Many bacteria and archaebacteria possess a two-dimensional protein array, or S-layer, that covers the cell surface and plays crucial roles in cell physiology. Here, we report the crystal structure of SlpA, the main S-layer protein of the bacterial pathogen *Clostridioides difficile*, and use electron microscopy to study S-layer organisation and assembly. The SlpA crystal lattice mimics S-layer assembly in the cell, through tiling of triangular prisms above the cell wall, interlocked by distinct ridges facing the environment. Strikingly, the array is very compact, with pores of only ~10 Å in diameter, compared to other S-layers (30–100 Å). The surface-exposed flexible ridges are partially dispensable for overall structure and assembly, although a mutant lacking this region becomes susceptible to lysozyme, an important molecule in host defence. Thus, our work gives insights into S-layer organisation and provides a basis for development of *C. difficile*-specific therapeutics.
The surfaces of most bacteria and archaea are covered with a proteinaceous coat, the surface or S-layer, which is formed through the self-assembly of individual protein subunits into a regularly spaced, two-dimensional array. The tendency of S-layer proteins to spontaneously form 2D assemblies has hampered structure determination and restricted understanding of both their function and architecture. To date, all the S-layer proteins studied in detail have a two-domain organisation, where the assembly domain is responsible for the formation of the paracrystalline layer, and the anchoring domain allows attachment of the S-layer to the cell wall. Considerable variation is observed in the functioning of target cells causing rounding and ultimately death. Toxins that disrupt the cytoskeletal structure and the tight junctions of the S-layer largely consist of a major S-layer component forms an intricate complex of representatives of two SLCTs (SLCT17, strain CD630 and SLCT17b, strain R7404) were determined using X-ray crystallography; we combined single anomalous dispersion sulfur data (S-SAD) and molecular replacement using substructures of the interacting domains (SLP1, interacting domain, LID, and SLP1H, interacting domain, HID; PDB ID: 7ACY) and SLP1/HID (PDB ID: 7ACV). Final models to a resolution of 2.55 Å (PDB ID: 7ACY) and 2.65 Å (PDB ID: 7ACX) for SlpACD630 and SlpAR7404, respectively, reveal a nearly identical fold (RMSD over aligned Ca − core 2.1 Å) and hence allow our analysis of the best overall model, SLP1_CD630. Our H/L structural model reveals three distinct regions: a pseudo-threefold symmetric SLP1H tile, an intricate LID/HID interacting motif and a third region composed of two domains, D1 and D2, of SlpP, (Fig. 1b). These regions define two separate planes, with the SLP2 spanning ~35 Å above the SLP1H plane, linked by the LID/HID motif (Fig. 1b).

SLP1H is composed of three sequence-conserved cell wall binding motifs 2 (CBW2), which define the C. difficile cell wall protein (CWP) family, and the HID. The three CBW2 motifs form a triangular prism and adopt an intertwined fold, with a three-strand from one CBW2 inserting into the neighbouring domain to complete a three-stranded (β23 in CBW2, complete CBW22, strand β64 from CBW2 inserts into CBW2, and β104 in CBW23 is inserted into CBW23) sandwiched between two α-helical regions (Supplementary Fig. 1). At the core of the face sits a helical bundle with each individual CBW2 containing one α-helix (CBW2, α32, CBW21−α74, CBW21−α111) while two α-helices define the vertices of the pseudo-threefold arrangement (Supplementary Fig. 1). This trimeric arrangement generates a distinctive charge distribution across each prism face, creating surfaces with complementary charges (Fig. 1c). The pseudo-threefold organisation of the CBW2s is conserved in the structures of two minor constituents of the S-layer, Cwp6 and Cwp8 (Supplementary Fig. 2). The environment- and cell-facing sides of SLP3 exhibit considerable charge differences, with a mostly negatively charged external surface and a largely non-polar cell wall-facing base, decorated by positive patches (Fig. 1c). The positive patches at the cell-wall base could provide the mechanism for anchoring SlpA to the cell wall via interactions with the anionic secondary cell wall polymer PSII.

It is worth noting that, despite the lack of homology between S-layer proteins from different species, all anchoring domains determined to date have a predominantly α-helical structure, which suggests that helical bundles might be a favoured organisation for S-layer attachment to the cell. Notably, the S-layer homology domain (SLH) responsible for anchoring the S-layer in Bacillus anthracis has a central three α-helical bundle similar to the core of the CBW2 triad organisation in SLP1H.

The HID motif interlocks with LID in an arrangement reminiscent of a paperclip: LID α-helices α41, and α54 and HID α114 pack against a β-sheet formed by insertion of β114 from HID between LID β-strands, β183, and β211 (Fig. 2a and Supplementary Fig. 1). This novel structural motif locks SLP1 and SLP1H together, providing a structural basis for the stability of the H/L complex.

Results

C. difficile major S-layer component forms an intricate complex. The major protein of the C. difficile S-layer, SlpA, is post-translationally cleaved into two S-layer proteins (SLPs): the high molecular weight (HMW) and low molecular weight (LMW) α-helices, herein referred to as SLP1 and SLP1H, respectively. These subunits then form a complex (referred to as H/L) that is incorporated into the S-layer (Fig. 1a). Unlike in other S-layer structures determined to date, in C. difficile, the anchoring domain, responsible for linking the S-layer to the cell wall, and the assembly domain, which forms the 2D paracrystalline array, are one and the same, both proposed to be present in SLP1H.

The structures of the full-length H/L complexes of representatives of different bacteria makes these arrays attractive targets for species-specific therapeutic interventions, provided sufficient structural and functional data are available.

Here we present the complete atomic-level model of the assembly and anchoring domains of the S-layer in C. difficile, generated by combining high-resolution X-ray crystallography with electron microscopy.

C. difficile major S-layer component forms an intricate complex. The major protein of the C. difficile S-layer, SlpA, is post-
To identify residues essential for the interaction of the SLP₁ and SLP₁(H/L) subunits, we analyzed H/L complex formation in an enzyme-linked immunosorbent assay (ELISA) with a panel of individual point mutants (Fig. 2c and Supplementary Fig. 3). Mutations of a single amino acid within LID (F274A) or HID (Y27A) in SlpA CD630 were sufficient to destabilize the H/L complex (Fig. 2c). Moreover, expression of either point mutant in an SlpA-null background resulted in SLPL shedding from the cell surface of C. difficile and detection of a fraction of SLPH in the culture supernatant. Loss of SLP₁ also resulted in partial degradation of SLP₁ (Fig. 2d): N-terminal sequencing revealed truncation of the HID, indicating that this region is unstable in the absence of the LID/HID interaction (Fig. 2d).

