nature microbiology

LETTERS

PUBLISHED: 18 APRIL 2017 | VOLUME: 2 | ARTICLE NUMBER: 17059

Structure of the hexagonal surface layer on *Caulobacter crescentus* cells

Tanmay A. M. Bharat[1,2*], Danguole Kureisaite-Ciziene[1], Gail G. Hardy[3], Ellen W. Yu[1], Jessica M. Devant[1], Wim J. H. Hagen[4], Yves V. Brun[3], John A. G. Briggs[1,4] and Jan Löwe[1*]

Many prokaryotic cells are encapsulated by a surface layer (S-layer) consisting of repeating units of S-layer proteins. S-layer proteins are a diverse class of molecules found in Gram-positive and Gram-negative bacteria and most archaea[1–5]. S-layers protect cells from the outside, provide mechanical stability and also play roles in pathogenicity. *In situ* structural information about this highly abundant class of proteins is scarce, so atomic details of how S-layers are arranged on the surface of cells have remained elusive. Here, using purified *Caulobacter crescentus*' sole S-layer protein RsaA, we obtained a 2.7 Å X-ray structure that shows the hexameric S-layer lattice. We also solved a 7.4 Å structure of the S-layer through electron cryotomography and sub-tomogram averaging of cell stalks. The X-ray structure was docked unambiguously into the electron cryotomography map, resulting in a pseudo-atomic-level description of the *in vivo* S-layer, which agrees completely with the atomic X-ray lattice model. The cellular S-layer atomic structure shows that the S-layer is porous, with a largest gap dimension of 27 Å, and is stabilized by multiple Ca^{2+} ions bound near the interfaces. This study spans different spatial scales from atoms to cells by combining X-ray crystallography with electron cryotomography and sub-nanometre-resolution sub-tomogram averaging.

S-layer proteins (SLPs) assemble to form planar sheets called S-layers on the surface of cells, where they are anchored usually through non-covalent interactions with other surface molecules (such as lipopolysaccharide (LPS) in Gram-negative bacteria)[2,6]. S-layers act as the outermost permeability barriers, protecting prokaryotic cells from extracellular attack and providing mechanical support to membranes[7]. S-layers also play a role in the pathogenicity of some bacteria, including *Clostridium difficile* and *Bacillus anthracis*[2]. The regularity and symmetry of S-layers have fascinated cell biologists, microscopists and structural biologists for many decades[8–10]. Low-resolution electron microscopy (EM) studies have been conducted on bacterial and archaeal S-layers[11,12], and the X-ray structures of some SLPs have been elucidated[13–16]. Thus far however, no high-resolution structure of an intact S-layer has been resolved directly on a prokaryotic cell surface.

To address this problem, we turned to *Caulobacter crescentus*, a well-studied Gram-negative alphaproteobacterium with a characteristic ultrastructure and a complex life cycle[17]. An ~120-nm-thick extension of the cell envelope called the stalk emanates from one pole of the *C. crescentus* cell body[18]. Electron cryotomography (cryo-ET) analysis of *C. crescentus* CB15 cells showed that the cells are covered with an S-layer that is continuous between the cell body and the stalk (Fig. 1a and Supplementary Video 1). The density corresponding to the S-layer of *C. crescentus* CB15 is almost perfectly hexameric (Fig. 1a, inset) with an ~220 Å repeat distance seen in tomographic top views of the cell surface, confirming previous EM and tomography studies of the *C. crescentus* S-layer[19,20]. Tomographic side views showed that the S-layer density is located ~180 Å away from the outer membrane (Fig. 1b,c). Two discrete densities were observed in the S-layer, the outer, highly connected S-layer lattice, and the discrete inner domains located around the centres of the hexamers. Weak, fuzzy density could be seen between the outer membrane and the inner domain of the S-layer, presumably corresponding to LPS molecules in which the S-layer is most probably anchored[21].

We purified the sole component of the *C. crescentus* CB15N (NA1000) S-layer, the ~98 kDa RsaA protein (Fig. 2a), directly from cells using a previously described S-layer extraction procedure that requires low pH[22]. Large quantities of pure RsaA protein could be obtained from cells using this procedure (Supplementary Fig. 1a). We confirmed that the purified protein retained its characteristic polymerization function by reconstituting the S-layers in solution at physiological pH (Fig. 2b). Incubation with divalent alkaline earth cations Ca^{2+} or Sr^{2+} resulted in the formation of two-dimensional sheets showing the characteristic 220 Å hexagonal lattice (Supplementary Fig. 1b,c and Supplementary Video 2). The reconstituted RsaA sheets showed only short-range order and were not perfectly planar (Supplementary Video 2), indicating that monomers of RsaA in the two-dimensional lattice and lattice contacts possessed significant conformational flexibility. Although purified RsaA had a tendency to degrade in solution even at 4 °C, the addition of divalent cations such as Ca^{2+} or Sr^{2+} moderately protected the protein from cleavage (Supplementary Fig. 1d), presumably because monomeric RsaA polymerized into sheets.

To resolve the structure of RsaA, we screened a variety of crystallization conditions and obtained diffraction-quality co-crystals of RsaA with Ca^{2+} or Sr^{2+}. Using the obtained crystals we determined a 2.7 Å X-ray structure by holmium SAD (single-wavelength anomalous dispersion) coupled with six-fold NCS (non-crystallographic symmetry) averaging and phase extension (Fig. 2c,d and Supplementary Table 1). RsaA amino-acid residues 249–1026 were resolved in the resulting electron density map (Supplementary Video 3). Peptide mapping of protein obtained from the dissolved crystals indicated that N-terminal residues of RsaA had been cleaved during crystallization. At least amino-acid residues 256–1026 were still present in the crystallized protein, as determined by mass spectrometry after protease cleavage into peptides, a method that does not reveal the exact N and C termini of the analysed polypeptide (Supplementary Fig. 2a,c).

$RsaA_{249–1026}$ adopts a predominantly β-helical fold with about 30 right-handed turns (Fig. 2c,d), forming an overall L-shape (Fig. 2c).

