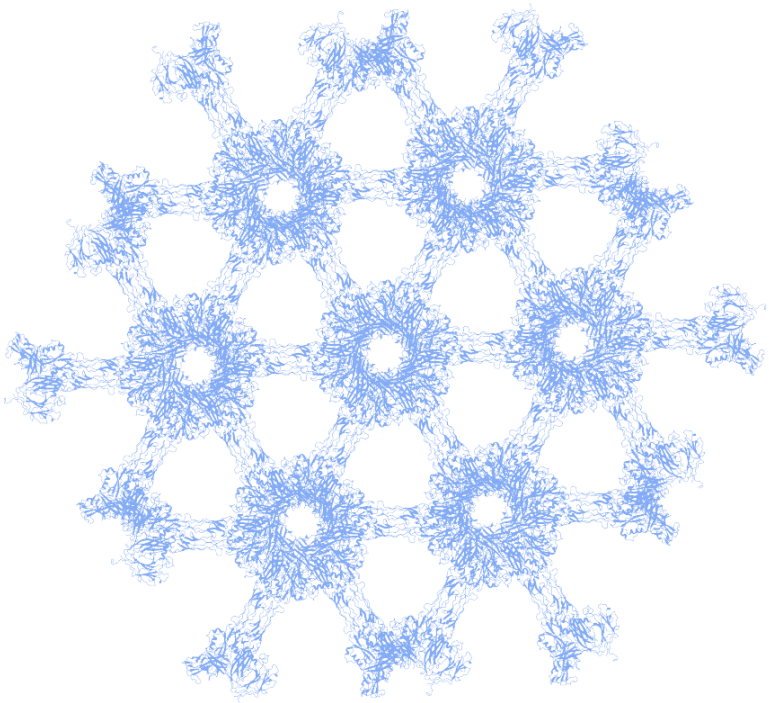
Sofie van Dorst | 6170870 | 23-07-2022

A closer look at the surface

Structural biology of prokaryotic cell surface molecules



Utrecht University

UNIVERSITY OF OXFORD

SIR WILLIAM DUNN SCHOOL OF PATHOLOGY

Minor Research Project for MSc Molecular and Cellular Life Sciences.
November 2021 – May 2022

Daily supervision: Andriko von Kügelgen, MSc.

Supervisor: Dr. Tanmay Bharat

Secondary examiner: Prof. Dr. Friedrich Förster

Bacterial Biofilms Group
Sir William Dunn School of Pathology
University of Oxford

Cover image: atomic model of Deinococcus radiodurans *surface layer lattice.*

Abstract

Most prokaryotes are encapsulated by a para-crystalline (glyco)protein array termed surface layer or S-layer, which mediates the interaction between the cell and its environment. Despite wide variation between prokaryotes and their S-layer protein sequences, S-layers share underlying organisational and assembly principles. Here, I used cryogenic electron microscopy (cryo-EM) and cryogenic electron tomography (cryo-ET) to study the molecular structure and organisation of the cell surfaces of archaea *Nitrosopumilus maritimus* and *Pyrobaculum arsenaticum,* the Gram-positive bacterium *Paenibacillus alvei* and the evolutionarily deep-branching diderm bacterium *Deinococcus radiodurans*. My work provides biological context to our structural knowledge of S-layer proteins and reveals common themes in S-layer protein structures. This work, along with other results, demonstrates that S-layers display a remarkable level of adaptation to the physical challenges of the prokaryotic environment, showing that higher order symmetric organisation and the presence of Immunoglobulin-like domains are ubiquitous in prokaryotic S-layers. Our combined approach of cryo-EM and cryo-ET shows potential for high-throughput S-layer structure determination and great promise for harvesting these structures for application in biotechnology. With our results, we have contributed to broadening the diversity of S-layer structural knowledge and of the underlying principles of prokaryotic cell surface biology.