SLP₁ protrudes from the interacting domains, with D1 closest to the SLP₁ plane and D2 extending outwards at an angle of ~120°, away from the long axis of D1. As we previously described, whilst D1 is well ordered, formed by a 5-strand β-sheet packed against two α-helices (Supplementary Fig. 1), D2 is predominantly composed of long, flexible loops, particularly at the externally exposed surface (Fig. 1b). Charge distribution across CWB2 motifs in SlpACD630 shown as a gradient (positive in blue to negative in red). Views are shown from the extracellular and cell wall surfaces, followed by side views of the lateral faces defined by two interacting CWB2s. Putty representations of SlpACD630 H/L complex showing B-factors ranging from low (blue and narrow) to high (red and wide). High B-factors are indicative of disorder/flexible regions. Conservation of the SlpA sequence across annotated SlpA cassette types (SLCTs) depicted on putty representations of SlpACD630 H/L complex, coloured from conserved (purple) to variable (cyan). Conservation was calculated using Consurf web server.

Available structures of assembly domains of other S-layers show a predominance of β-strand rich domains, organised to create the 2D array. The S-layer assembly domains from two Bacilli rely on multiple structurally related β-sandwich protomers. These form a central tile-like body, with an arm, composed of the more external protomers, extending outwards and providing both extra lattice contacts as well as a certain...
In *C. difficile*, the β-strand rich D1 and D2 domains also extend outwards from the SLPH tiles and seem to adopt a similar conformation; even though they are not thought to be part of the core assembly domain, they could thus also contribute to S-layer assembly contacts.

Conformational flexibility in the organisation of SlpA is further demonstrated by different arrangements observed in the structure of a truncated derivative of SlpA (SLP L/HID) and the H/L complex. In the structure of the SLPL/HID complex (PDB ID: 7ACV), the D1-D2 domains exhibit an orientation relative to the interacting domains, which is different from that seen in the corresponding H/L complex (R7404 strain, SLCT-7b, PDB ID: 7ACX). Our models indicate the presence of a hinge, formed by the D1-LID linker (Fig. 2b). The calculated rotation angle of the centres of mass of the D1-D2 region relative to LID/HID motif of 

Fig. 2 Interactions and flexibility in *C. difficile* SLPl/HID (H/L) complex. a Paperclip organisation of the interacting domains LID/HID is maintained by a range of interactions, with selected interface residues identified in strain R7404 (SLCT-7b, PDB ID: 7ACW) depicted as sticks. 2mFo-DFc electron density map is shown on the interacting amino acid pairs as a grey mesh contoured at 1.5 σ. Specific interatomic interactions identified with PDBePISA are represented as a dashed line. b Superimposing structures of SLPl/HID (gold/ slate blue, PDB ID: 7ACY) onto the native complex of SlpA (SLCT-7b, PDB ID: 7ACX) (blue/white) reveals the flexibility of the LID-D1 linker, as illustrated by rotation of D1-D2 domains in relation to fixed position of LID/HID motif (left). The hinge loop enabling this conformational flexibility (determined by DynDom6D) is coloured in red. The backbone displacement (coloured from blue – low, to red – high Ca deviation) is shown on the alignment of D1-D2 region of both structures (middle; SLPl/HID - opaque, H/L - semi-transparent) with the rotation angle of the LID/HID motif indicated with an arrow. Structural dynamics (right) of SLPl/HID, represented as increasing mobility (coloured blue – rigid, to red – mobile), calculated based on elastic network models implemented in DynOomics ENM version 1.0 server. c Probing of CD630 H/L complex interactions in vitro with ELISA, comparing effects of intact SLPl (gold circles), SLPl (slate blue circles), variants lacking interacting domains (black squares) and substitution mutants of F274A (structurally equivalent to F270 in R7404 LID/HID depicted in c, dark green triangles) and Y27A (structurally equivalent to Y26 in R7404 LID/HID in c, light green triangles) on H/L complex formation. Graphs represent mean ± standard deviation (SD) of *n* = 3 experiments, with least-squares curve fit of product formed upon the interaction of the two subunits. Source data provided in Source data file. d Western blot of cell surface extracts and culture supernatants, detecting (black arrowhead) SLPl (left) and SLPl (right) in strains devoid of endogenous slpA and expressing plasmid-borne SlpA (SLP L/HID) and SLPl variants with either F274A or Y27A substitution mutants in SLPl or SLH (denoted in subscript), respectively (*n* = 1). Detected product of partial degradation of SLPl indicated with an asterisk. Source data provided in Source Data file.
**Fig. 3 Planar crystal packing in the X-ray structure fits the in situ packing of the native S-layer.** a 2D schematic of H/L complex crystal packing, indicating the interaction network linking a single H/L (slate blue/gold) complex with six other molecules in a planar arrangement generated by SLPA1 tiling. Array is depicted as seen from the extracellular environment, with symbols representing key interaction types in the crystal lattice, detailed in Supplementary Fig. 4. b Cartoon representation of the H/L planar array (PDB ID 7ACY, coloured as in a, views as defined in Fig. 1b). c Native C. difficile S-layer ghosts (electron micrograph, negatively stained, left. Scale bar: 2 µm) were used to compute Fourier transforms (middle). Micrograph is a lower magnification (×3000) of ghosts used to collect the images used (82351x) for Fourier transform computation and is representative of the morphology of S-layer ghosts. Typically spots from two or more lattices were observed. Fourier transform is representative of 36 images collected for R20291. Reciprocal lattice axes (red and white axes) are indicated for two observed lattices (scale bar 0.0125 Å). d Orthogonal views of the 3D reconstruction of negatively stained S-layer ghost indicating the overall envelope in the native lattice. A rigid body fit of the structure of H/L complex determined by X-ray crystallography (PDB ID: 7ACY, cartoon representation, SLPA1 - gold, SLPL - slate) indicates a similar arrangement in the native S-layer ghosts and crystal packing. Reconstruction is shown from the environment (top left) and cell wall (top right), and side views in the 2D plane (bottom panels).

models, with no apparent effect on the fold of individual regions, suggests how the effector domains of other CWPs inserted in the functional S-layer can be accommodated by flexible rearrangement of D2. Interestingly, despite the overall similarity between our models, different SLPS can adopt alternative conformations: in SlpaR7404 (PDB ID: 7ACW), one SLPL molecule packs away from SLPA1 when compared to SlpaACD630 (PDB ID: 7ACY) due to a rotation of 9° around the LID/HID axis, further strengthening the idea that D1-D2 and D1-LID linkers provide flexibility to the structure. This flexibility among the most external domains, with certain loops acting as hinges for flexible conformation of parts of the structure, has also been described in other known S-layer structures and could have a functional role in S-layer assembly.