[1]Structural Studies Division, MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK. [2]Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK. [3]Department of Biology, Indiana University, Bloomington, Indiana 47405, USA. [4]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, Heidelberg 69117, Germany. *e-mail: tbharat@mrc-lmb.cam.ac.uk; jyl@mrc-lmb.cam.ac.uk

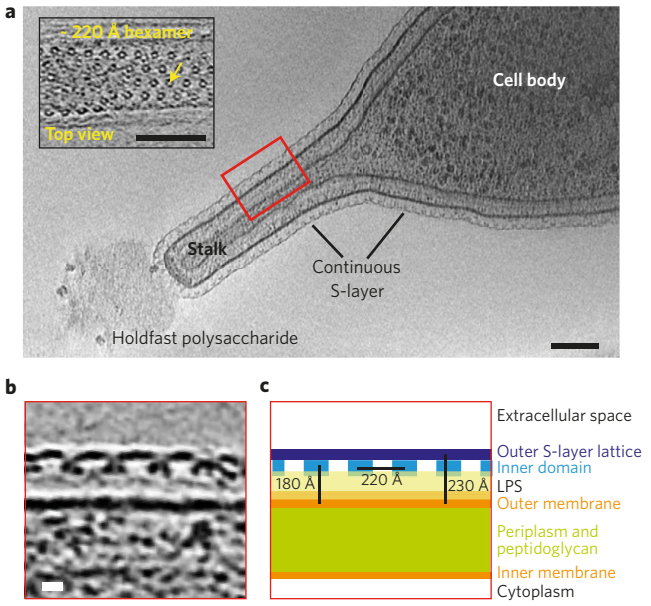

Figure 1 | Arrangement of the *C. crescentus* layer on cells and stalks. **a**, Tomographic slice of a CB15 *C. crescentus* cell. The S-layer is continuous between the cell body and the stalk. Inset: magnified tomographic slice through an S-layer of a stalk showing a hexameric lattice with ~220 Å spacing (Supplementary Video 1). Scale bar, 1,000 Å. **b**, Magnified tomographic slice of a side view of the cell surface. The S-layer is arranged in two layers and is seen ~180 Å away from the outer membrane of the cell. The outer S-layer lattice is highly interconnected, and the inner domains are ~220 Å apart from each other. Scale bar, 100 Å. **c**, Schematic representation of the cell surface at the same scale as panel **b**.

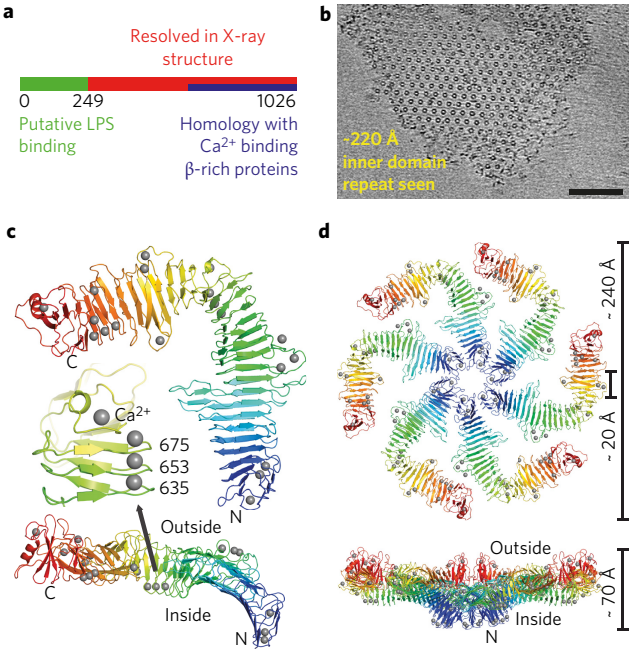

Figure 2 | The 2.7 Å X-ray structure of the outer S-layer lattice. **a**, Schematic of the RsaA protein sequence. The N-terminal region of RsaA is a putative LPS-binding region (Supplementary Fig. 2). Residues 249–1026 are resolved in the X-ray structure, and residues 618–1026 have a predicted sequence similarity with Ca^{2+}-dependent β-sheet rich proteins (HHPred server[51]). **b**, On addition of Ca^{2+} ions to RsaA, two-dimensional sheets were obtained with the same appearance and repeat as S-layers on cells. A cryo-ET slice through a reconstituted sheet is shown (Supplementary Video 2). Scale bar, 1,000 Å. **c**, Crystal structure of one monomer of RsaA (Supplementary Table 1), shown as a ribbon diagram coloured as a rainbow from the N terminus to the C terminus (blue to red). Two orthogonal orientations are shown, with the N and C termini of the protein marked. Ca^{2+} ions are shown as grey spheres. Inset: a region around amino acid 653, where three tightly bound structural Ca^{2+} ions stabilize the fold. **d**, Non-crystallographic hexamer of RsaA (which forms the asymmetric unit of the crystals), resolved in the X-ray structure and shown in two orthogonal orientations. The central pore is ~20 Å wide, and the lattice is ~70 Å thick. The N termini of the six monomers of RsaA point towards the cell surface.

The fold is reminiscent of β-helical proteins that protect cells against freezing, some of which also bind Ca^{2+} (ref. 23). A domain formed by the C-terminal residues 933–1026 is not β-helical and also contains the only α-helix in the entire structure. In the crystals, the L-shaped monomers are arranged into hexamers with a central pore with dimensions of ~20 Å (Fig. 2d). The N termini of all six RsaA monomers point towards the central pore, and the C termini are located on the outside of the hexamer, as dictated by the linear β-helical fold. Although the space group of the crystal is $P2_1$, the unit cell is almost hexameric (Supplementary Table 1), with an inter-hexameric spacing of between ~215 and ~220 Å. Intriguingly, this pseudo-hexameric repeat is close to the repeat observed in cellular S-layers in tomograms (Fig. 1). We thus hypothesized that the structure of $RsaA_{249–1026}$ corresponds to the outer S-layer lattice (Fig. 1b), and that the N-terminal residues missing in the crystal structure contain the membrane-proximal LPS-associated domain. We tested whether removal of the N terminus would remove LPS association to RsaA. In protein gels of full-length RsaA stained for LPS, a significant band was observed corresponding to LPS, which was not seen when Coomassie-stained (Supplementary Fig. 2b). We compared full-length RsaA with a gently trypsin-cleaved version missing the first 184 N-terminal residues, and found that, on removal of the N-terminal domain, the LPS band was also lost (Supplementary Fig. 2b). This indicates that, indeed, the X-ray structure contains only the outer S-layer lattice and not the inner LPS-associated domain of the S-layer.

