
2 M DTT, 95°C	No		

<sup>a</sup>Crystalline order is scored as follows: High, diffraction beyond second order; Low-high, diffraction between first order and second order; Low, diffraction up to first order. Crystal abundance is scored as follows: High, more than 5 crystals/grid square; Low, fewer than 1 crystal/grid square.



**FIG 9** Alignment of the genomic context of *CsxA* in *C. sporogenes* and *C. botulinum* group I strains, from *pheT* to a U32 family peptidase gene. Regions of sequence similarity between the strains are shown in red, with brighter shades indicating higher levels of sequence identity. The region between *zapA* (gray: cell division gene) and the gene encoding a U32 peptidase (olive green) contains 7 ORFs in *C. sporogenes* NCIMB 10696, several of which correspond to protein sequences detected in the exosporium. The *csxA* gene is shown in orange and is located at the right hand end of the cluster. Three ORFs in red at the left end each contain a central domain of collagen-like repeats (these are [left to right] *bclB* [CLSPO\_c32280] and *bclA* [CLSPO\_c32270], which encode proteins identified in the *C. sporogenes* exosporium [20], followed by CLSPO\_c32260, which encodes a protein with a BclB-like C-terminal domain. The adjacent gene (shown in pink) encodes a glycosyl transferase, which may be responsible for glycosylation of one or more of the collagen-like repeat proteins, and is followed by an ORF encoding a cysteine-rich protein (white) and a highly conserved ORF (blue), organized divergently from *csxA*. *C. sporogenes* strain ATCC 15579 and *C. botulinum* Osaka 05 have a smaller complement of *bcl*-like ORFs, and *C. botulinum* strains Hall and Langeland retain intact homologues of only *csxA* and its divergent partner.

DTT treatment of *CsxA* crystals at room temperature leads to increased crystal disorder but not to complete disassembly (Fig. S4).

**The *CsxA* protein is essential for formation of the exosporium.** A *csxA* mutant of *C. sporogenes* was constructed in the more genetically amenable ATCC 15579 strain. Spores of this mutant no longer have a typical 110-Å lattice exosporium, consistent with the interpretation that *CsxA* is the core structural component of the outermost exosporium basal layer. Instead, spores appeared partially wrapped in thin broken sheets of material, fragments of which were frequently sloughed off the spores (Fig. 1C). These sheets formed 2D crystals but with a trigonal rather than hexagonal unit cell measurement of ~65 Å on a side (Fig. S5). This loose proteinaceous layer may derive either from the coat or from some additional exosporium sublayer that is normally more tightly associated with the spore core (23).

**Genes encoding cysteine-rich exosporium proteins and collagen-like repeat (CLR) domain proteins are collocated in distantly related spore formers.** The *CsxA* protein is conserved across *C. sporogenes* and related group I *C. botulinum* species (20), and *CsxA* homologues are also present in a wide variety of other *Clostridium* species (Fig. S6). In *C. sporogenes* strain ATCC 15579, and others, the *CsxA* protein is encoded in a gene cluster between *zapA* and a gene encoding a U32 family peptidase (Fig. 9). In ATCC 15579, the cluster also encodes a glycosyl transferase and proteins containing collagen-like repeats (CLR domains), including the BclA protein. BclA is found associated with *CsxA* in high-molecular-weight exosporium protein complexes, and the CLR domain protein, BclB, is detected in bulk exosporium (20). By analogy with *B. cereus* and *B. anthracis*, at least some of these CLR domain proteins are likely glycoproteins forming the hairy nap on the surface of the exosporium. This is reminiscent of *B. anthracis*, in that genes encoding the three proteins that are found in the high-molecular-weight exosporium complexes (BclA, ExsFA, and the cysteine-rich protein ExsY) are present in a cluster, along with those encoding glycosyl transferase and other genes (24). In *C.*

*sporogenes*, no other identified exosporium proteins are encoded at this position, although other conserved open reading frames (ORFs) may be present.

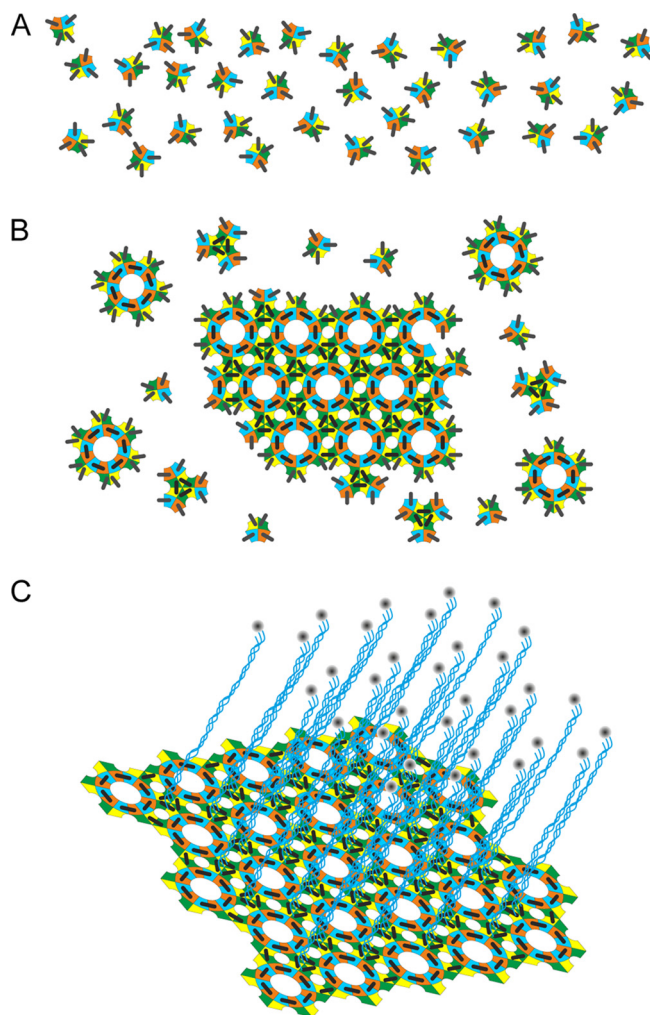
Gene synteny, albeit with variations in the CLR-encoding ORFs, is found in other *C. sporogenes* and group I *C. botulinum* strains (Fig. 9); more-distant *csxA* homologues (e.g., those in *C. tyrobutyricum*, *C. puniceum*, and *C. beijerinckii*) are also frequently collocated with genes encoding CLR proteins and glycosyl transferases, though not necessarily in the same genomic context. More generally, CLR domain proteins and cysteine-rich proteins can be found encoded within 20 kb of each other across many spore formers (see Data Set 1 in the supplemental material); a number of these cysteine-rich proteins are annotated as spore coat proteins, and others are potential candidates for identification as spore coat and/or exosporium proteins.