Crystal lattice reflects in situ S-layer assembly. Due to the propensity of S-layer proteins to spontaneously self-assemble into 2D structures, we hypothesised that the packing of our crystal structures might be reflective of native S-layer assembly. In the crystal, two H/L complexes, related by pseudo-twofold symmetry, are present in the P1 asymmetric unit, packed in a 2D planar array (Fig. 3) which is then stacked to extend the crystal into the third dimension. The 2D lattice is achieved by tiling of SLPA1, with interlocked ridges of SLPA1 molecules covering gaps between the tiles, creating a tightly packed layer (Fig. 3a, b). Lattice contacts between CWB2 motifs of neighbouring SLPA1 molecules involve helix-helix interactions between the symmetry-related copies of α12 (see topology in Supplementary Fig. 1), as well as electrostatic interactions generating a tightly bonded network (Fig. 3a and Supplementary Fig. 5a).

We investigated if the planar crystal packing observed in the X-ray structure reflects the in situ packing of the native S-layer assembly. Intact S-layer extracted from C. difficile vegetative cells formed collapsed capsules exactly mimicking the size and shape of the originating cell (native S-layer ghosts). These double-layered 2D crystals were interrogated by electron crystallography, with rotationally separated diffraction patterns observed from images of the superimposed layers (Fig. 3c). As hypothesised, the p2 symmetric 2D lattices of native S-layer ghosts (a = b = 80 Å, c = 100°) were consistent with unit cell parameters of the stacked lattices in the 2D plane of the X-ray crystals (b = 78 Å, c = 80 Å, α = 100°) (Supplementary Tables 1–3), pointing towards a similar packing (Fig. 3c). The 3D reconstruction from images of the native ghosts revealed a molecular envelope with a staggered ridged surface on one face of the S-layer, with deep grooves between the parallel ridges, and an opposing surface defined by paired, globular domains arranged in rows (Fig. 3d). These features recapitulate the surface characteristics of the H/L array in the X-ray crystal structure. Indeed, manual fitting of the 2D X-ray...
lattice as a single rigid body into the EM density matches the ridged surface to SLPL$_1$ with the paired globular domains on the opposite face corresponding to the SLPL$_H$ CWB2 motifs (Fig. 3d). The ridges are also observed in cryo-electron microscopy (cryo-EM) side views of intact cells (Fig. 3c). This confirms that the X-ray crystal lattice of the H/L complex has the same overall arrangement as the in situ lattice of a mature S-layer in intact cells, therefore establishing our crystallographic model as a template for S-layer assembly in *C. difficile*.

**Principles of S-layer assembly.** Analysis of crystal packing of the H/L complex revealed key principles governing S-layer assembly. The charge distribution generated by the trimeric arrangement of the CWB2s provides complementary charges across the lateral faces of the SLPL$_H$ triangular prism tile (Fig. 1c), allowing these interactions to be established (Fig. 3a and Supplementary Fig. 5a). The structural analysis of homology models of other SLCTs suggests that most of the interactions between neighbouring SLPH molecules, which define those interfaces, are conserved across different SLCTs (Supplementary Fig. 5b), indicating that they are likely to be important for *C. difficile* S-layer assembly. It is worth noting that the most conserved interactions are at the interface of neighbouring CWB2–CWB2 motifs but also involve residues from SLPL$_1$ (Fig. 3a and Supplementary Fig. 5b, top 6 rows). This conservation suggests that SLPL$_1$ is also important for maintaining S-layer packing, together with the CWB2 assembly motifs, and that targeting these interactions could lead to disruption of the array.

Our analysis of the charge distribution of the CWB2 motifs in Cwp6 and Cwp8 (Supplementary Fig. 2) indicates that the charge complementarity between the H/L complex and these minor S-layer components would also be possible. However, homology between CWPs and SlpA is restricted to the CWB2 trimeric motif as CWPs have distinct accessory domains, replacing the SLPL$_1$. These structurally diverse domains are presumably accommodated in the S-layer whilst maintaining the integrity of the crystalline array. Based on the observation that most of the X-ray crystallographic model of the H/L complex fits well into the envelope defined by the EM reconstruction except for part of the D2 domain, we propose that this region of SLPL$_1$ might confer further flexibility to the assembled S-layer. It is possible that D2 adopts a slightly different position relative to SLPL$_1$ in the mature S-layer as it could be better accommodated by a rotation of 11° relative to D1 into the EM envelope (Supplementary Fig. 6). It is worth noting that this slight reorientation of D2 does not alter the overall packing of H/L complex into the 2D array (Supplementary Fig. 6). The strikingly different arrangements of the D1-D2 domains in SLPL$_1$ relative to the interacting domains observed in the H/L and SLPL$_1$/HID structures (Fig. 2a), are consistent with this proposed flexibility.

While increasing numbers of S-layer structural models are available,

The X-ray crystal structure of the resulting H/L complex (Fig. 4a, PDB ID: 7ACZ) superimposes readily onto the full-length SlpA$_{C. difficile}$ model (core RMSD 1.5 Å), with the absence of D2 not perturbing the overall protein fold. Moreover, the crystal lattice is similar to wild type, with equivalent interactions between SLPH$_H$ tiles and D1 domains (Supplementary Fig. 7).

**Permeability of the S-layer.** S-layers have been proposed to act as a molecular filter but given the tight packing of the S-layer, the bulk of the *C. difficile* cell surface seems virtually impenetrable to large molecules; this is despite being able to secrete toxin in the apparent absence of cell lysis. Other S-layer proteins have been proposed to create relatively permeable arrays with pores ranging from ~30 Å up to 100 Å in diameter and possibly wide. The interlocked D2 domains of adjacent SLPL$_1$ molecules cap this pore, further reducing access from the external environment to the cell wall (Fig. 5b–e, top). The second pore, formed between two SLPL$_1$ (Fig. 5a, molecules 1 and 2) is fully accessible from both outer and inner surfaces of the layer. It has a width of approximately 11 Å (pore 2, Fig. 5a, pink arrow), but is narrowed to 8 Å by two pseudo-symmetry equivalent arginine residues within 10 Å of the pore outward side (Fig. 5b–e, bottom). Importantly, the absence of D2 exposes pore 1 between SLPL$_1$ tiles (Fig. 5b–e, middle), which is occluded by interlocking D2 domains in the full-length structure. This creates two openings in the array of about 16 Å, which could indicate a more permeable S-layer than in the wild-type structure, with twice as many pores, of slightly increased size. Moreover, many of the residues lining the two exposed pores in this lattice are not resolved in the electron density of the SlpA$_{C. difficile}$ and could not be modelled, suggesting higher flexibility and, therefore, potentially weaker interactions.