Table of contents

Layman's summary

Prokaryotes (bacteria and archaea) are single celled organisms that form the majority of the diversity of life on earth. Prokaryotes are found in every nook and cranny of the earth and beyond and can live in extreme conditions such as freezing cold, boiling heat, sites of nuclear disaster or even in outer space. Unlike multicellular species like humans, prokaryotes cannot rely on specialised cells to protect them from their immediate environment, or an immune system to protect them from infections. Instead, most prokaryotic cell surfaces are covered in a layer of protein called the surface layer or S-layer. The S-layer surrounds prokaryotes like a chainmail, covering the outside of the cell in a regular pattern to help the cell interact with its environment. Because the shape or structure of a protein determines its function, it is important to find the molecular structure of S-layer proteins to try and understand how they interact with and protect the cell against its surroundings. In this work, we describe the structure of the S-layer of several prokaryotes, aiming to understand them within their biological context. We see that these S-layers have adapted to their specific environments, allowing the prokaryotes to survive in conditions of food scarcity, high temperatures and toxic radiation. We also show that despite the large differences between these species, prokaryotic S-layers share common structural features. While the overall structure of the proteins is different between species, all S-layers we have studied here are built from the same type of building block, the immunoglobulin-like domain, found also in the human immune system. This domain must have played an important role in the evolution of S-layer structures. Our work using state-of-the-art electron microscopy techniques gives us a better view of the cell surface of prokaryotes and helps us understand how S-layers aid bacteria and archaea to survive in the face of extreme conditions.

Introduction

The interaction of a prokaryote with its direct environment is critical for its survival. Vital processes like successful nutrient uptake, adaptation to physical stress and defence against pathogens and predators are examples of this interaction. As the first point of contact with the immediate surroundings, the cell surface mediates this interplay between cell and environment. In most prokaryotes, a para-crystalline lattice of (glyco)proteins termed the surface layer or S-layer forms the outer boundary of the cell (Bharat et al., 2021). The S-layer plays a critical role in mediating the interaction with the environment and is known to modulate cell permeation, nutrient uptake, cell-cell interactions, can serve as a binding site for external molecules and even as a point of entry for phages and viruses (Fagan & Fairweather, 2014; Sleytr et al., 2019). Moreover, the S-layer is thought to be important for maintaining cell shape integrity, especially in the face of harsh environments where extremophilic prokaryotes can reside (Sleytr et al., 2014). The wide prokaryotic diversity is reflected in a large variation among the protein sequences of S-layers across the bacterial and archaeal domains. Nonetheless, S-layers appear to share basic principles of organisation (Bharat et al., 2021). They consist of polymeric subunits with often high degrees of symmetry, which form a lattice tiling the cell surface. S-layers of both bacteria and archaea can functionally be divided into two parts, one forming the para-crystalline lattice, and a second anchoring the S-layer lattice to the cell surface **(Fig. 1)**.

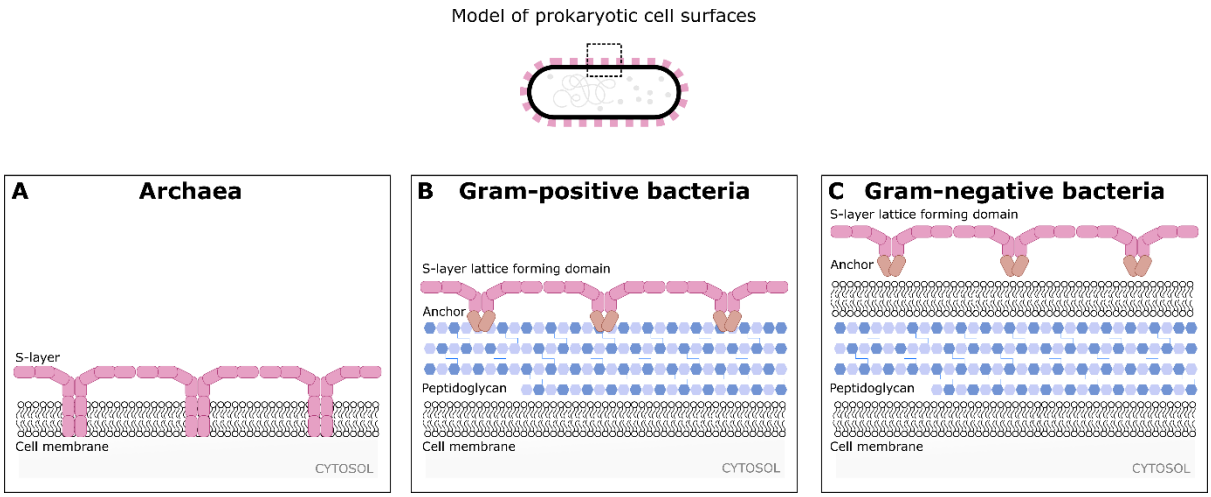


FIGURE 1 **Schematic representation of the prokaryotic cell surface.** A) Model of archaeal cell surface. The S-layer consists of a single protein embedded directly into the membrane. B) Model of the Gram-positive bacterial cell surface. The S-layer commonly consists of two proteins, one forming the S-layer lattice, and one anchoring the lattice to the peptidoglycan cell wall. C) Model of the Gram-negative bacterial cell surface. An anchoring domain links the S-layer lattice forming domain to the cell wall by binding the outer membrane, secondary cell wall polymers or lipopolysaccharides. Adapted from Bharat et al. (2021).