**Distantly related spore formers use different proteins but adopt similar design principles to build the spore envelope.** The representatives of the *Bacillales* and *Clostridiales* that we have studied (*B. cereus*-*B. anthracis*, *C. sporogenes*-*C. botulinum*, *C. acetobutylicum*, *C. tyrobutyricum*, *C. puniceum*, and *C. pasteurianum*) possess exosporia of remarkably similar morphologies, with a crystalline basal layer enveloping the spore core and decorated by a more extensively disordered hairy nap (Fig. 1A and B) (25); in the cases of *B. cereus* and *C. sporogenes*, a disordered outer surface and ordered inner surface are observed by AFM (Fig. 3) (7, 20). In *B. cereus*, *B. anthracis*, *C. botulinum*, *C. sporogenes*, and other clostridial species, the basal layer of the exosporium has a regular tiling pattern of interlinked 6-fold symmetric oligomers (compare data in Fig. 2 and 3 and Fig. S1 and S2 with reference 7 data). In *B. cereus*, the core components are ExsY and CotY (8); these are cysteine rich and analogous to but not homologous to CsxA. The lattice spacings in the *Clostridia* tested were larger ( $\sim 110$  to  $\sim 127$  Å versus  $\sim 80$  Å in *B. cereus* and *B. anthracis*), and the *C. sporogenes* basal layer revealed an apparently more permeable structure, with pores of  $\sim 55$  Å, compared to  $\sim 20$  Å (25) (Fig. 2; see also Fig. S2). Although the effective diameters for diffusion would be less than those physically measured in the reconstructions (26), all would allow the permeation of small-molecule germinants (27, 28) but might still exclude hostile enzymes and antibodies.

The similar designs of the hexagonally tiled meshes across the basal layers of spore formers reflect apparently common functions with respect to acting as semipermeable molecular filters and forming a platform onto which other proteins may bind, including proteinaceous appendages. In the *B. cereus* group, the CLR-domain protein BclA is attached to ExsY via an anchor protein, ExsFA/BxpB (10). There is no ExsF homologue in *Clostridia*, and there is no evidence of a third protein in the CsxA-BclA complexes extracted from *C. sporogenes* (20), so whether BclA is directly anchored to the CsxA basal layer is not known. Notably, a second *C. sporogenes* cysteine-rich protein, CsxB, was detected in association with BclA, but not in association with CsxA, in somewhat smaller complexes by SDS-PAGE, and CsxB was much more easily dissociated into likely oligomers and monomers; its role in the exosporium, like that of BclA of *C. sporogenes* and those of other exosporium proteins identified by Janganan et al. (20), remains to be directly tested.

**Self-assembling cysteine-rich proteins form robust protective layers in spores from distantly related species.** The presence of cysteine-rich proteins is emerging as a characteristic feature of the proteomes of spores (8, 13, 20, 29–32). We have now identified a variety of spore proteins, some with very different sequences, that have the common properties of being cysteine rich and capable of self-organization into extended 2D ordered arrays resembling the natural assemblies found in the native spore. It is notable that the identical crystal packing symmetries that we see in CsxA assemblies have been found in the unrelated cysteine-rich proteins, including CotY from *B. subtilis* spore coat (31) and ExsY and CotY from *B. cereus* exosporium (8). Arrays of these proteins isolated from an *E. coli* expression host also require harsh reducing conditions and boiling for complete disassembly. This robust cross-linking of arrays is likely to reflect the situation in the native spore—harsh denaturing and reducing conditions are required for complete disintegration of the hexagonal lattice of *C. botulinum* and *B.*





**FIG 10** Two groups of distantly related spore formers show the same hierarchy of assembly states on the spore surface. The schematic diagram indicates the hierarchy of assembly states of exosporium in the *B. cereus* group and the *C. botulinum* group. (A and B) Monomeric units of ExsY in the *B. cereus* group and CsxA in the *C. botulinum* group (A) self-assemble into a hexameric disulfide cross-linked array (B). The high symmetry of the growing lattice would enhance the avidity for disulfide bond formation—represented by the black bars. Note that the number of cysteine residues participating in cross-linking is unknown. The colors represent schematically different surface regions of the asymmetric monomers. (C) In both groups, the surface lattice is embellished with a “hairy nap,” known to consist mostly of the CLR protein BclA in the case of the *B. cereus* group. In the case of the *C. botulinum* group, it is likely to consist of the corresponding BclA equivalent.

*cerus* exosporium (8, 33). While the cellular milieu of the native mother cell, where the exosporium is assembled, or of the heterologous *E. coli* host is generally considered reducing and intracellular disulfide bonding is rare, the ordered lattice and high symmetry of the respective proteins are likely to be critical factors in the assembly mechanism. As shown in Fig. 10, the high symmetry is likely to provide sufficient avidity to overcome the reducing environment and drive cooperative intracellular disulfide cross-linking.