The lining of the pores observed in the crystal lattices (Fig. 3) are highly hydrophilic (Fig. 5d), suggesting that only small
hydrated ions could easily diffuse into the cell. Pore 1, mostly occluded by D2 in the full H/L complex, has a mixed charged distribution, with patches of both positive and negative charges throughout (Fig. 5b–e). In contrast, the fully exposed pore 2 is mostly negatively charged, indicating that positively charged small metabolites could preferentially diffuse via this pore. The electropositive patch formed by two pseudo symmetry-related lysines covering the outermost opening of this pore and the two arginines at the innermost side (Fig. 5e) could provide some restricted access for charged species.

It is worth noting that interacting D1 domains from neighbouring SLP1 molecules completely cover the widest cavity in the SLP1 H/L CWB2s tiling. This interface, defined by neighbouring CWB21-CWB22 motifs, at around 20 Å wide but spanning over 100 Å across the triangular prism tiles, is also hydrophilic, with complementary charges (Fig. 6a). The SLP1 H/L CWB2 motif tiling also creates a cavity of ~70 Å between symmetry-related molecules which is partly occluded by the HID and LID domains, with the interlocking D1 domain ridges covering this gap (Fig. 6a). If the interacting D1 domains are flexible and can at least partially expose these cavities, it could potentially allow diffusion of larger, charged molecules through the S-layer.

Surprisingly, strain RAD2 displays susceptibility to lysozyme, unlike the parental R20291 wild-type, which is completely resistant (Fig. 6b). Importantly, strains devoid of SlpA had been previously reported to also be lysozyme-sensitive 13. Although lysozyme has an average diameter of about ~30 Å and the pores observed in the RAD2 strain are only about 16 Å, absence of D2 seems to be sufficient to allow lysozyme to kill C. difficile. It is possible that less stable D1-LID/HID interactions with the CWB2s due to absence of the D2 domain leads to wider openings that render the S-layer permeable to lysozyme at discrete positions. In any case, our work points to an important role of D2 in preventing access of at least some antimicrobials, although the precise mechanism is still unclear.

The fitting of our H/L crystal model in the in situ EM reconstruction shows that a similarly tight packing is present in the cells. It is important to note that, unlike a previously reported electron microscopy analysis of S-layer, which included nanobodies intercalated among SbsB molecules 26, our EM reconstruction is calculated from the native S-layer, containing not only H/L complexes but possibly other CWPs as well. However, as H/L constitutes 80–95% of the S-layer protein content, the contributions of the other CWPs are likely to have averaged out in the final reconstruction. Moreover, our reconstructions have been calculated from relatively small, well-ordered regions of the S-layer; it is possible that these regions are enriched in H/L. Any differences in packing and pore size would also not necessarily be visible after averaging.

**S-layer assembly: how can a 2D crystal array bend around a cell?** To act efficiently as a molecular sieve, the paracrystalline S-layer must be accommodated around the curved surface of the C. difficile cell and allow cell growth and division. Our recent work showed formation of C. difficile S-layer at specific sites that coincide with cell wall synthesis 28, suggesting discrete S-layer assembly points. Indeed, dynamic flexibility between S-layer protein domains has been shown to promote efficient crystal nucleation on the curved cellular surface in Caulobacter crescentus 27. Fourier analysis of S-layer ghosts, and tomographic reconstruction shows that a similarly tight packing is present in the cells. It is important to note that, unlike a previously reported electron microscopy analysis of S-layer, which included nanobodies intercalated among SbsB molecules 26, our EM reconstruction is calculated from the native S-layer, containing not only H/L complexes but possibly other CWPs as well. However, as H/L constitutes 80–95% of the S-layer protein content, the contributions of the other CWPs are likely to have averaged out in the final reconstruction. Moreover, our reconstructions have been calculated from relatively small, well-ordered regions of the S-layer; it is possible that these regions are enriched in H/L. Any differences in packing and pore size would also not necessarily be visible after averaging.
**Fig. 5 C. difficile S-layer is a tightly packed array with very narrow pores.**

**a** Surface representation of wild-type H/L (7ACY, left) and Slp$\Delta$D2 H/L (7ACZ, right) crystal packing showing pores in the 3D crystal lattice. Positions of pores marked with arrowheads (pore 1 in magenta, pore 2 in cyan) are equivalent in both lattices. **b** Zoomed-in view of the pores generated by H/L multimerization. Pore 1, top view covered by D2 in SlpACD630 (first panel) and in Slp$\Delta$D2 (second panel). Pore 2 – top view for SlpACD630 (third panel) and Slp$\Delta$D2 (fourth panel). Widest openings are labelled for each pore. Arrows indicate the widest points in each pore, which are exposed in SlpACD630 due to the lack of D2 that completely covers it in SlpACD630. **c** Cross-section views of pore 1 and pore 2 in SlpACD630 (first and third panels, respectively) and Slp$\Delta$D2 (second and fourth panels, respectively). Neighbouring SLPH (slate blue) and SLPL (gold) molecules that create the pores are shown in surface representation. **d** Hydrophobicity characteristics of the residues lining pore 1 (top) and 2 (bottom) calculated in ChexVis (see Methods section for details) according to Kyte-Doolittle scale, ranging from hydrophilic (green) to hydrophobic (blue), as per hydrophobicity gradient key. **e** Poisson-Boltzmann electrostatic potential calculated for residues lining pore 1 (first panel) and 2 (second panel) in SlpACD630 represented as a charge distribution (positive in blue and negative in red, as per electronegativity gradient key). Views and scale are as in **c** (left) and as a slice across the largest pore surface (right). Pseudo-symmetry-related lysine residues at the top and arginine residues at the bottleneck of pore 2 are highlighted.
imaging (Supplementary Fig. 9), indicate a highly mosaic surface, with many crystal defects, particularly at the cell poles, where the paracrystalline array must be distorted to allow for cell curvature. The observed pattern of crystalline patches with grain boundaries is consistent with the random secretion of S-layer protein monomers and self-assembly of 2D crystals occurring at gaps and grain boundaries within the curved S-layer, as proposed for other organisms \(^{28,29}\). It is possible that these mismatch points are wider than the observed pores and would therefore be sufficient for access of proteins and other molecules. However, whether this is the only mechanism for import/export of molecules across the S-layer and how these processes can be controlled are key questions to pursue to further our understanding of S-layer in \(C.\) \(difficile\).