Structural knowledge of S-layer proteins is of major importance to understand how prokaryotes interact with their environment at a molecular level. However, only a handful of S-layer protein structures have been resolved, most of individual domains (Arbing et al., 2012; Baranova et al., 2012; Fioravanti et al., 2019; Lanzoni-Mangutchi et al., 2022; von Kügelgen et al., 2020, 2021)**.** This gap in our structural understanding of prokaryotic cell surface biology is due in large part to technical limitations. Because of their two-dimensional crystalline arrangement, S-layers are commonly not amenable to popular structural biology techniques such as X-ray crystallography or nuclear magnetic resonance. X-ray crystallography has been successfully applied to achieve atomic models of several S-layers protein domains (Arbing et al., 2012; Baranova et al., 2012; Bharat et al., 2017; Fioravanti et al., 2019; Lanzoni-Mangutchi et al., 2022), but is often a tedious process and does not provide biological context to the protein structure. As a result, there is only a limited record of high-resolution S-layer protein structures, and biological context in our structural understanding of S-layers is missing.

More recently, a combined approach using cryogenic electron microscopy (cryoEM) and cryogenic electron tomography (cryoET) was successfully applied to resolve the molecular structure of the S-layers of the Gram-negative bacterium *Caulobacter crescentus* and the archaeon *Haloferax volcanii* in their native contexts (von Kügelgen et al., 2020, 2021). These studies demonstrate purified S-layer proteins can be reconstituted into an S-layer lattice. This allows to obtain a high-resolution structure of the S-layer lattices using single particle cryoEM, that can be related to medium to high resolution structures obtainable by cryoET of prokaryotic cell surfaces from cells. It has been shown for a wide range of species that S-layers can be readily isolated from cells as fully assembled sheets (Messner & Sleytr, 1988). Using crude isolation of S-layer sheets, single particle cryoEM can thus be utilised to study S-layers at a structural level on a wider range of organisms, including biologically challenging species where sample availability and genetic tractability pose problems for conventional protein purification. In parallel, the application of cellular cryoET allows molecular structures to be contextualised and related to the biological functions of prokaryotic S-layers (Turk & Baumeister, 2020). The power of this combined approach was demonstrated by von Kügelgen et al. (2021), who used cryoET to show that while the *H. volcanii* S-layer is hexameric in isolated sheets, pentameric units are required to fully tile the cell surface *in situ*.

In this project, I addressed the lack of species diversity and biological context in our structural knowledge of prokaryotic cell surfaces by applying cryoEM and cellular cryoET to study cell surface proteins of 1) *Nitrosopumilus maritimus,* 2) *Pyrobaculum arsenaticum,* 3) *Paenibacillus alvei* and 4) *Deinococcus radiodurans* in relation to their biological function. I aimed to resolve high resolution structures of these proteins to start uncovering the some of the underlying principles of prokaryotic cell surface biology.