As we did previously for ExsY (8), for CsxA we can describe a hierarchy of units of assembly in the protein lattice, from monomer through hexamer to the extended cross-linked 2D hexagonal array. In both cases, this array of cysteine-rich proteins can then be decorated by filamentous appendages, including proteins with CLR domains (Fig. 10). We propose that the many cysteine-rich coat and exosporium proteins identified in the genomes of spore formers are likely to have structural roles similar to those of CsxA, ExsY, and CotY. Indeed, disulfide bonding plays a central role in the

development of fully assembled spores. Reduction of spore disulfide bonds sensitizes *Bacillus* and *Clostridium* spores to lysozyme and hydrogen peroxide, thus emphasizing their protective role (34). Moreover, Aronson and Fitz-James (35) found evidence for the role of self-organization by partially reconstituting outer-coat layers of *B. cereus* that had been treated initially with DTT and urea.

It may not be a general requirement for self-assembly of cysteine-rich proteins that they form the extensive crystalline arrays that we have described; some may be only locally ordered. Nevertheless, extensive crystalline layers of unidentified proteins are associated with spore coats and exosporium-like layers in a wide range of spore-forming species (36–44).

**Role of the exosporium in outgrowth.** We have shown that those parts of the *in situ* exosporium that are accessible for diffraction analysis adopt a basal layer structure identical to that of recombinant CsxA crystals (Fig. 1A; see also Fig. S1 and Fig. 6A). However, it is possible that the spore is not completely enclosed by a uniform and continuous CsxA lattice. Although low-resolution scanning electron micrographs suggest a mostly uniform exosporium fully enveloping the *C. sporogenes* spore, there does appear to be a relatively weak structure (sporiduct) at one pole through which the outgrowing cell emerges (45). This may be made up of different proteins and could incorporate a “cap” as seen in *B. anthracis* (46, 47). The exosporium does not appear to tear beyond this region, indicating that the CsxA lattice is relatively resistant to any pressure from the emerging vegetative cell.

**Conclusion.** We have achieved the first three-dimensional molecular reconstruction of a clostridial spore surface. The *C. sporogenes* spore is enveloped by an exosporium built with a paracrystalline basal layer, the core component of which is the cysteine-rich protein CsxA. This is the first example of a clostridial protein capable of self-assembly to form a paracrystalline cross-linked spore layer, a phenomenon that we previously demonstrated in unrelated spore surface proteins of the *Bacillales* (8, 31). We have demonstrated how apparently unrelated proteins from different species can assemble to form similar highly symmetric tiled arrays. The proteins are cysteine rich and self-assemble to form highly resilient spore layers. We propose that diverse cysteine-rich proteins identified in the genomes of a broad range of spore formers may adopt a similar strategy for assembly.

## MATERIALS AND METHODS

**Spore preparation.** Spores of *C. sporogenes* NCIMB 701792 (NCDO1792) were grown on BHIS agar (brain heart infusion agar supplemented with 0.1% L-cysteine and 5 mg/ml yeast extract) as described previously (20, 48). Spores of *C. acetobutylicum* NCIB 8052, *C. tyrobutyricum* NCDO 1756, *C. puniceum* BL 70/20, and *C. pasteurianum* NCDO 1845 were grown on potato extract agar plates. Aliquots (100  $\mu$ l) of a meat culture that had been cooked overnight were spread on duplicate potato extract agar plates and incubated in an anaerobic chamber for 3 days at 30°C. Colonies were harvested from the surface of the plates using a sterile spreader and 1 ml of chilled saline solution. Plates were washed with a further 5 ml of saline solution, and the spore suspension and washings from duplicate plates were combined in a centrifuge tube. Cultures containing spores, cells, and cell debris were pelleted by centrifugation at 3,000  $\times g$  for 15 min at 4°C. The pellets were washed three times in 10-ml saline solution with centrifugation at 2,000  $\times g$  for 10 min at 4°C. Spores were separated from vegetative cells and debris on the basis of density. Washed pellets were resuspended in 1 ml saline solution, layered onto 10 ml of a 50% solution of sodium/meglumine diatrizoate (50 ml Urografin 370 [76%] mixed with 26 ml saline solution) and centrifuged at 5,500  $\times g$  for 60 min at 4°C. The top layers were discarded, while the pellet was resuspended in 20 ml saline solution. Pellets were then washed a further four times in saline solution (2,000  $\times g$ , 15 min, 4°C), resuspended in 0.5 ml saline solution, and stored at 2°C until required.

**csxA mutant construction.** Mutants of *C. sporogenes* strain ATCC 15579 were generated using a Clostron system as described previously (27). *C. sporogenes* strain NCIMB 701792 (NCDO1792) was not used for mutational studies as this strain is erythromycin resistant and not amenable to Clostron mutation. Target sites were identified in the *csxA* gene (gene CLOSPO\_00498 [*csxA*], insert site 345/346s) using a method described previously by Perutka et al. (49), and mutants were generated as described previously (50). Retargeted introns were synthesized and ligated into the pMTL007C-E2 vector by the use of DNA 2.0 (ATUM, Menlo Park, USA). Plasmids were transformed into *E. coli* CA434 and then transferred by conjugation into *C. sporogenes* strain ATCC 15579.

**Expression of *csxA* in *E. coli*.** Genomic DNA from *C. sporogenes* NCIMB 701792 was used as a template to amplify the *csxA* sequence by PCR using Phusion DNA polymerase (NEB), with forward primer ATCTA**CATATG**GCTATTAATTCAAAGATTTATTCCAC and reverse primer ATATT**CTCGAG**ATTATTAGTTAT TACTGCTAGTTATC (where the boldface underlined characters indicate the restriction sites of NdeI and

XhoI, respectively). Oligonucleotides were designed based on the genome sequence of *C. sporogenes* PA 3679 strain. The amplified product, containing a C-terminal His<sub>6</sub> fusion, was cloned between the NdeI and XhoI sites in pET21a. The resulting plasmid was transformed into *E. coli* BL21(DE3) pLysS for protein production.