All residues in the $RsaA_{249–1026}$ X-ray structure are well ordered, reflecting a requirement that the S-layer lattice must avoid flexible loops, as they would be easy targets for hostile extracellular proteases. The fully ordered C terminus of RsaA up to residue 1026 is almost completely buried in the C-terminal domain that deviates significantly from a canonical β-helical fold (Fig. 2c, residues 933–1026). Amino-acid side chains buried inside the β-helix domain of RsaA are either hydrophobic or are negatively charged and coordinated tightly to structural Ca^{2+} ions (Fig. 2c,d, grey spheres), with very few exceptions. Some Ca^{2+} ions are in close proximity to the central pore of the hexamer (Fig. 2d and Supplementary Video 4), where they appear to stabilize the hexameric interface. Inspection of the crystal packing revealed that RsaA hexamers were packed into regular, near-hexagonal, two-dimensional lattices that show repeat distances closely related to S-layers observed on cells and those seen reconstituted *in vitro* as described above.

To verify that the lattice found in the crystals is the same as the *in vivo* S-layer structure, we attempted to resolve the structure of the *C. crescentus* S-layer using cryo-ET of cells and subsequent sub-tomogram averaging. Because the quality of tomograms is affected by the thickness of the sample, we used a mutant of *C. crescentus*, strain YB2811[24], which efficiently sheds cell stalks for cryo-ET imaging (Supplementary Fig. 3 and Supplementary Video 5). Using a regularized likelihood algorithm for refinement[25], we obtained an *ab initio* sub-tomogram averaging model of the S-layer at ~30 Å resolution.

Next, to resolve a higher-resolution structure of the hexameric repeating asymmetric unit, we used a pipeline for data collection and image processing that has recently allowed near-atomic-resolution

sub-tomogram averaging refinement[26]. The initial model obtained above was used as input for high-resolution sub-tomogram refinement. The final refined 7.4 Å structure (Supplementary Fig. 4a) allowed us to locate different domains of the S-layer in the map (Supplementary Fig. 4b,c) and to identify the secondary structure in the S-layer (Fig. 3a–f and Supplementary Fig. 4d–f). As expected from inspection of the tomograms (Fig. 1b), the outer S-layer lattice is highly interconnected and consists mainly of β-sheets (Fig. 3a,b), whereas the inner domain is discrete and appears to contain cylindrical densities that are most probably α-helical (Fig. 3e,f).

We next docked the hexamer from the $RsaA_{249–1026}$ X-ray structure into the cryo-ET density map of the outer S-layer lattice as a rigid body (Fig. 3a–d, Supplementary Fig. 5a–f and Supplementary Video 6). The fit shows the precise position of the outer S-layer lattice in the cryo-ET map and shows that the inner domains are separated from the outer lattice by a short linker region between the two domains (Fig. 3c and Supplementary Fig. 4e). The size of the resolved inner domain density suggests that it probably contains all 248 N-terminal residues of RsaA missing from the X-ray structure. Extended densities emanating from the inner domains towards the outer membrane of the cell could be seen at low threshold levels (Supplementary Fig. 5c,d). Importantly, the docked atomic model very convincingly shows that the lattices formed by RsaA in the crystals and on cells are

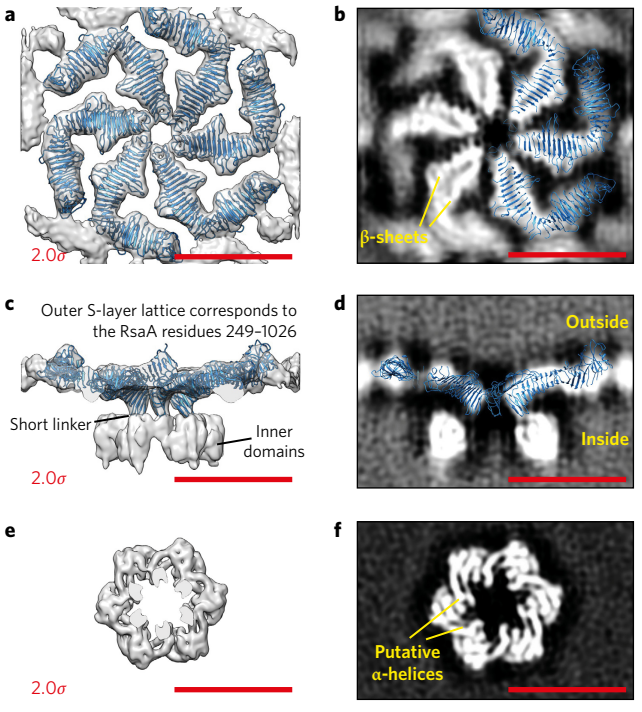

Figure 3 | A 7.4 Å cryo-ET and sub-tomogram averaging map of the *C. crescentus* S-layer. a, Top view (from outside the cell) of the outer layer from the sub-tomogram averaging map (grey density) with the X-ray structure (blue ribbon) fitted into the density (Supplementary Fig. 4). The cryo-EM density is shown at an isosurface threshold level of 2.0σ away from the mean (red number, bottom left). The entire hexamer as determined by X-ray crystallography is fitted into the outer S-layer lattice as a rigid body without further adjustments, demonstrating very strong agreement between the X-ray structure determination and cryo-ET. **b**, Slice through the outer S-layer lattice with three of the six RsaA molecules from the X-ray hexamer overlaid (Supplementary Figs 4 and 5 and Supplementary Video 6). **c**, Side view of the sub-tomogram averaging map, showing that N-terminal RsaA residue 249 in the X-ray structure is located in the short linker between the two domains of the S-layer (Supplementary Fig. 4e). **d**, Single slice through the cryo-ET map, with the fitted RsaA hexamer from the X-ray model overlaid (a clipping plane has been applied at the front and back). **e**, Top view of the hexamer of the inner domain of the S-layer. **f**, Slice through the inner domain density, which suggests that it might contain some α-helical secondary structure elements. Scale bars, 100 Å.