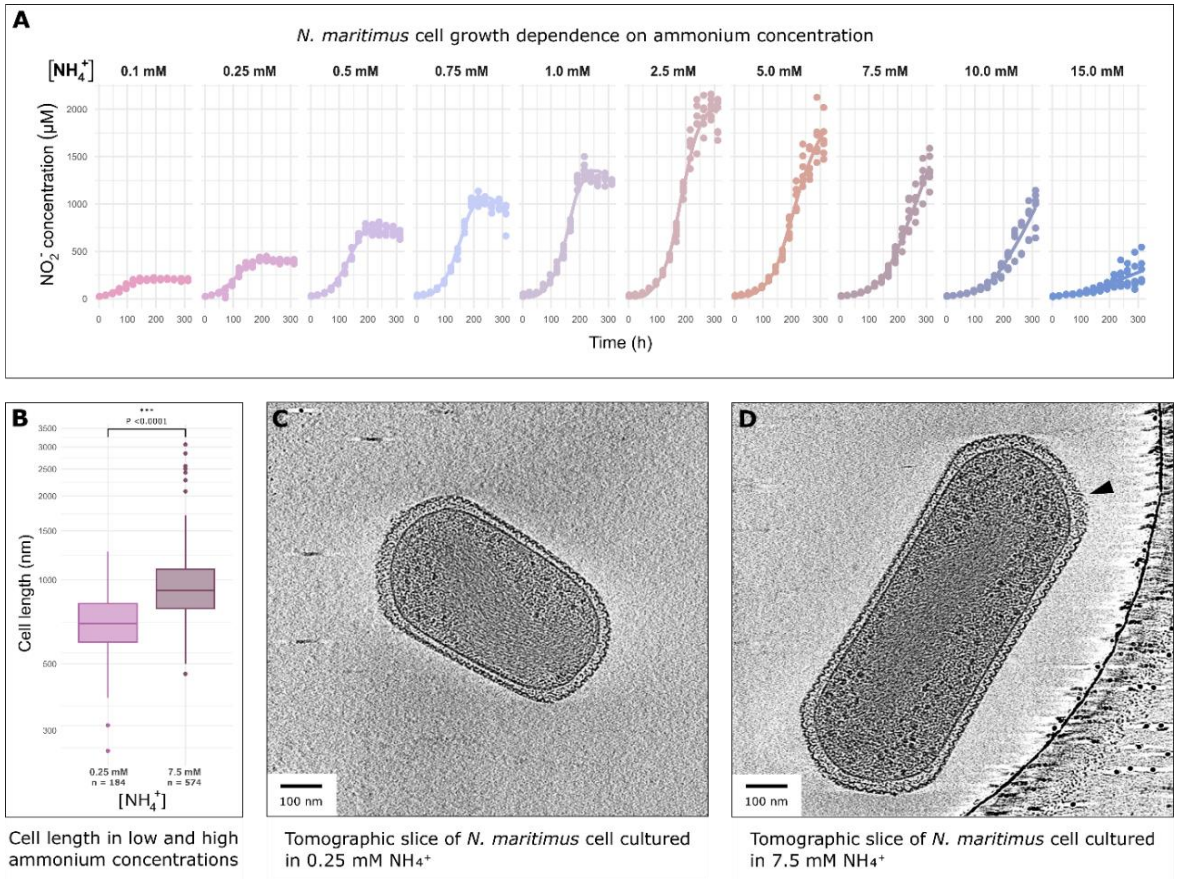
Results – S1 | **The *Nitrosopumilus maritimus* S-layer acts as a molecular sieve**

The Thaumarchaeon *Nitrosopumilus maritimus* is a chemolithoautotrophic archaeon isolated from marine environments. *N. maritimus* is predominantly found in the deep sea, where it survives in a nutrient deprived environment by using ammonium ions as its food source (Könneke et al., 2005). At just 200 nanometres in width and less than 1 micrometres in length, it is, to the best of our knowledge, the smallest free-living organism characterised. Despite its small size, *N. maritimus* and other ammonium oxidising archaea (AOA) enable a significant part of the global nitrogen cycle by performing the first step in nitrification. They are often considerably more abundant than ammonium oxidising bacteria (AOB), especially so in nutrient deprived environments (Leininger et al., 2006; Wuchter et al., 2006). In *N. maritimus*, ammonium ions are converted to nitrite ions by an ammonia monooxygenase enzyme in the cell membrane. The archaeal ammonia monooxygenase complex is not well described, and how AOA's can survive in nutrient concentrations two orders of magnitude lower than AOB is poorly understood (Park et al., 2010). Previous theoretical studies have hypothesized the S-layer plays a key role in the uptake of ammonium in *N. maritimus* (Li et al., 2018). Based on computational modelling, it was suggested that the presence of an ammonium converter separated from its environment by a lattice with pores will lead to a flux of ammonium into the cell periphery. Upon the introduction of negative charges along the pore, this effect is exacerbated. However, without structural and experimental evidence, the role of the S-layer in AOA nutrient-harvesting remains unclear. Here, I studied (in collaboration) the *N.* maritimus S-layer in relation to its nutrient source ammonium to understand the role of the S-layer in ammonium uptake.