For protein overexpression, *E. coli* cells were grown to an optical density at 600 nm (OD<sub>600</sub>) of 0.5 in LB broth and induced with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) for 3 h at 37°C. Cells were harvested and resuspended in spore resuspension buffer (SRB; 25 mM Tris [pH 8], 150 mM NaCl, 1 mM phenylmethylsulfonyl fluoride [PMSF]) and sonicated for 30 s at an amplitude of 10 μm; sonication was repeated 3 to 5 times with 1-min intervals. The cell lysates were mixed with nickel-nitrilotriacetic acid (Ni-NTA) agarose and incubated on an end-over-end rotator at room temperature for 60 min. The Ni-NTA agarose was allowed to settle by gravity, and the supernatant was discarded. The Ni-NTA beads were resuspended in buffer (25 mM Tris [pH 8], 150 mM NaCl), packed in a gravity flow column, and washed several times with buffer. The Ni-NTA agarose was transferred to a Falcon tube, and the protein was eluted with buffer (25 mM Tris [pH 8], 150 mM NaCl, 1 M imidazole). The eluted protein was centrifuged at 100,000 × *g* for 1 h, and the pellet containing CsxA 2D crystals was washed and resuspended in Tris buffer.

CsxA crystals were examined by negative-stain EM (see below) after incubation under a variety of combined conditions, including 95°C heat treatment and 8 M urea and 2 M DTT treatments (see Table 2).

**Electron microscopy of *E. coli* cell sections.** *E. coli* cells overexpressing CsxA were pelleted (100 μl) and fixed with fresh 3% glutaraldehyde–0.1 M phosphate buffer overnight at 4°C. The sample was then washed in 0.1 M phosphate buffer two times at 30-min intervals at 4°C. Secondary fixation was carried out in 2% aqueous osmium tetroxide for 2 h at room temperature, and the sample was washed as described above. The sample was dehydrated using a series of ethanol washes and finally dried over anhydrous copper sulfate for 15 min. The sample was then placed in two changes of propylene oxide for 15 min. Infiltration was achieved by placing the sample in a 50/50 mixture of propylene oxide/araldite resin overnight at room temperature. The sample was left in full-strength araldite resin for 6 to 8 h at room temperature, after which it was embedded in fresh araldite resin for 48 to 72 h at 60°C. Ultrathin sections, approximately 70-nm to 90-nm thick, were cut on a Reichert Ultracut E ultramicrotome and stained for 25 min with 3% uranyl acetate followed by staining with Reynold's lead citrate for 5 min. Sections were examined on a FEI Tecnai 120 G2 Biotwin electron microscope at 80 kV with a Gatan Orius SC 1000B digital camera (bottom mounted).

**Negative-stain electron microscopy (EM).** Spore suspension or sonicated *E. coli* cells (2 μl) were applied to glow-discharged carbon-coated copper palladium grids, incubated for 1 min, and then washed, stained (20 s) with uranyl formate (0.75%), and vacuum dried. For CsxA crystals extracted from sonicated *E. coli* cells, after sample incubation, the grids were washed once in distilled water before being stained. Grids were examined in a Philips CM100 transmission electron microscope operating at 100 kV. Micrographs were recorded under low-dose conditions on a Gatan MultiScan 794 1,000-pixel-by-1,000-pixel charge-coupled-device (CCD) camera at nominal ×52,000 magnification with 0.5-to-1.2-μm underfocus and at specimen tilts over a range of ±55° in 10° steps.

**Electron cryomicroscopy (CryoEM).** Recombinant CsxA 2D crystals purified as described above were resuspended in Tris buffer (pH 7.5; 150 mM NaCl and 1 mM EDTA) and loaded onto glow-discharged carbon-coated molybdenum grids or Quantifoil R2/2 grids with carbon support and incubated for 1 min. Grids were blotted and plunge-frozen in liquid ethane using an FEI Vitrobot plunge freezer or Leica plunge freezer; the blotting time used was 25 s or 4 to 6 s, respectively.

Samples were examined on a Philips CM200 field emission gun (FEG) EM, equipped with an Oxford Instruments CT3500 cold stage, or a FEI Tecnai Arctica FEG EM, each operated in low-dose mode at 200 kV and at liquid-nitrogen temperature. Data collected on the CM200 FEG EM were at a nominal magnification of ×66,000 and were recorded on a CCD Gatan UltraScan 890 (US4000SP) 2,000-pixel-by-2,000-pixel camera, with defocus values of ~1 μm and 0.5-s exposure. Micrographs were recorded using a Tecnai Arctica EM at magnifications of ×39,000 to ×78,000 on a Falcon 3 direct electron 4,000-pixel-by-4,000-pixel detector at defocus values of ~1.5 to 6.5 μm with either a 1-s single exposure or a 2-s exposure over 79 frames and were motion corrected with MotionCorr2.

**Image processing.** Electron micrographs were processed within the 2dx software suite (48) based on the MRC suite of programs (51). Images were subjected to two cycles of unbending. For all subsequent analyses, *p*6 symmetry was enforced. Phase origins for individual images were refined against each other using ORIGINLTD, sequentially adding images of increasing tilt angle to the refinement. Initial estimates of tilt angle were made from the more highly tilted members of a series using EMTILT (52). The common phase origin was found by comparing the phases of the reflections on each image within a *z*\* value of 0.01 Å<sup>-1</sup> to those of all the other images. At least two cycles of refinement of the phase origin, tilt angle, and tilt axis were performed for each image. The program LATLINE (53) was used to determine interpolated amplitudes and phases on a regular lattice of 1/270 Å<sup>-1</sup> in *z*\* for data up to 1/25-Å<sup>-1</sup> resolution. A generous real-space envelope of approximately twice the estimated exosporium thickness (70 Å) was applied as a constraint. The output interpolated lattice lines were used as references for two cycles of crystal tilt and phase origin refinement. The variation of amplitude and phase along 0, 0, and *l* was estimated from the maximum contrast on each *Z*-section in the 3D map (54). Density maps were calculated within the CCP4 suite of crystallography programs (55). 3D surface representations were rendered with CHIMERA (56).