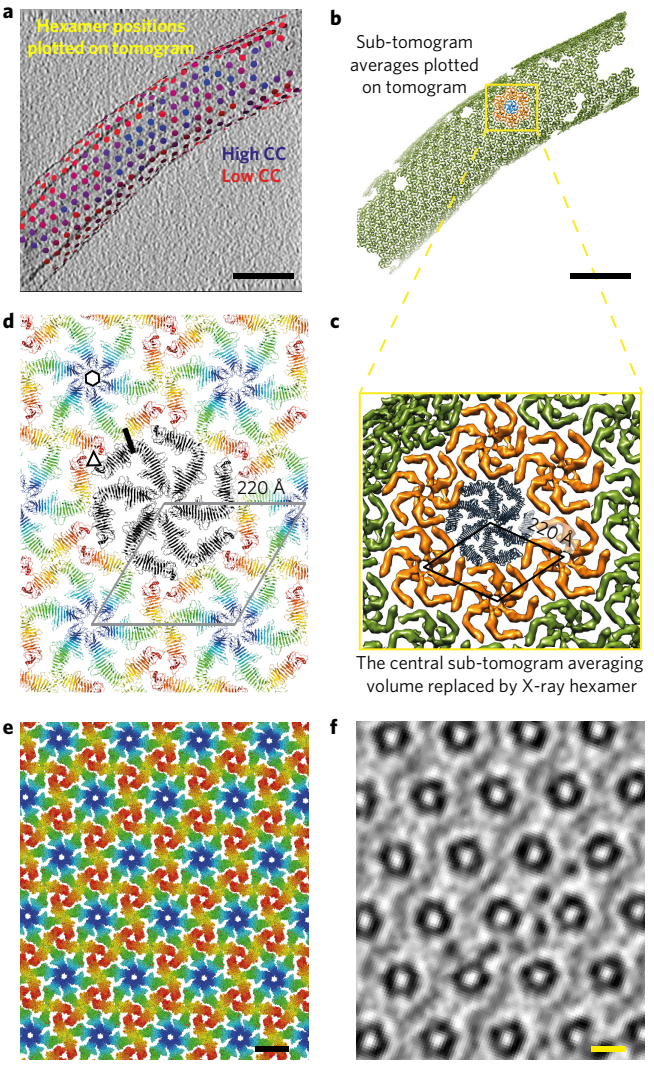

Figure 4 | S-layer lattices observed in the X-ray structure and the cryo-ET map are exceptionally similar. a, Final refined positions of sub-tomograms are shown plotted back onto a tomogram of a cell stalk with the corresponding refined orientations (Supplementary Video 7). Positions are coloured from blue (high cross-correlation (CC) of alignment) to red (low CC). One slice of the tomogram is shown, with protein density in black. **b**, The same plot as in **a**, except that each hexamer position is illustrated with the sub-tomogram average (green volumes). One hexamer is highlighted in blue and the hexamers directly contacting it are shown in orange. Scale bars in **a** and **b**, 1,000 Å. **c**, Zoomed view of the hexameric lattice revealed by cryo-ET and sub-tomogram averaging. The 220 Å hexamer–hexamer spacing is highlighted. The central blue hexamer from **b** is replaced by one copy of the X-ray hexamer. **d**, The RsaA lattice formed in the crystals through crystal packing is the same as the S-layer lattice observed on cells through cryo-ET. The 220 Å hexamer–hexamer repeat is indicated. Examples (one each) of hexameric, trimeric and dimeric interfaces are indicated by a black hexagon, triangle and line, respectively. In conclusion, both X-ray crystallography and cryo-ET used in this study show essentially the same lattice and arrangement of RsaA in *C. crescentus* S-layers. See also Supplementary Videos 4 and 7. **e**, Atomic structure of the S-layer. **f**, Slice through an *in vitro* assembled RsaA sheet (Supplementary Video 2) for comparison, at the same scale as in **e**. Scale bars in **e** and **f**, 100 Å.

extremely similar (Fig. 4a–d), allowing us to describe the arrangement of the S-layer lattice on cells at the atomic level (Fig. 4a–d and Supplementary Video 7).

In the two-dimensional lattice formed by RsaA in the crystals, shown above by cryo-ET to be equivalent to the lattice formed in intact S-layers at 7.4 Å resolution, the hexameric interface is mediated mainly by amino-acid residues between positions 256 and 260, making contact with the loop between residues 271 and 275 in the neighbouring subunit (Supplementary Fig. 6a,b, Supplementary Video 4 and Supplementary Table 2). Conspicuously, Ca^{2+} ions are bound near this interface (Supplementary Fig. 6b) and could play a role in stabilizing the RsaA hexamer and the S-layer lattice. The hexameric interface is the most extensive protein–protein interface in the S-layer lattice (Supplementary Table 2). Formation of this interface requires six RsaA molecules to assemble into a hexameric structure, and it is probably highly energetically favourable due to the tight packing of the monomers around the six-fold axis.

In addition to the hexameric interface, there are connections between residues 667–668 and 688–693 on one RsaA molecule with residues 713–715 and 756–758 on another RsaA molecule, in a dimeric interface between adjoining hexamers (Supplementary Fig. 6c,d and Supplementary Table 2). This dimeric interface is built up of two spatially distinct smaller interfaces (brown circles in Supplementary Fig. 6d), thus making it another strong interface stabilizing the S-layer lattice (Supplementary Fig. 6d). Again, like the hexameric interface described above, Ca^{2+} ions appear to be tightly bound to some interface residues. A less extensive trimeric contact is also observed in the S-layer lattice involving residues in the loop between positions 952–956 and residues 854–857 (Supplementary Fig. 6e,f and Supplementary Video 4). This interface is smaller than the hexameric and dimeric interfaces, and Ca^{2+} ions are located near some of the interface residues (Supplementary Fig. 6e,f).