First, we asked whether environmental ammonium concentrations affect *N. maritimus* cell growth and cell cycle regulation. To this end, we analysed *N. maritimus* ammonium conversion to nitrite as a measure of cell growth over time in a range of ammonium concentrations **(Fig. 2A)**. As can be expected from growth in food scarcity, reduced concentrations of ammonium lead to reduced growth rates and reduced time to reach stationary phase. A clear optimum growth can be observed from 1 mM, at which *N. maritimus* is commonly cultured, to 2.5 mM of ammonium. Higher concentrations of ammonium show an unanticipated response. Where most species saturate at surplus concentrations of their food source, *N. maritimus* shows reduced growth rates, as well as latency to reach stationary phase. Notably, cryoEM images of cells grown at low (0.25 mM) and high (7.5 mM) ammonium concentrations show that the environmental concentration of ammonium affects cell size. The average cell size of cells growing in low ammonium concentration is significantly smaller than cells growing in high ammonium concentrations (two sample t-test, $P < 2.2E\text{-}16$) and in cultures grown at high ammonium concentrations, there is a notable number of highly elongated cells

(Fig. 2B). Cellular tomograms show that in reduced ammonium concentrations cells have a dwarfed morphology but seem organisationally intact **(Fig. 2C)**. At high ammonium concentrations on the other hand, cells are often elongated and sometimes show disruption of the long-range crystalline order of the S-layer **(Fig. 2D)**. This implies that *N. maritimus* is able to adapt its cell size in response to environmental nutrient availability and suggests the cell cycle and cell growth, including S-layer elongation are affected by this adaptive response.

To further understand the role of the *N. maritimus* S-layer in relation to ammonium uptake, we next utilised state of the art cryoEM to solve the molecular structure of SlaA, the *N. maritimus* S-layer protein. Tomograms of *N. maritimus* show the S-layer fully tiles the cells surface **(Fig. 3A)**. Micrographs of isolated S-layer sheets were subjected to single particle analysis, which enabled us to resolve a cryoEM map to a 2.7 Å resolution **(Fig. 3B-C)**. SlaA consists of a string of immunoglobulin



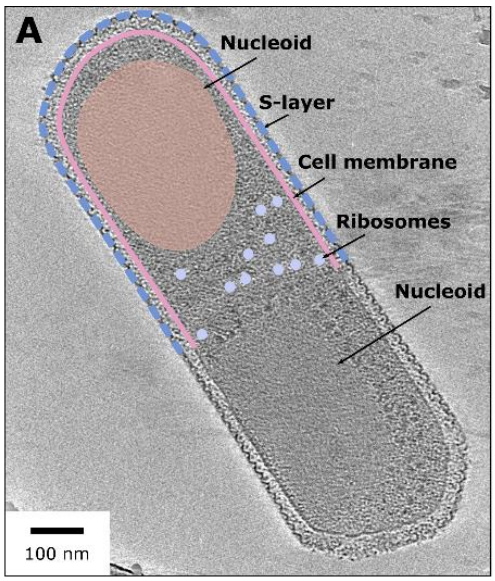
Cell length in low and high ammonium concentrations

Tomographic slice of *N. maritimus* cell cultured in 0.25 mM NH_4^+

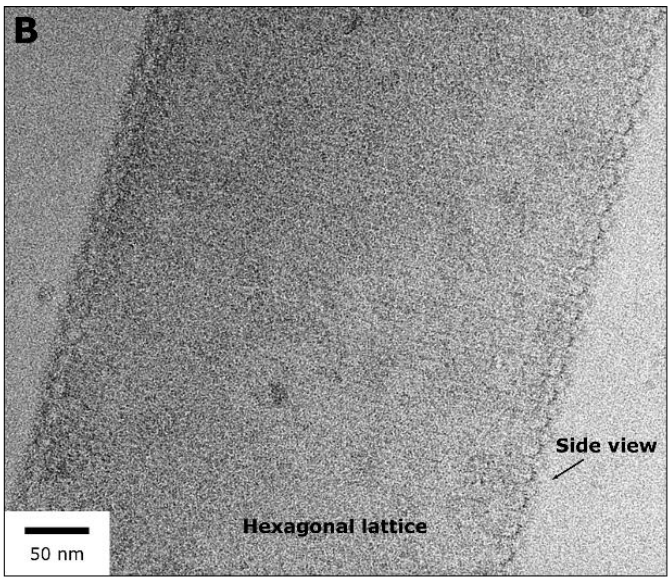
Tomographic slice of *N. maritimus* cell cultured in 7.5 mM NH_4^+

FIGURE 2 ***N. maritimus*** adaptation to environmental ammonium concentrations. A) *N. maritimus* cell growth over the course of 14 days in a range of ammonium concentrations (low to high concentrations are depicted from left to right). Culture density is measured as the concentration of metabolic product NO_2^-. B) Quantification of *N. maritimus* cell size (nm) in 0.25 mM (n = 184) and 7.5 mM (n = 574) respectively. Welch two sample t-test, P < 2.2E-16 (logarithmic scale). C) Tomographic slice of *N. maritimus* cell cultured in 0.25 mM ammonium. D) Tomographic slice of *N. maritimus* cell cultured in 7.5 mM ammonium. Arrow (black) indicates loss of long-range S-layer lattice order.