For analysis of frozen hydrated CsxA crystals, we merged and averaged 14 images in 2dx (57) to 9-Å resolution. A negative-temperature factor (B-factor) was estimated and applied to the projection map by scaling averaged image amplitudes against bacteriorhodopsin electron diffraction amplitudes (58).



**Atomic force microscopy (AFM).** Exosporium fragments were prepared as described previously (20). A 10- $\mu$ l volume of spore fragment solution (concentration, 2.2 mg/ml) was diluted in 100  $\mu$ l citric acid/sodium phosphate buffer (150 mM KCl [pH 4]) and incubated on poly-D-lysine-coated coverslips (Corning BioCoat) for 30 min at room temperature.

2D crystals of CsxA were prepared for AFM by incubating 2 to 5  $\mu$ l of crystals in storage buffer (20 mM Tris, 150 mM NaCl [pH 8]) on freshly cleaved mica for  $\sim$ 30 min at room temperature. After binding, all samples were washed with 10 $\times$  1 ml high-performance liquid chromatography (HPLC)-grade deionized water and either imaged in water or dried with filtered nitrogen and imaged in air.

Imaging in air was performed using a JPK NanoWizard Ultra AFM in AC mode with TESPA V2 cantilevers (nominal stiffness, 37 N/m; nominal resonant frequency, 320 kHz) in a home-built vibration and acoustic isolation system. The free amplitude was approximately 8 nm and the relative set point 90% to 95%. Imaging in water was performed using a Bruker Dimension FastScan AFM in tapping mode with FastScan D cantilevers (nominal stiffness, 0.25 N/m; nominal resonant frequency [in water], 100 kHz). The cantilever holder was washed in a mixture of household detergent, isopropanol, and pure water before each experiment. Imaging was performed in a small ( $\sim$ 200- $\mu$ l) water droplet. The free amplitude was approximately 1 nm and the relative set point 80% to 90%.

To minimize the effects of thermal drift, the sample was placed into the AFM and allowed to settle for approximately 1 h (experiments in water) or 2 to 24 h (experiments in air) with the isolation hood closed. In all cases, the probe was tuned close to the surface after engagement to compensate for squeeze film damping. The feedback gains, scan rate, pixel density, and amplitude set point were adjusted for optimal image quality while scanning. To ensure accuracy of all measurements, scanner calibration was checked using a 3- $\mu$ m pitch and 180-nm depth grating and found to be within 1% for all 3 axes.

Images were processed (flattening, plane fitting) and analyzed using JPK DP software, Gwyddion, and NanoScope.

**Phylogenetic analysis of CsxA homologues.** To identify homologues of *csxA*, all available complete and draft genome sequences of spore-forming bacteria were downloaded from GenBank, and the annotated protein sequences were extracted and converted into a BLAST database for each strain. These were interrogated with BLASTp, using the CsxA sequence of *Clostridium sporogenes* ATCC 15579 as a query. Any hits with >20% amino acid identity to CsxA over at least 50% of the length of the query sequence were retained, resulting in the identification of 265 potential CsxA homologues. Duplicate sequences were removed, giving a data set of 117 distinct protein sequences of CsxA homologues. These were aligned using Muscle version 3.8.31 (59), and a phylogenetic tree was constructed with RAxML version 8.2.12 (60) using the VT model of amino acid substitution (61), which was selected as the best-scoring amino acid model by RAxML, and a gamma model of rate heterogeneity. Alignments of the genomic context of CsxA in *C. sporogenes* and *C. botulinum* group I strains were performed using NUCmer (62) and displayed using a custom version of the xBASE database (63).

**Data availability.** Raw data and materials can be provided by the authors on request.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, PDF file, 0.8 MB.

**FIG S2**, PDF file, 1.4 MB.

**FIG S3**, JPG file, 0.9 MB.

**FIG S4**, PDF file, 1.7 MB.

**FIG S5**, PDF file, 1.9 MB.

**FIG S6**, PDF file, 0.4 MB.

**TABLE S1**, PDF file, 0.04 MB.

**TABLE S2**, PDF file, 0.05 MB.

**DATA SET S1**, XLSX file, 0.8 MB.

**MOVIE S1**, MOV file, 11.7 MB.

## ACKNOWLEDGMENTS

We thank Chris Hill for help with EM of thin sections. All electron microscopy work was carried out in the University of Sheffield's Faculty of Science Electron Microscopy Facility. We also thank Paul Kemp-Russell and Simon Dixon, who fabricated the acoustic enclosure for AFM.

J.K.H. and N.M. gratefully acknowledge the Imagine: Imaging Life initiative at the University of Sheffield and the EPSRC for financial support through its Program Grant scheme (Grant No. EP/I012060/1). P.A.B. and T.K.J. gratefully acknowledge the Wellcome Trust for financial support. A.D.-S. was in receipt of a White Rose BBSRC DTP studentship (grant no. BB/M011151/1).

T.K.J., N.M., and P.A.B. designed research; T.K.J., N.M., A.D.-S., J.B., S.B.T., S.S., R.C., and

P.A.B. performed research; T.K.J., N.M., A.D.-S., A.M., R.C., R.P.F., J.K.H., and P.A.B. analyzed data; T.K.J., N.M., A.M., R.P.F., J.K.H., and P.A.B. wrote the paper.

We declare no competing interests.