**Discussion**

Here we report the experimentally determined structure of a complete S-layer from a medically significant pathogen, which allows us to understand the organisation of the paracrystalline array at an atomic level. The packing of H/L complexes in the crystal replicates the assembly of the functional S-layer observed in situ by electron microscopy.

One of the most surprising features of the S-layer in \(C.\) \(difficile\) is its compactness. While other S-layer assemblies proposed to date \(^{2,5,6}\) suggest arrays porous enough (30–100 Å) to allow substantial molecules to be imported or exported, pores in the \(C.\) \(difficile\) S-layer would only allow passage of small metabolites or hydrated ions. While the assembly domain SLPH tiles create a 2D array, the SLP\(_H\) ridges cover the pores present within the triangular SLPH prism packing, generating a structure impermeable to large molecules, including folded proteins such as lysozyme. How large molecules, such as the \(C.\) \(difficile\) toxins \(^8\) and other secreted proteins, are exported to the environment and how nutrients are acquired remains unclear. Moreover, S-layers must also be able to accommodate cell growth and division, while adapting to the curvature of the cellular poles. Having tightly packed core sub-units or domains, maintained by interchangeable electrostatic interactions, decorated with more flexible regions, is a simple and effective way to achieve both requirements. Points of mismatched symmetry, as observed by tomography (Supplementary Fig. 9), could provide increased flexibility and permeability, perhaps creating discrete points for import of nutrients and export of larger molecules.

Currently known S-layer structures show great variability in terms of domain organisation, fold and assembly, as well as little sequence conservation across species \(^{30,31}\). Surprisingly, the tight packing of SlpA goes against our understanding of the need for a degree of permeability of the molecular sieve. Nevertheless, comparing the different known structures of individual S-layer domains and proposed assemblies does reveal some common features. Notably, all known anchoring domains, from \(B.\) \(anthracis\) \(^{19}\) and \(Pseudo\) \(bacillus\) \(alvei\) \(^{32}\) SLHs to the \(Geobacillus\) \(steamthermophilus\) secondary cell wall polysaccharide binding motif \(^{18}\), and \(Caulobacter\) \(crescentus\) C-terminal domain \(^6\), have an \(\alpha\)-helical fold. Conversely, known assembly domains rely on \(\beta\)-strand rich protomers that pack to create an array, with some domains extending to link neighbouring molecules \(^{2,5,18,33}\). Strikingly, in \(C.\) \(difficile\), the assembly and anchoring functions
are combined in the largely α-helical SLP₄ but the β-strand-rich, environment-exposed, SLP₃ also contributes to assembly contacts. Therefore, the preference for helical domains facing the cell wall and β-strand rich external domains seems to be maintained. Moreover, a two-level organisation of domains creating a 2-layer array, was reported recently using cryo-electron tomography and cryo-EM imaging of whole cells of B. anthracis, Sulfolobus spp. and C. crescentus.⁵⁶,²⁵. This feature is also seen in C. difficile (Fig. 3b), further suggesting that these might be general S-layer features. Further structural and functional studies of other S-layers will improve our understanding of the role and mechanisms of these fascinating 2D arrays.

Importantly, our characterisation of the S-layer assembly in C. difficile also reveals new potential therapeutic avenues. Recently reported antibodies targeting a conserved region of SLP₄ recognize a region in CWB₂, facing the cell wall. It will be interesting to investigate if these antibodies might disrupt the S-layer in C. difficile, as has been observed in B. anthracis using nanobodies.⁵ If that is the case, molecules which affect S-layer assembly, by targeting the interacting SLP₄ subunits and the flexible D₂ domains, are attractive therapeutic agents.

Methods

Strains and growth conditions. C. difficile and E. coli strains are described in Supplementary Table 4. E. coli strains were routinely grown at 37 °C in LB broth and on LB agar (Oxoid). C. difficile strains were routinely grown under anaerobic conditions at 37 °C on brain heart infusion (BHI, Oxoid) or supplemented BHI (BHI-S)⁵⁵ agar and in TY broth. Growth media were supplemented with chloramphenicol (15 μg ml⁻¹), thiopanamycin (15 μg ml⁻¹) or kanamycin (30 μg ml⁻¹) as required.

Construction of RΔD₂. DNA oligonucleotides are described in Supplementary Table 3. Plasmid pRP223, containing a copy of the complete slpA gene from C. difficile, was amplified by PCR using oligonucleotides RP102 and RP103 to delete the coding sequence of SlpA residues 115–259 and replace with GGA GGT, encoding two glycine residues. The resulting plasmid, pOB001, was transferred to the C. difficile S-layer mutant strain FM2.⁵¹ by conjugation. FM2.5 displays an aberrant colony morphology that is easily distinguished from wild-type C. difficile. Recombination between the plasmid-borne slpA and chromosomal slpA gene was confirmed by PCR and the resulting protein profile was determined by SDS PAGE of S-layer proteins isolated using standard methods (see below).

Concentration and purifications. Cell lysis was performed using BugBuster Protein Extraction Reagent (Novagen). Pellets were resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 100 μM NaCl, 1x Complete EDTA-free protease inhibitors (Roche), 100 μg ml⁻¹ lysozyme, 10 μg ml⁻¹ DNase I, 1x BugBuster) and incubated at room temperature for 30 min. Extracts were centrifuged at 20,000 × g for 30 min, supernatant was filtered and separated on a 5 ml HisTrap (GE Healthcare) column in 50 mM Tris-HCl pH 8.0, 250 mM NaCl, 10 mM imidazole, with a linear gradient of imidazole (5–30 mM). Eluates were subjected using low pH glycine as described above and normalised to an equivalent OD₅₆₂₀ of 25. Culture supernatants were filtered, concentrated to an equivalent OD₅₆₂₀ of 50 using a Vivaspun column with a 10 kDa MWCO. Samples were then subjected to SDS PAGE and western blotting with specific antibodies for the C. difficile SlpA (pRPF170; dilution 1:2000), or SlpP (dilution 1:1000) and detected using goat HRP-conjugated secondary antibody (Invitrogen) used at 1:10,000.