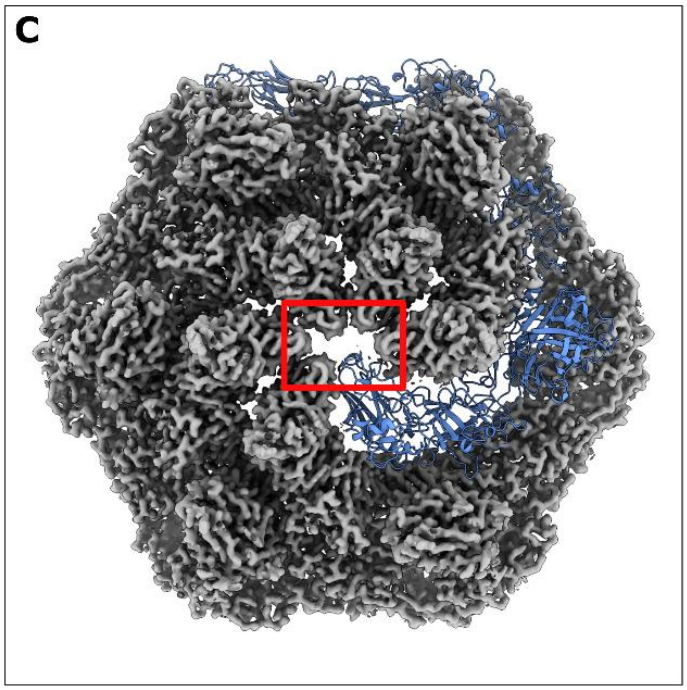
(Ig)-like domains, ordered in a crescent shape to form a hexameric protein complex. Notably, the hexameric centre forms a 20 Å pore, lined by negatively charged Asp and Glu residues **(Fig. 3D)**. To confirm that the negatively charged central pore could mediate ammonium import, we resolved the structure of SlaA from S-layer sheets isolated from cells growing in elevated ammonium concentrations (2.5 mM) to a resolution of 3.1 Å. In the cryoEM map derived from S-layers at higher ammonium concentrations, we find increased electron density within the central S-layer pore **(Fig. 3E)**. This shows increased presence of cations within the S-layer pore, likely to be ammonium, since only that concentration was varied between experiments. While the ion identity cannot be



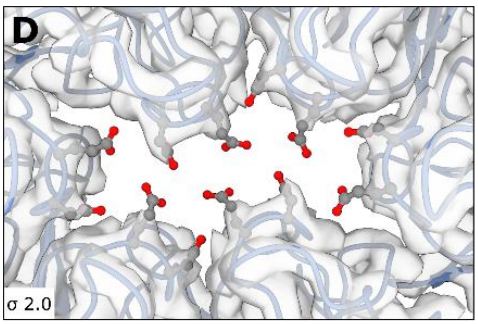
Tomographic slice of dividing *N. maritimus* cell



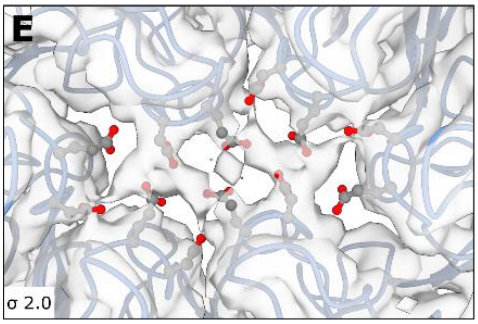
Electron micrograph of isolated *N. maritimus* SlaA sheet