## REFERENCES

- Setlow P. 2014. Germination of spores of bacillus species: what we know and do not know. *J Bacteriol* 196:1297–1305. <https://doi.org/10.1128/JB.01455-13>.
- Galperin MY, Mekhedov SL, Puigbo P, Smirnov S, Wolf YI, Rigden DJ. 2012. Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. *Environ Microbiol* 14:2870–2890. <https://doi.org/10.1111/j.1462-2920.2012.02841.x>.
- Driks A, Eichenberger P. 2016. The spore coat. *Microbiol Spectr* 4(2). <https://doi.org/10.1128/microbiolspec.TBS-0023-2016>.
- Stewart GC. 2015. The exosporium layer of bacterial spores: a connection to the environment and the infected host. *Microbiol Mol Biol Rev* 79:437–457. <https://doi.org/10.1128/MMBR.00050-15>.
- Gu C, Jenkins SA, Xue Q, Xu Y. 2012. Activation of the classical complement pathway by *Bacillus anthracis* is the primary mechanism for spore phagocytosis and involves the spore surface protein BclA. *J Immunol* 188:4421–4431. <https://doi.org/10.4049/jimmunol.1102092>.
- Wang Y, Jenkins SA, Gu C, Shree A, Martinez-Moczygemba M, Herold J, Botto M, Wetsel RA, Xu Y. 2016. *Bacillus anthracis* spore surface protein BclA mediates complement factor H binding to spores and promotes spore persistence. *PLoS Pathog* 12:e1005678. <https://doi.org/10.1371/journal.ppat.1005678>.
- Kailas L, Terry C, Abbott N, Taylor R, Mullin N, Tzokov SB, Todd SJ, Wallace BA, Hobbs JK, Moir A, Bullough PA. 2011. Surface architecture of endospores of the *Bacillus cereus/anthracis/thuringiensis* family at the subnanometer scale. *Proc Natl Acad Sci U S A* 108:16014–16019. <https://doi.org/10.1073/pnas.1109419108>.
- Terry C, Jiang S, Radford DS, Wan Q, Tzokov SB, Moir A, Bullough PA. 2017. Molecular tiling on the surface of a bacterial spore: the exosporium of the *Bacillus anthracis/cereus/thuringiensis* group. *Mol Microbiol* 104:539–552. <https://doi.org/10.1111/mmi.13650>.
- Sylvestre P, Couture-Tosi E, Mock M. 2005. Contribution of ExsFA and ExsFB proteins to the localization of BclA on the spore surface and to the stability of the *Bacillus anthracis* exosporium. *J Bacteriol* 187:5122–5128. <https://doi.org/10.1128/JB.187.15.5122-5128.2005>.
- Steichen CT, Kearney JF, Turnbough CL. 2005. Characterization of the exosporium basal layer protein BxpB of *Bacillus anthracis*. *J Bacteriol* 187:5868–5876. <https://doi.org/10.1128/JB.187.17.5868-5876.2005>.
- Abhyankar W, Hossain AH, Djajasaputra A, Permpoonpattana P, Ter Beek A, Dekker HL, Cutting SM, Brul S, de Koning LJ, de Koster CG. 2013. In pursuit of protein targets: proteomic characterization of bacterial spore outer layers. *J Proteome Res* 12:4507–4521. <https://doi.org/10.1021/pr4005629>.
- Permpoonpattana P, Phetcharaburanin J, Mikelsone A, Dembek M, Tan S, Brisson M, Ragione RL, Brisson AR, Fairweather N, Hong HA, Cutting SM. 2013. Functional characterization of *Clostridium difficile* spore coat proteins. *J Bacteriol* 195:1492–1503. <https://doi.org/10.1128/JB.02104-12>.
- Calderón-Romero P, Castro-Córdova P, Reyes-Ramírez R, Milano-Céspedes M, Guerrero-Araya E, Pizarro-Guajardo M, Olguín-Araneda V, Gil F, Paredes-Sabja D. 2018. *Clostridium difficile* exosporium cysteine-rich proteins are essential for the morphogenesis of the exosporium layer, spore resistance, and affect *C. difficile* pathogenesis. *PLoS Pathog* 14:e1007199. <https://doi.org/10.1371/journal.ppat.1007199>.
- Collins MDM, Lawson PA, Willems A, Cordoba JJ, Fernandez-Garayzabal J, Garcia P, Cai J, Hippe H, Farrow JA. 1994. The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol* 44:812–826. <https://doi.org/10.1099/00207713-44-4-812>.
- Yutin N, Galperin MY. 2013. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol* 15:2631–2641. <https://doi.org/10.1111/1462-2920.12173>.
- Lawson PA, Citron DM, Tyrrell KL, Finegold SM. 2016. Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prévot 1938. *Anaerobe* 40:95–99. <https://doi.org/10.1016/j.anaerobe.2016.06.008>.
- Peck MW, Stringer SC, Carter AT. 2011. *Clostridium botulinum* in the post-genomic era. *Food Microbiol* 28:183–191. <https://doi.org/10.1016/j.fm.2010.03.005>.
- Butler RR, Schill KM, Wang Y, Pombert J-F. 2017. Genetic characterization of the exceptionally high heat resistance of the non-toxic surrogate *Clostridium sporogenes* PA 3679. *Front Microbiol* 8:545. <https://doi.org/10.3389/fmicb.2017.00545>.