A fascinating overall feature of the *C. crescentus* S-layer lattice is the presence of multiple Ca^{2+} ions bound tightly to RsaA, either tightly coordinated within the RsaA fold or bound near the hexameric, trimeric and dimeric interfaces (Supplementary Fig. 6). Without a divalent alkaline earth ion like Ca^{2+} or Sr^{2+}, no sheet formation was observed in our cryo-EM experiments, nor were crystals of RsaA sheets obtained, indicating that Ca^{2+} ions could play a role in lattice stabilization and assembly on cells. A strong Ca^{2+}-dependence of RsaA polymerization and S-layer lattice formation has been observed previously[22]. Our results are also in line with previous analysis of a *Geobacillus stearothermophilus* SLP SbsB (ref. 13), which also showed a strong dependence on Ca^{2+} for oligomerization.

The atomic structure of the S-layer lattice (Fig. 4e,f and Supplementary Fig. 7) suggests how the surface layer acts as a tough, yet flexible, permeability barrier, protecting cells from extracellular attack. The central circular pore is only ~20 Å wide (Fig. 2d), and other gaps between RsaA monomers in the lattice are no larger than ~27 Å along the largest dimensions, smaller than previous estimates of ~35 Å from low-resolution EM studies of purified RsaA sheets[20]. This pore size is larger than that observed in purified S-layer sheets of the archaeal species *Methanosarcina acetivorans*[14]. A 20-Å-wide pore would not only occlude harmful entities like phages, but would also prevent most macromolecules and even thin filaments (those larger than ~27 Å) from reaching the bacterial cell surface through the S-layer. Despite forming a continuous sheet, there is a requirement for flexibility in the S-layer lattice. RsaA monomers in the lattice must encase membranes with widely differing membrane curvatures, from 550-nm-wide *C. crescentus* cell bodies (Supplementary Fig. 7a,b) to 120-nm-wide cell stalks (Supplementary Fig. 3), requiring RsaA to adopt slightly different conformations. Furthermore, gaps in the lattice will be needed for insertion of new RsaA molecules on growing cells. The surface of RsaA is lined with rows of threonines, and this, together with similarities to certain anti-freeze proteins[23], might suggest that the *C. crescentus* S-layer plays an additional role in surface hydration and/or ice crystal avoidance.

The cryo-ET and X-ray structures together explain the modular architecture and function of RsaA, which was suggested by previous low-resolution cryo-ET studies[19]. Such a modular organization in p6 S-layers has been suggested previously for the bacterium *Acetogenium kivui*[12]. Having more than one domain can allow SLPs to perform the dual functions of S-layer lattice formation and membrane anchoring. This is an elegant solution to the problem of placing a rigid, lattice-forming protein in the correct position in a cell with only one part of the protein mediating anchoring (Supplementary Fig. 7a, inset).

The S-layer lattice structure on the surface of cells revealed in this study immediately suggests a number of experiments and potential applications. Residues exposed to the extracellular environment could be fluorescently tagged to investigate S-layer biogenesis and its morphological changes during the cellular life cycle of *C. crescentus* at the molecular and atomic levels. Also guided by the S-layer structure reported here, *C. crescentus* cells carrying mutated or fused versions of the *rsaA* gene may be used for surface display of molecules for nanotechnological or antigen display applications[27,28].

Recent improvements in cryo-ET imaging have revealed the architecture of protein complexes (for example, type IV pili) on bacterial cells[29], and novel image-processing techniques have allowed high-resolution refinement from purified virus (HIV-1) particles[26]. Here, we combine X-ray crystallography with cryo-ET imaging from cell stalks and sub-nanometre-resolution refinement to bridge different spatial scales from atoms to cells (Supplementary Fig. 7). The integrated approach (Supplementary Fig. 7) used in this study to obtain the atomic S-layer structure on the surface of bacterial cells illustrates convincingly how *in vitro* reconstitution may be combined with high-resolution *in vivo* imaging to shed light on an important biological problem. We believe that we are now entering an era of cellular structural biology where it will be increasingly possible to determine the structures of molecules *in situ* in their natural environments.

Methods

Purification of RsaA. *C. crescentus* CB15N (NA1000) cells were grown in peptone yeast extract (PYE) medium[30] for 48 h at 30 °C. Two litres of the resulting culture were centrifuged (4,000 r.c.f., 4 °C, 30 min). The pelleted cells were frozen in liquid nitrogen and stored at −80 °C. On the day of the purification, pellets were thawed and the cells were resuspended in 50 mM HEPES buffer at pH 2.0 and incubated at 37 °C for 10 min with vigorous shaking. The suspension was then centrifuged (16,000 r.c.f., 4 °C, 30 min). The pellet was discarded and the pH of the supernatant was adjusted to 7.0 with 5 M NaOH. The resulting liquid was filtered and from this point onwards kept at 4 °C. The protein solution was loaded onto a 5 ml HiTrap Q HP column (GE Healthcare), washed with 50 mM HEPES pH 7.0 and then eluted with the same buffer containing increasing concentrations of NaCl. Fractions containing pure RsaA were collected and concentrated to ~30 mg ml^{-1}. Aliquots of concentrated RsaA were frozen in liquid nitrogen and stored at −80 °C.

Crystallization of *C. crescentus* RsaA. The RsaA protein solution was supplemented with 5 mM $CaCl_2$ before crystallization. Initial screens of full-length RsaA were set up using the in-house robotic nanolitre crystallization facility of the MRC Laboratory of Molecular Biology[31]. After optimization, the native RsaA crystals were grown at 19 °C by sitting-drop vapour diffusion in a drop composed of 100 nl of reservoir solution (0.07 M KSCN, 24% (wt/vol) PEG 8000, 0.075 M TAPS pH 8.5) and 100 nl of protein solution at 30 mg ml^{-1}. Plate-like crystals appeared in 3–10 days and continued growing to a final size of 300 × 300 × 30 μm^3. Crystals were flash-frozen in liquid nitrogen for data collection using an additional 25% (vol/vol) PEG 200 as cryo-protectant. Heavy-atom crystal derivatives, selected based on their similarity to calcium, were obtained using a similar technique. Again, 100 nl of the reservoir solution (0.43 M KSCN, 18.5% (wt/vol) PEG 8000, 0.02 M Sr $(CH_3COO)_2$, 0.075 M TAPS pH 8.5) was mixed with 100 nl of 26 mg ml^{-1} RsaA solution supplemented with 5 mM $SrCl_2$. The resulting crystals were soaked for 7 min in reservoir solution additionally containing 0.04 M $HoCl_3$, 25% (vol/vol) PEG 200, then flash-frozen in liquid nitrogen. Holmium was used for phasing, as described below in the section describing RsaA structure determination by X-ray crystallography.