cryoEM map of SlaA hexameric unit at 2.7 Å resolution



C2 pore at 1 mM ammonium, 2.7 Å resolution

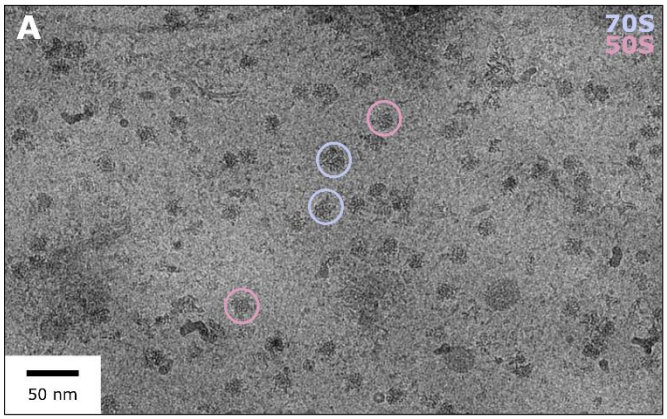


C2 pore at 2.5 mM ammonium, 3.1 Å resolution

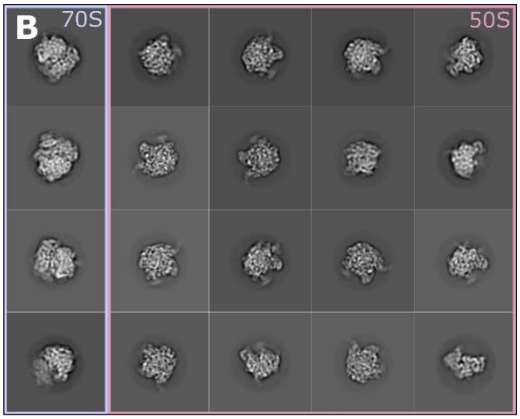
FIGURE 3 **Structure of the *N. maritimus* S-layer protein SlaA.** A) Tomographic slice of dividing *N. maritimus* cell. S-layer, cell membrane, ribosomes and nucleoids are shown. B) cryoEM micrograph of isolated *N. maritimus* S-layer sheet. C) CryoEM map (2.7 Å resolution) and ribbon representation of SlaA. CryoEM map of the SlaA hexameric unit is shown in grey, ribbon representation of atomic model of a single subunit is shown in blue. Red frame highlights the central pore. D) Central pore of SlaA hexameric unit from cells grown in 1 mM ammonium. E) Central pore of SlaA hexameric unit from cells grown in 2.5 mM ammonium. The atomic model is shown in blue, cryoEM density map (sigma level 2.0) is shown in transparent grey, oxygen atoms of acidic residues are shown in red.

determined confidently from the cryoEM map, the density in the pore is distinct from tightly coordinated calcium densities that are found between subunits and contribute to the structural integrity of the hexamer. We therefore propose this ion density to be ammonium, being attracted into the S-layer pore by the negatively charged residues lining the pore periphery.

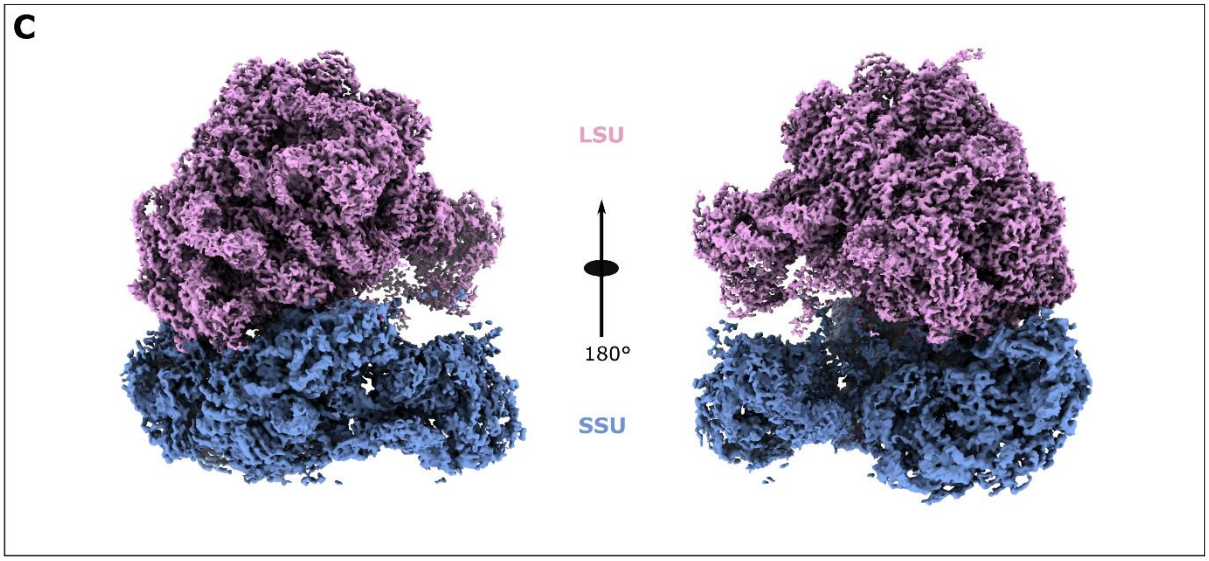
It is apparent the *N. maritimus* S-layer plays an important role in nutrient uptake. The S-layer is the most abundant protein in most prokaryotes thus forms a significant share of the translation load (Fagan & Fairweather, 2014). To accommodate the role of the S-layer in ammonium uptake as a well as observed adaptivity to fluctuating ammonium concentrations, there must be a coordination between translation of S-layer proteins and S-layer insertion. To understand how translation is regulated in relation to S-layer insertion, we aimed to study the spatial distribution of different translational states of the ribosome using cellular tomography. As an initial step to this, we resolved the molecular structure of the *N. maritimus* ribosome. Using a crude ribosome purification, we performed cryoEM single particle analysis to reconstruct the structure of this archaeal ribosome. Using a TOPAZ convolutional neural network (Bepler et al., 2020), we selected ribosome positions from electron micrographs **(Fig. 4A)**, resulting in 2D classes of the full, 70S ribosome as well as the dissociated 50S ribosomal subunit **(Fig. 4B)**. 50S and 70S ribosomes were separated using 3D classification, and structural refinement resulted in a structure of 3.4 Å global resolution for the 70S ribosome **(Fig. 4C-D)**. Further multibody refinement allowed resolution of the small subunit (SSU) of the ribosome to 3.8 Å and the large subunit (LSU) to 3.1 Å **(Fig. S1)**. Initial attempts at classification to separate ribosome translational states proved unsuccessful, presumably due to data quantity limitations.