- Masuda K, Kawata T, Takumi K, Kinouchi T. 1980. Ultrastructure of a hexagonal array in exosporium of a highly sporogenic mutant of *Clostridium botulinum* type A revealed by electron microscopy using optical diffraction and filtration. *Microbiol Immunol* 24:507–513. <https://doi.org/10.1111/j.1348-0421.1980.tb02854.x>.
- Janganan TK, Mullin N, Tzokov SB, Stringer S, Fagan RP, Hobbs JK, Moir A, Bullough PA. 2016. Characterization of the spore surface and exosporium proteins of *Clostridium sporogenes*; implications for *Clostridium botulinum* group I strains. *Food Microbiol* 59:205–212. <https://doi.org/10.1016/j.fm.2016.06.003>.
- Panessa-Warren BJB, Tortora GTG, Warren J. 1997. Exosporial membrane plasticity of *Clostridium sporogenes* and *Clostridium difficile*. *Tissue Cell* 29:449–461. [https://doi.org/10.1016/S0040-8166\(97\)80031-6](https://doi.org/10.1016/S0040-8166(97)80031-6).
- Panessa-Warren BJB, Tortora GTG, Warren J. 2007. High resolution FESEM and TEM reveal bacterial spore attachment. *Microsc Microanal* 13:251–266. <https://doi.org/10.1017/S1431927607070651>.
- Stevenson KE, Vaughn RH. 1972. Exosporium formation in sporulating cells of *Clostridium botulinum* 78A. *J Bacteriol* 112:618–621. <https://doi.org/10.1128/JB.112.1.618-621.1972>.
- Todd SJ, Moir AJG, Johnson MJ, Moir A. 2003. Genes of *Bacillus cereus* and *Bacillus anthracis* encoding proteins of the exosporium. *J Bacteriol* 185:3373–3378. <https://doi.org/10.1128/jb.185.11.3373-3378.2003>.
- Ball DA, Taylor R, Todd SJ, Redmond C, Couture-Tosi E, Sylvestre P, Moir A, Bullough BA. 2008. Structure of the exosporium and sublayers of spores of the *Bacillus cereus* family revealed by electron crystallography. *Mol Microbiol* 68:947–958. <https://doi.org/10.1111/j.1365-2958.2008.06206.x>.
- Dechadilok P, Deen WM. 2006. Hindrance factors for diffusion and convection in pores. *Ind Eng Chem Res* 45:6953–6959. <https://doi.org/10.1021/ie051387n>.
- Brunt J, Plowman J, Gaskin DHJ, Itchner M, Carter AT, Peck MW. 2014. Functional characterisation of germinant receptors in *Clostridium botulinum* and *Clostridium sporogenes* presents novel insights into spore germination systems. *PLoS Pathog* 10:e1004382. <https://doi.org/10.1371/journal.ppat.1004382>.
- Barlase PJ, Houston CW, Clements MO, Moir A. 2002. Germination of *Bacillus cereus* spores in response to L-alanine and to inosine: the roles of gerL and gerQ operons. *Microbiology* 148:2089–2095. <https://doi.org/10.1099/00221287-148-7-2089>.
- Lawley TD, Croucher NJ, Yu L, Clare S, Sebahia M, Goulding D, Pickard DJ, Parkhill J, Choudhary J, Dougan G. 2009. Proteomic and genomic characterization of highly infectious *Clostridium difficile* 630 spores. *J Bacteriol* 191:5377–5386. <https://doi.org/10.1128/JB.00597-09>.
- Barra-Carrasco J, Olguín-Araneda V, Plaza-Garrido A, Miranda-Cárdenas C, Cofré-Araneda G, Pizarro-Guajardo M, Sarker MR, Paredes-Sabja D. 2013. The *Clostridium difficile* exosporium cysteine (CdeC)-rich protein is required for exosporium morphogenesis and coat assembly. *J Bacteriol* 195:3863–3875. <https://doi.org/10.1128/JB.00369-13>.
- Jiang S, Wan Q, Krajcikova D, Tang J, Tzokov SB, Barak I, Bullough PA. 2015. Diverse supramolecular structures formed by self-assembling proteins of the *Bacillus subtilis* spore coat. *Mol Microbiol* 97:347–359. <https://doi.org/10.1111/mmi.13030>.
- Rabi R, Larcombe S, Mathias R, McGowan S, Awad M, Lyras D. 2018. *Clostridium sordellii* outer spore proteins maintain spore structural integrity and promote bacterial clearance from the gastrointestinal tract. *PLoS Pathog* 14:e1007004. <https://doi.org/10.1371/journal.ppat.1007004>.
- Takumi K, Kinouchi T, Kawata T. 1979. Isolation and partial characterization of exosporium from spores of a highly sporogenic mutant of *Clostridium botulinum* type A. *Microbiol Immunol* 23:443–454. <https://doi.org/10.1111/j.1348-0421.1979.tb00484.x>.
- Gould GW, Hitchins AD. 1963. Sensitization of bacterial spores to lysozyme and to hydrogen peroxide with agents which rupture disulphide bonds. *J Gen Microbiol* 33:413–423. <https://doi.org/10.1099/00221287-33-3-413>.