Protein analysis by western immunoblotting. For analysis of H/L complex interactions on the surface of C. difficile, plasmids carrying a tetacycline-inducible copy of CDS₀₃ slpA (pRPFI70) or derivatives with a point mutation in the LID (F247A, pRPFI90) or the HID (Y27A, pAK186) were transferred into the slpA null strain FM2.5 by conjugation.⁵⁵. Strains were grown to an OD₅₆₂₀ of ~0.4 in TY broth and induced with anhydroe ROTAC and IPTG (20 μg ml⁻¹). Surface localised H/L complex subunits were extracted using low pH glycine as described above and normalised to an equivalent OD₅₆₂₀ of 25. Culture supernatants were filtered, concentrated to an equivalent OD₅₆₂₀ of 50 using a Vivaspun column with a 10 kDa MWCO. Samples were then subjected to SDS PAGE and western blotting with specific antibodies for the C. difficile SlpA (pRPF170; dilution 1:2000), or SlpP (dilution 1:1000) and detected using goat HRP-conjugated secondary antibody (Invitrogen) used at 1:10,000.

Protein expression and purification. S-layer was extracted as previously described.²⁶ Briefly, Maxisorp microplate plates (Nunc) were coated with 10 μg ml⁻¹ of SlpP or SlpP and their variants (Supplementary Table 4), blocked with 3% (w/v) milk in PBS/0.05% Tween-20 and overlaid with respective interacting partner SlP or SlP, and resuspended in 0.0001–100 μg ml⁻¹ range. Binding was assessed with polyclonal rabbit primary antibodies against the overlay protein (α-Slp, at 1:3000 and α-Slp at 1:15,000 dilution). Spectrophotometric detection of product of horse radish peroxidase (HRP)-conjugated secondary anti-rabbit antibody used at 1:2,500 dilution with an amperometric assay with a phenylenediamine dihydrochloride (OPD; 1 mg ml⁻¹) with hydrogen peroxide to the reaction mix was carried out in 490 nm using Biotek ELx800 plate reader.

X-ray crystallography. Purified and concentrated proteins (recombinant LID/ HID-6xHis-tag at 38 mg ml⁻¹ and SlpP/HID-6xHis-tag at 21 mg ml⁻¹, CDS₀₃, R7404 and RΔD₂ L/H at 10 mg ml⁻¹) were subjected to crystallisation using a Mosquito liquid handling robot (TTP Labtech), with the sitting drop vapour diffusion method at 20 °C. Native H/L complex crystallised in 0.1 MES pH 6, 0.12 M magnesium chloride, 400 μl per well. Crystals were obtained in 0.1 MES pH 6, 0.12 M magnesium chloride, 16% PEG 6000 and 10 % glycerol. Recombinant SlpA/HID-6xHis-tag produced diffraction quality crystals in 0.2 M ammonium sulphate, 0.1 M MES pH 6.5, 35 % PFD, while LID/HID-6xHis-tag was crystallised in 1.6 M sodium citrate tribasic dihydrate pH 6.5. Data were collected on the I04 (λ = 0.98 Å), I23 (λ = 2.75 Å) and I42 (λ = 0.97 Å) beamlines at the Diamond Light Source Sconfentro (Didcot, UK) at 100 K. The data were collected from the automatic software pipeline xia2 within the Information System for Protein Crystallography Beamline (ISPBy), processed with DXS, ⁵, MissFOM²⁸ or DIAlS²⁹ and scaled with Aimless within CCP4i or CCP4i2 software, when needed. Density modification was performed with PARROT³⁰. The structure of LID/HID was solved de novo using Arcimboldo_lite within CCP4i and re-determined with Coot with DIALS and resolution 1.9 A. The structure of SlpP/HID-6xHis-tag was determined by sequential molecular replacement in Phaser³¹ searching first for SlpP, D1-D2 domains and

ARTICLE NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-022-28196-w | www.nature.com/naturecommunications

10 NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-022-28196-w | www.nature.com/naturecommunications
model (3CVZ), followed by the search with LID/HID structure into a fixed SLPļ solution, and subsequent manual building (COOT) and refinement (Phenix). Initial attempts to solve the structure of the complete H/L complex by S-SAD provided only weak phases, and were improved by combining molecular replacement in Phasor using the CWR2 domain core of C. difficile Casp8 (PDB ID: 5J6Q) with sulphur anomalous difference Fourier maps using Anodo. This solution provided an S-SAD in Phenix of 3.3 Å resolution. However, a fit with LID/HID and D2 from SLPļ/HID structure and D1 from CD630 SLPļ (PDB ID: 3CVZ) with loops removed. Final models were obtained after iterative cycles of manual model building with Coot and refinement in Phenix_refine and REFMAC with sulphur anomalous difference Fourier maps using Anodo. This solution provided an S-SAD in Phenix of 3.3 Å resolution. To improve electron density maps we used molecular replacement runs using Phaser in CCP4i2 combining the SLPL model with the obtained LID/HID, D2 from SLPļ/HID structure and D1 from CD630 SLPļ (PDB ID: 3CVZ) with loops removed. Final models were obtained after iterative cycles of manual model building with Coot and refinement in Phenix_refine and REFMAC as well as DPD-REDO. Applied strategies included refinement of XYZ coordinates, real space, individual B-factors, TLS parameters and occupancies. Validation of final models was performed using Coot and Phenix internal tools, as well as MOLPROBITY. Data collection and refinement statistics are summarised in Supplementary Table 1. Structural representations were generated using PyMOL Molecular Graphics System (Schrodinger, LLC) or Chimera 1.13.2.

Electron crystallography data collection. To allow visualisation by electron microscopy, S-layers were either removed from C. difficile cells in a single piece following centrifugation or digestion (S-layer ghosts) or cells were mechanically fragmented (S-layer/cell wall fragments). C. difficile cells were harvested by centrifugation and resuspended to an OD600nm of 10 in 20 mM HEPES pH 7.5, 150 mM NaCl, 500 mM sucrose. For S-layer ghosts, cells were digested using purified CD27 endolysin for 30 min at 37 °C. The resulting membrane-bound spheroplasts were separated from the sample by centrifugation at 2000 µg for 15 min, and the supernatant, containing S-layer ghosts, was retained for imaging. 5 µl of S-layer ghosts were loaded on glow-discharged, amorphous carbon-coated 300 mesh copper EM grids and stained with 2% uranyl acetate, as previously described. Samples were examined on a Philips CM200 FEG transmission electron microscope at 200 kV. Images were collected on a 4096 × 4096 pixel Gatan UltraScan 4000SP Model 890 CCD camera (Gatan Inc.), with 15 µm pixel size. A total of 36 micrographs of R20291 and 29 micrographs for RΔ55 were collected using SerialEM, at a pixel size of 5.47 Å, with each tilt series covering ±60° with a tilt increment of 3° and collected with a grouped dose-symmetric acquisition scheme and group sizes of 4. Samples from each tilt series received 100 e⁻/Å² total dose with 20 frames per tilt. Tomograms were constructed using the IMOD package. Tilt series were tracked and aligned based on fiducial markers, and then tomograms were reconstructed by weighted back projection with 1x binning.