Trypsin cleavage and LPS staining. Full-length RsaA (~2 mg ml^{-1}) was incubated at a ratio of 1,000:1 with trypsin at room temperature (22 °C) for 24 h with gentle shaking. The pH of the resulting solution was lowered to 2.0 with HCl to inactivate trypsin and then restored to pH 7.0 with NaOH. The solution was centrifuged, filtered and then subjected to gel fitration chromatography on a Superose 6 10/300 column (GE Healthcare) using 100 mM NaCl, 1 mM EGTA, 50 mM HEPES pH 7.0 buffer. Fractions containing cleaved RsaA were pooled and concentrated. Staining of full-length and trypsin-cleaved RsaA was performed using the Pro-Q Emerald 300 lipopolysaccharide gel stain kit (Thermo Fisher Scientific).

RsaA structure determination by X-ray crystallography. The structure was solved by holmium SAD, phase transfer to a high-resolution native data set by molecular replacement, and by phase extension using six-fold, non-crystallographic (NCS) symmetry. A SAD data set obtained from a holmium-soaked RsaA crystal was collected on beamline I04 at the Diamond Light Source (Harwell, UK) at the holmium edge to 3.3 Å resolution (Supplementary Table 1). Diffraction images were indexed and integrated with XDS (ref. 32) and scaled and merged with SCALA (ref. 33). Forty-two holmium sites were located with SHELXD (ref. 34). PHASER (ref. 35) extended this to 49 sites, and an interpretable electron density map was obtained in one of the two possible hands after SAD phasing in PHASER, coupled with density modification in DM (ref. 36). BUCCANEER (ref. 37) was used to build a rough initial model that indicated six-fold NCS symmetry, but was not sufficient to complete the atomic interpretation. Because the crystals were not isomorphous to each other, the density around one of the six monomers was cut out using CCP4 methods and used for molecular replacement using PHASER into the native data set, which extended to 2.7 Å resolution in two directions (Supplementary Table 1). The native data set was collected on beamline I04 at the Diamond Light Source and was first modified for severe anisotropy in the B* direction using the UCLA diffraction anisotropy server (https://services.mbi.ucla.edu/anisoscale/). Phases for the native data set were slowly extended from 8 to 2.7 Å using DM and the six-fold NCS symmetry over 2,000 cycles. The resulting electron density map was readily interpretable and all six monomers of RsaA were traced almost completely and automatically with BUCCANEER. The model was improved manually using MAIN (ref. 38) and refined with PHENIX and REFMAC (refs 39, 40), keeping the 5% randomly selected test set of reflections separate at all times. Nineteen calcium ions per RsaA chain were added in positions of otherwise unexplained strong electron density peaks located in chemically plausible environments, such as being close to several carbonyl groups and aspartate and glutamate side chains. Structure factors and coordinates of the final model have been deposited in the Protein Data Bank (PDB) with accession code 5N8P.

Purification of *C. crescentus* stalks. Purification of stalks from cells was performed using a stalk-shedding *C. crescentus* strain (YB2811), with a differential centrifugation procedure described previously, but slightly modified here[24]. Briefly, 5 ml of starter culture of strain YB2811 in PYE medium was used to inoculate a 250 ml flask containing PYE medium. The new culture was grown at 30 °C for 2–3 days. The culture was imaged using a wide-field light microscope to confirm efficient stalk shedding. The culture was then centrifuged to pellet whole *C. crescentus* cells (17,000 rcf, 4 °C, 30 min). The pellet was discarded and the supernatant was centrifuged again to pellet stalks (41,000 rcf, 4 °C, 1 h). The supernatant was carefully decanted and the pellet was resuspended in 100 µl PYE and stored at 4 °C or used immediately for cryo-EM sample preparation.

Cryo-EM sample preparation. Whole cells, purified stalks or reconstituted sheets were mixed with 10 nm protein-A gold (CMC Utrecht) and pipetted onto a freshly glow-discharged Quantifoil R2/2 Cu/Rh 200 mesh grids in a Vitrobot Mark IV (FEI Company) maintained at 10 °C at 100% relative humidity. Grids were plunge-frozen into liquid ethane and stored in liquid nitrogen until the day of microscopy.

Cryo-EM and cryo-ET data collection. Initial cryo-EM assessment of samples was conducted on a Spirit microscope (FEI Company) with a Gatan 626 side-entry holder. Data collection for initial sub-tomogram averaging model generation was performed on a Titan Krios microscope (FEI) fitted with a Quantum energy filter (slit width 20 eV) and a K2 direct electron detector running in counting mode with a dose rate of ~10 e$^-$ per pixel per second (Gatan). Tilt series (22 in total) of purified stalks were collected between +60° and −60° in two directions with a 2° increment using SerialEM (ref. 41). A total dose of 70 e$^-$ Å^{-2} was applied, and the data were sampled at a pixel size of 4.45 Å. To obtain a sub-nanometre-resolution cryo-ET structure of the S-layer on cell stalks, we adopted a data collection strategy that has recently allowed high-resolution refinement from cryo-ET[26]. Briefly, a dose-symmetric data collection scheme was used[42] and 110 tomograms were collected at a pixel size of 1.3 Å. A total dose of 140 e$^-$ Å^{-2} was applied over the tilt series collected between +60° and −60°, with 3° spacing between tilts.