35. Aronson AI, Fitz-James PC. 1971. Reconstitution of bacterial spore coat layers in vitro. *J Bacteriol* 108:571–578. <https://doi.org/10.1128/JB.108.1.571-578.1971>.
36. Mackey BM, Morris JG. 1972. The exosporium of *Clostridium pasteurianum*. *J Gen Microbiol* 73:325–338. <https://doi.org/10.1099/00221287-73-2-325>.
37. Aronson AI, Fitz-James P. 1976. Structure and morphogenesis of the bacterial spore coat. *Bacteriol Rev* 40:360–402. <https://doi.org/10.1128/MMBR.40.2.360-402.1976>.
38. Lund B, Gee J, King N, Horne R, Harnden J. 1978. The structure of the exosporium of a pigmented *Clostridium*. *J Gen Microbiol* 105:165–174. <https://doi.org/10.1099/00221287-105-2-165>.
39. Ebersold HR, Lüthy P, Cordier JL, Müller M. 1981. A freeze-substitution and freeze-fracture study of bacterial spore structures. *J Ultrastruct Res* 76:71–81. [https://doi.org/10.1016/s0022-5320\(81\)80051-2](https://doi.org/10.1016/s0022-5320(81)80051-2).
40. Plomp M, Leighton TJ, Wheeler KE, Malkin AJ. 2005. Architecture and high-resolution structure of *Bacillus thuringiensis* and *Bacillus cereus* spore coat surfaces. *Langmuir* 21:7892–7898. <https://doi.org/10.1021/la050412r>.
41. Plomp M, McCaffery JM, Cheong I, Huang X, Bettgowda C, Kinzler KW, Zhou S, Vogelstein B, Malkin AJ. 2007. Spore coat architecture of *Clostridium novyi* NT spores. *J Bacteriol* 189:6457–6468. <https://doi.org/10.1128/JB.00757-07>.
42. Plomp M, Carroll AM, Setlow P, Malkin AJ. 2014. Architecture and assembly of the *Bacillus subtilis* spore coat. *PLoS One* 9:e108560. <https://doi.org/10.1371/journal.pone.0108560>.
43. Rabi R, Turnbull L, Whitchurch CB, Awad M, Lyras D. 2017. Structural characterization of *Clostridium sordellii* spores of diverse human, animal, and environmental origin and comparison to *Clostridium difficile* spores. *mSphere* 2:e00343-17. <https://doi.org/10.1128/mSphere.00343-17>.
44. Holt SC, Leadbetter ER. 1969. Comparative ultrastructure of selected aerobic spore-forming bacteria: a freeze-etching study. *Bacteriol Rev* 33:346–378. <https://doi.org/10.1128/MMBR.33.2.346-378.1969>.
45. Brunt J, Cross KL, Peck MW. 2015. Apertures in the *Clostridium sporogenes* spore coat and exosporium align to facilitate emergence of the vegetative cell. *Food Microbiol* 51:45–50. <https://doi.org/10.1016/j.fm.2015.04.013>.
46. Steichen CT, Kearney JF, Turnbough CL. 2007. Non-uniform assembly of the *Bacillus anthracis* exosporium and a bottle cap model for spore germination and outgrowth. *Mol Microbiol* 64:359–367. <https://doi.org/10.1111/j.1365-2958.2007.05658.x>.
47. Thompson BM, Hoelscher BC, Driks A, Stewart GC. 2012. Assembly of the BclB glycoprotein into the exosporium and evidence for its role in the formation of the exosporium “cap” structure in *Bacillus anthracis*. *Mol Microbiol* 86:1073–1084. <https://doi.org/10.1111/mmi.12042>.
48. Smith CJ, Markowitz SM, Macrina FL. 1981. Transferable tetracycline resistance in *Clostridium difficile*. *Antimicrob Agents Chemother* 19:997–1003. <https://doi.org/10.1128/aac.19.6.997>.
49. Perutka J, Wang W, Goerlitz D, Lambowitz AM. 2004. Use of computer-designed group II introns to disrupt *Escherichia coli* DExH/D-box protein and DNA helicase genes. *J Mol Biol* 336:421–439. <https://doi.org/10.1016/j.jmb.2003.12.009>.
50. Heap JT, Theys J, Ehsaan M, Kubiak AM, Dubois L, Paesmans K, Van Mellaert L, Knox R, Kuehne SA, Lambin P, Minton NP. 2014. Spores of *Clostridium* engineered for clinical efficacy and safety cause regression and cure of tumors in vivo. *Oncotarget* 5:1761–1769. <https://doi.org/10.18632/oncotarget.1761>.
51. Crowther RA, Henderson R, Smith JM. 1996. MRC image processing programs. *J Struct Biol* 116:9–16. <https://doi.org/10.1006/jsbi.1996.0003>.
52. Shaw PJ, Hills GJ. 1981. Tilted specimen in the electron-microscope: a simple specimen holder and the calculation of tilt angles for crystalline specimens. *Micron* 12:279–282. [https://doi.org/10.1016/0047-7206\(81\)90081-9](https://doi.org/10.1016/0047-7206(81)90081-9).
53. Agard DA. 1983. A least-squares method for determining structure factors in three-dimensional tilted-view reconstructions. *J Mol Biol* 167:849–852. [https://doi.org/10.1016/s0022-2836\(83\)80114-4](https://doi.org/10.1016/s0022-2836(83)80114-4).
54. Amos LA, Henderson R, Unwin PN. 1982. Three-dimensional structure determination by electron microscopy of two-dimensional crystals. *Prog Biophys Mol Biol* 39:183–231. [https://doi.org/10.1016/0079-6107\(83\)90017-2](https://doi.org/10.1016/0079-6107(83)90017-2).
55. Collaborative Computational Project, Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760–763. <https://doi.org/10.1107/S0907444994003112>.
56. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612. <https://doi.org/10.1002/jcc.20084>.
57. Gipson B, Zeng X, Zhang ZY, Stahlberg H. 2007. 2dx—user-friendly image processing for 2D crystals. *J Struct Biol* 157:64–72. <https://doi.org/10.1016/j.jsb.2006.07.020>.
58. Havelka WA, Henderson R, Oesterheld D. 1995. Three-dimensional structure of halorhodopsin at 7 Å resolution. *J Mol Biol* 247:726–738. <https://doi.org/10.1006/jmbi.1995.0176>.
59. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
60. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
61. Müller T, Vingron M. 2000. Modeling amino acid replacement. *J Comput Biol* 7:761–776. <https://doi.org/10.1089/10665270050514918>.
62. Delcher AL, Salzberg SL, Phillippy AM. 2003. Using MUMmer to identify similar regions in large sequence sets. *Curr Protoc Bioinformatics* 00:10.3.1–10.3.18. <https://doi.org/10.1002/0471250953.bi1003s00>.
63. Chaudhuri RR, Loman NJ, Snyder LA, Bailey CM, Stekel DJ, Pallen MJ. 2008. xBASE2: a comprehensive resource for comparative bacterial genomics. *Nucleic Acids Res* 36:D543–D546. <https://doi.org/10.1093/nar/gkm928>.