Lysosome resistance. To assess resistance to D. difficile cultures were grown in TY broth, subcultured to an OD600nm of 0.05 in 1 ml fresh TY in a 1.5 ml cuvette and then grown for 8 h with hourly OD600nm measurements. Where appropriate, lysosome (500 µg ml⁻¹) was added after 2.5 h growth. Experiments were performed in triplicate on biological duplicates and data expressed as the mean and standard deviation.

Other methods. Sequences of SLCTs were downloaded from D. difficile Multi Locus Sequence Typing website ([https://pubmlst.org/difficile/](https://pubmlst.org/difficile/)) and aligned using MAFFT. Multiple sequence alignments were visualised and annotated using Jalview or ALINE. Analysis of the evolutionary conservation of amino acids was performed using Consurf71 and the components of DynOomics webserver. Homology models for SLCT representatives were generated using ProtModel72, HingeProt72 and the components of DynOomics webserver. Structural alignments between homology model and template were performed using COOT. Data was analysed using Numpy v1.16.6, Pandas v0.24.2, and heatmap generated with Seaborn v0.10.174. Analysis of pores in the H/L array was carried out using ChexVis and hydropathy profiles for residues lining each pore calculated using the Kyte-Doolittle scale.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability. X-ray structural data are available in the PDB repository under PDB IDs: 7ACW, 7ACV, 7ACX, 7ACY and 7ACZ. Electron crystallography data are available in the EMDB repository under accession code EMD-13957 and the fitted model is available in the PDB repository with PDB ID: 7QGQ. Source data are provided with this paper. Any other datasets generated and/or analysed during the current study are available from the corresponding authors on request. Source data are provided with this paper.
References

1. Fagan, R. P. & Fairweather, N. F. Biogenesis and functions of bacterial S-layers. Nat. Rev. Microbiol. 12, 211–222 (2014).
2. Baranova, E. et al. Subunit structure and lattice reconstruction unveiling Ca2+-triggered S-layer assembly. Nature 487, 119–122 (2012).
3. Bharat, T. A. M. et al. Structure of the hexagonal surface layer on Caulobacter crescentus cells. Nat. Microbiol. 2, 17059 (2017).
4. Sichantra, D. et al. Molecular basis for the attachment of S-layer proteins to the cell wall of Bacillus anthracis. Biochemistry 57, 1949–1953 (2018).
5. Fioravanti, A. et al. Structure of S-layer protein Sap reveals a mechanism for therapeutic intervention in anthrax. Nat. Microbiol. https://doi.org/10.1038/s41564-019-0499-1 (2019).
6. von Kügelgen, A. et al. In situ structure of an intact lipopolysaccharide-bound bacterial surface layer. Cell https://doi.org/10.1016/j.cell.2019.12.006 (2019).
7. Smith, W. K., Lytras, D., Lacy, D. B., Wilcox, M. H. & Küppers, E. J. Clostridium difficile infection. Nat. Rev. Dis. Prim. 2, 16020 (2016).
8. Pruitt, R. N. & Lacy, D. B. Toward a structural understanding of Clostridium difficile toxins A and B. Front. Cell. Infect. Microbiol. 28, 2 (2012).
9. Ng, J. et al. Clostridium difficile toxin-induced inflammation and intestinal injury are mediated by the inflammasome. Gastroenterology 159, 542–552.e3 (2019).
10. Ryan, A. et al. A Role for TLR4 in recognition of surface layer proteins. Trends Microbiol. 29, 405–415 (2021).
11. Ravi, J. & Fioravanti, A. S-layers: the proteinaceous multifunctional armor of Gram-positive pathogens. Front. Microbiol. 12, 685 (2021).
12. Blackler, R. J. et al. Structural basis of cell wall anchoring by SLH domains in Paenibacillus alvei. Nat. Commun. 9, 1–11 (2018).
13. Arbing, M. A. et al. Structure of the surface layer of the methanogen Methanococcus marinus acetivorans. Proc. Natl Acad. Sci. USA 109, 11812–11817 (2012).
14. Lawry, B. M. et al. Species-specific detection of C. difficile using targeted antibody design. Anal. Chem. 90, 13475–13482 (2018).
15. Sorg, J. A. & Dineen, S. S. Laboratory maintenance of Clostridium difficile. Curr. Protoc. Microbiol. 12, 9.A.1–9.A.10 (2009).
16. Kirk, J. A. & Fagan, R. P. Heat shock increases conjugation efficiency of Clostridium difficile. Anaerobe 42, 1–5 (2017).
17. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D. Biol. Crystallogr. 66, 133–144 (2010).
18. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. D. Biol. Crystallogr. 67, 271–281 (2011).
19. Winter, G. et al. DIAls: implementation and evaluation of a new integration package. Acta Crystallogr. D. Struct. Biol. 74, 85–97 (2018).
20. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? Acta Crystallogr. D. Biol. Crystallogr. 69, 1204–1214 (2013).
21. Potoe, M. D. et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D. Biol. Crystallogr. 67, 235–242 (2011).
22. Potoe, M. et al. CCP4i2: the new graphical user interface to the CCP 4 program suite. Acta Crystallogr. D. Struct. Biol. 74, 68–84 (2018).
23. Cowtan, K. Recent developments in classical density modification. Acta Crystallogr. D. Biol. Crystallogr. 66, 470–478 (2010).
24. Rodriguez, D. D. et al. Crystallographic ab initio protein structure solution below atomic resolution. Nat. Methods. 6, 651–653 (2009).
25. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D. Biol. Crystallogr. 62, 1002–1011 (2006).
26. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. 60, 2126–2132 (2004).
27. Headd, J. et al. Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. Acta Crystallogr. D. Biol. Crystallogr. 68, 381–390 (2012).
28. McCoy, A. J. et al. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007).
29. Thorn, A. & Sheldrick, G. M. ANODE: anomalous and heavy-atom density calculation. J. Appl. Crystallogr. 44, 1285–1287 (2011).
30. Murshudov, G. N. et al. REFMACS for the refinement of macromolecular crystal structures. Acta Crystallogr. D. Biol. Crystallogr. 67, 355–367 (2011).
31. Foosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. The PDB-REDO server for macromolecular structure model optimization. J. Comput. Chem. 31, 213–220 (2010).
32. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D. Biol. Crystallogr. 68, 947–958 (2008).
33. Henderson, R., Baldwin, J. M., Downing, K. H., Lepault, J. & Zemlin, F. Structure of purple membrane from halobacterium halobium: recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. Ultramicroscopy 19, 147–178 (1986).
34. Crowther, R. A., Henderson, R. & Smith, J. M. X-ray image processing programs. J. Struct. Biol. 116, 9–16 (1996).
35. Gispen, W. H., Zeng, X. & Stahlberg, H. 2Dx: merge data management and merging for 2D crystal images. J. Struct. Biol. 160, 375–384 (2007).
36. Valpuesta, J. M., Carrascosa, J. L. & Henderson, R. Analysis of electron microscopy images and electron diffraction patterns of thin crystals of Ω29 connectors in ice. J. Mol. Biol. 240, 281–287 (1994).
37. Agard, D. A. & Moody, M. F. A least-squares method for determining structure factors in three-dimensional tilted-view reconstructions. J. Mol. Biol. 167, 849–852 (1983).
38. Amos, L. A., Henderson, R. & Unwin, P. N. T. Three-dimensional structure determination by electron microscopy of two-dimensional crystals. Proc. Natl. Acad. Sci. USA 39, 183–218 (1982).
39. Havelka, W. A., Henderson, R. & Osterhelt, D. Three-dimensional structure of halohorin at 7 Å resolution. J. Mol. Biol. 247, 726–738 (1995).
Croll, T. I. IUCr. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr. D Struct. Biol. 74, 519–530 (2018).