Sub-tomogram averaging image processing. Tomogram reconstructions were performed in IMOD (ref. 43). Sub-tomogram extraction from tomograms of tubular stalks was performed using previously described procedures implemented in MATLAB (ref. 44). Initial model generation was performed using a regularized likelihood refinement algorithm implemented in RELION software[45]. First, heavily binned sub-tomograms were subjected to RELION 3D classification (with a three-dimensional contrast transfer function (CTF) model used) with wide angular and translational search range to locate all the hexamers of the S-layer in three dimensions. Once all the hexamers had been located, sub-tomograms were re-extracted at the refined positions, and a three-dimensional auto-refinement sub-tomogram averaging was run in RELION. The final volume obtained had a resolution of 30 Å, and this was used as input for high-resolution refinement. High-resolution sub-tomogram averaging was performed using a recently described image-processing scheme[26,42]. CTF estimation of the input tilt series images was performed using CTFFIND (ref. 46). The estimated defocus values were used for CTF correction within IMOD to produce a CTF corrected tilt series stack. Next, dose filtration was performed on the stacks to down-weight the higher-resolution Fourier components in the high-tilt images[26]. The resulting dose-weighted and CTF-corrected stacks were used for tomogram generation in IMOD using weighted back-projection. The centres of all stalks were picked manually, and these central positions were used to generate the sub-tomogram extraction locations assuming a cylindrical arrangement of the RsaA hexamers around the stalk. Sub-tomograms were extracted from eight times binned tomograms, as before, and sub-tomogram alignment and averaging were performed to locate the positions of each hexamer in the tomogram[47,48]. After locating the hexamer positions, all poorly aligned sub-tomograms were removed, as judged from cross-correlation scores, as well from their position in lattice maps. Finally, re-extraction was performed with no binning on the data and with the extracted sub-tomograms centred on the six-fold axis of each hexamer. C6 symmetry was applied during the final refinement of a total of 51,866 hexamers of S-layer from 110 tomograms, leading to a 7.4 Å structure. Resolution estimation, Fourier shell correlation (FSC) weighting and map sharpening were conducted in RELION (ref. 49). The final map has not been explicitly symmetrized six-fold and has been deposited in the Electron Microscopy Data Bank (EMDB) with accession code EMD-3604. Rigid body docking of the hexamer of RsaA from X-ray crystallography into the cryo-ET map was performed with UCSF Chimera[50].

Data availability. Structure factors and the coordinates of the final model have been deposited in the Protein Data Bank (PDB) under accession code 5N8P. The final cryo-ET map has been deposited in the Electron Microscopy Data Bank (EMDB) under accession code EMD-3604. The data that support the findings of this study are available from the corresponding authors upon request.

Received 24 January 2017; accepted 24 March 2017; published 18 April 2017

References

1. Albers, S. V. & Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9,** 414–426 (2011).
2. Fagan, R. P. & Fairweather, N. F. Biogenesis and functions of bacterial S-layers. *Nat. Rev. Microbiol.* **12,** 211–222 (2014).
3. Glauert, A. M. The fine structure of bacteria. *Br. Med. Bull.* **18,** 245–250 (1962).
4. Sara, M. & Sleytr, U. B. S-layer proteins. *J. Bacteriol.* **182,** 859–868 (2000).
5. Zhu, C. *et al.* Diversity in S-layers. *Prog. Biophys. Mol. Biol.* **123,** 1–15 (2017).
6. Kirk, J. A., Banerji, O. & Fagan, R. P. Characteristics of the *Clostridium difficile* cell envelope and its importance in therapeutics. *Microb. Biotechnol.* **10,** 76–90 (2016).
7. Sleytr, U. B. & Beveridge, T. J. Bacterial S-layers. *Trends Microbiol.* **7,** 253–260 (1999).
8. Houwink, A. L. A macromolecular mono-layer in the cell wall of *Spirillum* spec. *Biochim. Biophys. Acta* **10,** 360–366 (1953).
9. Sleytr, U. B. & Glauert, A. M. Ultrastructure of the cell walls of two closely related *Clostridia* that possess different regular arrays of surface subunits. *J. Bacteriol.* **126,** 869–882 (1976).
10. Baumeister, W., Wildhaber, I. & Phipps, B. M. Principles of organization in eubacterial and archaebacterial surface proteins. *Can. J. Microbiol.* **35,** 215–227 (1989).
11. Kessel, M., Wildhaber, I., Cohen, S. & Baumeister, W. Three-dimensional structure of the regular surface glycoprotein layer of *Halobacterium volcanii* from the Dead Sea. *EMBO J.* **7,** 1549–1554 (1988).
12. Lupas, A. *et al.* Domain structure of the *Acetogenium kivui* surface layer revealed by electron crystallography and sequence analysis. *J. Bacteriol.* **176,** 1224–1233 (1994).
13. Baranova, E. *et al.* SbsB structure and lattice reconstruction unveil Ca^{2+} triggered S-layer assembly. *Nature* **487,** 119–122 (2012).
14. Arbing, M. A. *et al.* Structure of the surface layer of the methanogenic archaean *Methanosarcina acetivorans*. *Proc. Natl Acad. Sci. USA* **109,** 11812–11817 (2012).
15. Jing, H. *et al.* Archaeal surface layer proteins contain beta propeller, PKD, and beta helix domains and are related to metazoan cell surface proteins. *Structure* **10,** 1453–1464 (2002).
16. Kern, J. *et al.* Structure of surface layer homology (SLH) domains from *Bacillus anthracis* surface array protein. *J. Biol. Chem.* **286,** 26042–26049 (2011).

17. Jiang, C., Brown, P. J., Ducret, A. & Brun, Y. V. Sequential evolution of bacterial morphology by co-option of a developmental regulator. *Nature* **506,** 489–493 (2014).
18. Wagner, J. K. & Brun, Y. V. Out on a limb: how the *Caulobacter* stalk can boost the study of bacterial cell shape. *Mol. Microbiol.* **64,** 28–33 (2007).
19. Amat, F. *et al.* Analysis of the intact surface layer of *Caulobacter crescentus* by cryo-electron tomography. *J. Bacteriol.* **192,** 5855–5865 (2010).
20. Smit, J., Engelhardt, H., Volker, S., Smith, S. H. & Baumeister, W. The S-layer of *Caulobacter crescentus*: three-dimensional image reconstruction and structure analysis by electron microscopy. *J. Bacteriol.* **174,** 6527–6538 (1992).
21. Ford, M. J., Nomellini, J. F. & Smit, J. S-layer anchoring and localization of an S-layer-associated protease in *Caulobacter crescentus*. *J. Bacteriol.* **189,** 2226–2237 (2007).
22. Nomellini, J. F., Kupcu, S., Sleytr, U. B. & Smit, J. Factors controlling *in vitro* recrystallization of the *Caulobacter crescentus* paracrystalline S-layer. *J. Bacteriol.* **179,** 6349–6354 (1997).
23. Garnham, C. P., Campbell, R. L. & Davies, P. L. Anchored clathrate waters bind antifreeze proteins to ice. *Proc. Natl Acad. Sci. USA* **108,** 7363–7367 (2011).
24. Ireland, M. M., Karty, J. A., Quardokus, E. M., Reilly, J. P. & Brun, Y. V. Proteomic analysis of the *Caulobacter crescentus* stalk indicates competence for nutrient uptake. *Mol. Microbiol.* **45,** 1029–1041 (2002).
25. Bharat, T. A. M. & Scheres, S. H. W. Resolving macromolecular structures from electron cryo-tomography data using subtomogram averaging in RELION. *Nat. Protoc.* **11,** 2054–2065 (2016).
26. Schur, F. K. M. *et al.* An atomic model of HIV-1 capsid-SP1 reveals structures regulating assembly and maturation. *Science* **353,** 506–508 (2016).
27. Howorka, S. Rationally engineering natural protein assemblies in nanobiotechnology. *Curr. Opin. Biotechnol.* **22,** 485–491 (2011).
28. Mark, S. S. *et al.* Bionanofabrication of metallic and semiconductor nanoparticle arrays using S-layer protein lattices with different lateral spacings and geometries. *Langmuir* **22,** 3763–3774 (2006).
29. Chang, Y.-W. *et al.* Architecture of the type IVa pilus machine. *Science* **351,** aad2001 (2016).
30. Poindexter, J. S. Biological properties and classification of the *Caulobacter* group. *Bacteriol. Rev.* **28,** 231–295 (1964).
31. Stock, D., Perisic, O. & Löwe, J. Robotic nanolitre protein crystallisation at the MRC Laboratory of Molecular Biology. *Prog. Biophys. Mol. Biol.* **88,** 311–327 (2005).
32. Kabsch, W. XDS. *Acta Crystallogr. D* **66,** 125–132 (2010).
33. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D* **67,** 282–292 (2011).
34. Sheldrick, G. M. in *Direct Methods for Solving Macromolecular Structures* (ed. Fortier, S.) 401–411 (Springer, 1998).
35. McCoy, A. J. Solving structures of protein complexes by molecular replacement with *Phaser*. *Acta Crystallogr. D* **63,** 32–41 (2007).
36. Cowtan, K. Recent developments in classical density modification. *Acta Crystallogr. D* **66,** 470–478 (2010).
37. Cowtan, K. The *Buccaneer* software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D* **62,** 1002–1011 (2006).
38. Turk, D. *MAIN* software for density averaging, model building, structure refinement and validation. *Acta Crystallogr. D* **69,** 1342–1357 (2013).
39. Adams, P. D. *et al. PHENIX*: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
40. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).
41. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152,** 36–51 (2005).
42. Hagen, W. J. H., Wan, W. & Briggs, J. A. G. Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *J. Struct. Biol.* **197,** 191–198 (2016).
43. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116,** 71–76 (1996).
44. Bharat, T. A. M. *et al.* Cryo-electron tomography of Marburg virus particles and their morphogenesis within infected cells. *PLoS Biol.* **9,** e1001196 (2011).
45. Bharat, T. A. M., Russo, C. J., Löwe, J., Passmore, L. A. & Scheres, S. H. W. Advances in single-particle electron cryomicroscopy structure determination applied to sub-tomogram averaging. *Structure* **23,** 1743–1753 (2015).
46. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142,** 334–347 (2003).
47. Briggs, J. A. *et al.* Structure and assembly of immature HIV. *Proc. Natl Acad. Sci. USA* **106,** 11090–11095 (2009).
48. Förster, F., Medalia, O., Zauberman, N., Baumeister, W. & Fass, D. Retrovirus envelope protein complex structure *in situ* studied by cryo-electron tomography. *Proc. Natl Acad. Sci. USA* **102,** 4729–4734 (2005).
49. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180,** 519–530 (2012).
50. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).
51. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33,** W244–W248 (2005).

Acknowledgements

The authors thank M. Skehel and F. Begum (MRC Laboratory of Molecular Biology, LMB) for mass-spectrometric identification of proteins, M. Yu (LMB) for help with X-ray data collection, C. Savva (LMB) for help with cryo-EM data collection, S. Scheres (LMB) for help with RELION software, F. Schur and W. Wan (European Molecular Biology Laboratory, EMBL) for providing high-resolution image-processing code and scripts before publication and for advice on their implementation, F. van den Ent (LMB) for advice on protein purification and T. Darling and J. Grimmett (LMB) for help with high-performance computing. The authors also thank Y. Modis (Cambridge University) for pointing out the similarity to anti-freeze proteins. Part of this work was funded by the European Molecular Biology Organization (aALTF 778-2015 to T.A.M.B.), the Medical Research Council (U105184326 to J.L.), the Wellcome Trust (095514/Z/11/Z to J.L.) and the National Institutes of Health (GM51986 to Y.V.B.). This work was supported by iNEXT, project no. 1482, funded by the Horizon 2020 programme of the European Union.

Author contributions

T.A.M.B., Y.V.B. and J.L. designed the research. T.A.M.B., D.K.-C., G.G.H., E.W.Y., J.M.D., W.J.H.H. and J.L. performed experiments. W.J.H.H. and J.A.G.B. supported high-resolution cryo-ET and image processing. TA.M.B., D.K.-C., J.M.D., E.W.Y. and J.L. analysed the data. T.A.M.B. and J.L. wrote the manuscript with support from all authors.

Additional information

Supplementary information is available for this paper.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to T.A.M.B. and J.L.

How to cite this article: Bharat, T. A. M. *et al.* Structure of the hexagonal surface layer on *Caulobacter crescentus* cells. *Nat. Microbiol.* **2**, 17059 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Competing interests

The authors declare no competing financial interests.