Kremer, J. R., Mastronarde, D. N. & McLintoch, J. R. Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. 116, 71–76 (1996).

Bond, C. S. & Schütte-Kopf, A. W. ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. Acta Crystallogr. D Biol. Crystallogr. 65, 510–512 (2009).

Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. 60, 2256–2268 (2004).

Krissinel, E. & Hendrick, K. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797 (2007).

Laskowski, R. A., Jalloulka, J., Pravda, L., Váckrová, R. S. & Thornton, J. M. PDBeum: structural summaries of PDB entries. Protein Sci. 27, 129–134 (2018).

Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J. Chem. Inf. Model. 51, 2778–2786 (2011).

Veevers, R. & Hayward, S. Methodological improvements for the analysis of domain movements in large biomolecular complexes. Biophys. Physicochem. 16, 328–336 (2019).

Emekli, U., Schneidman-Duhovny, D., Wolfson, H. J., Nussinov, R. & Haliloglu, T. HingeProt: automated prediction of hinges in protein structures. Proteins Struct. Funct. Genet. 70, 1219–1227 (2008).

Van Der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: a structure for efficient numerical computation. Comput. Sci. Eng. 13, 22–30 (2011).

Mckinney, W. Data Structures for Statistical Computing in Python. (SCIPY, 2010).

Waskom, M. et al. mwaskom/seaborn: v0.8.1. Zenodo https://doi.org/10.5281/zenodo.883859 (2017).

Masood, T., Bin, Sandhiya, S., Chandra, N. & Natarajan, V. CHEXVIS: a tool for molecular channel extraction and visualization. BMC Bioinform. 16, 119 (2015).

Kyte, J. & Doolittle, R. F. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. 157, 105–132 (1982).

Author contributions
P.L.M., O.B., J.W. and A.B.S. contributed equally to the work presented. P.L.M. and A.B.S. carried out X-ray crystallographic experiments, collected and analyzed X-ray data, determined and refined X-ray structural models, wrote and revised the manuscript. O.B. and J.W. carried out E.M. experiments, collected and analyzed data, revised the manuscript. A.B.S. carried out ELISA experiments and analyzed data. J.A.K. carried out in vitro experiments, analyzed the data and revised the final manuscript. F.V. contributed to data analysis, project discussions and revised the manuscript. S.O.B. contributed to in vitro experiments and revised the final manuscript. A.B.S., J.A.K., O.B., R.P.F. and S.O.B. constructed plasmids and strains. A.B.S. collected, analyzed X-ray data and supervised P.L.M. K.E.O., A.W. collected and analyzed the S-SAD data and revised the final manuscript. N.F. designed and supervised initial structural studies, revised the manuscript. G.R.D. designed experiments, analyzed the data, wrote and revised the manuscript. P.A.B. designed, supervised and analyzed E.M. studies, wrote and revised the manuscript. R.P.F. designed the study, prepared samples, collected X-ray data, analyzed data, supervised the study, wrote and revised the manuscript. P.S.S. designed the study, prepared samples, collected and determined X-ray structures, interpreted data, supervised the study, wrote and revised the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-28196-w.

Correspondence and requests for materials should be addressed to Per A. Bullough, Robert P. Fagan or Paula S. Salgado.

Peer review information Nature Communications thanks Han Remaut and the other, anonymous, reviewers for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

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

Acknowledgements
This work was supported by the Wellcome Trust [204877/Z/16/Z]. A.B.S., J.A.K. and F.V. were supported by this Wellcome Trust Collaborative Award, awarded to G.R.D., P.S.S. and R.P.F.; J.W. was supported by a BBSRC grant (BB/P02002X/1) awarded to P.A.B. and R.P.F.; P.L.M., O.B. and S.O.B. were supported by Faculty of Medical Sciences, Newcastle University PhD, University of Sheffield Imagine: Imaging Life initiative PhD and University of Sheffield PhD studentships, respectively. We thank Diamond Light Source for access to beamlines I23, I04 and 124 (ib15523-6, mx18598-24, mx18598-14, through the "Macromolecular Crystallography at Newcastle, Durham, Lincoln and Durham" BAG) and eBIC (through the EM19832 "Northern England Cryo-EM consortium" BAG). We thank Dr. Svetomir Tzokov for supporting EM data collection in the University of Sheffield Electron Microscopy Facility and Dr. Giuseppe Canonne for help with EM data collection at the MRC Laboratory of Molecular Biology, Cambridge. The contents of this work are solely the responsibilities of the authors and do not reflect the official views of any of the funders, who had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript.

© The Author(s) 2022