Cryo electron micrograph of crude ribosome purification



Subset of selected 2D classes of ribosome



cryoEM map of N. maritimus ribosome at 3.4 Å global resolution

FIGURE 4 Structure of the *N. maritimus* archaeal ribosome. A) cryoEM micrograph of crude ribosome purification. Examples of full, 70S ribosomes are circled in light blue. Examples of 50S subunits are circled in pink. B) Subset of selected 2D class averages of 70S (light blue) and 50S (pink) ribosome. C) CryoEM map of the full *N. maritimus* ribosome after multibody refinement. Global resolution 3.4 Å, 50S large subunit (LSU, pink) 3.1 Å resolution and 30S small subunit (LSU, blue) 3.8 Å resolution.

Based on these results, we propose the *N. maritimus* S-layer functions as a molecular sieve for ammonium. With this structural information, we can propose an experimentally supported model for the *N. maritimus* cell surface architecture **(Fig. 5).** Environmental ammonium is attracted towards the S-layer predominantly by the negatively charged pores in the centre of S-layer hexameric units, although density for ammonium was also seen in other parts of the lattice between monomers and between hexamers where acidic residues are found. Continuous conversion of ammonium to nitrite by the ammonium converter in the cell membrane means there is an ammonium sink in the cell periphery. As demonstrated by Li et al. (2018)**,** this will lead to a positive chemical gradient, and

increased flow of ammonium into the cell periphery. Such a mechanism of nutrient-harvesting allows *N. maritimus* to thrive in its nutrient deprived deep-sea environment and outcompete AOB that require higher levels of ammonium for survival. The role of the S-layer in ammonium uptake in part explains our observations of aberrant cell growth at increased and reduced concentrations of ammonium. Changing environmental ammonium concentrations will alter the chemical gradient that is required to locally increase ammonium concentrations in the cell periphery. *N. maritimus* is thus required to adapt to fluctuating ammonium concentrations. In the case of low ammonium concentrations, reducing cell size will increase the surface to volume ratio of the cell, thereby increasing the nutrient uptake per volume. High levels of environmental ammonium might shift the chemical gradient and obscure the necessity of the S-layer in ammonium uptake, leading to the observed lattice breaks and loss of long-range lattice order. This process of adaptivity is likely coordinated with translational regulation. With our high-resolution ribosome structure in hand, we aim to study this translational coupling with S-layer insertion in detail using cellular cryoET and subtomogram averaging in the future. *N. maritimus* thus shows a sophisticated level of adaptivity towards life in a nutrient deprived deep-sea environment, where fluctuations in environmental ammonium concentrations are expected due to currents and seasonal changes.

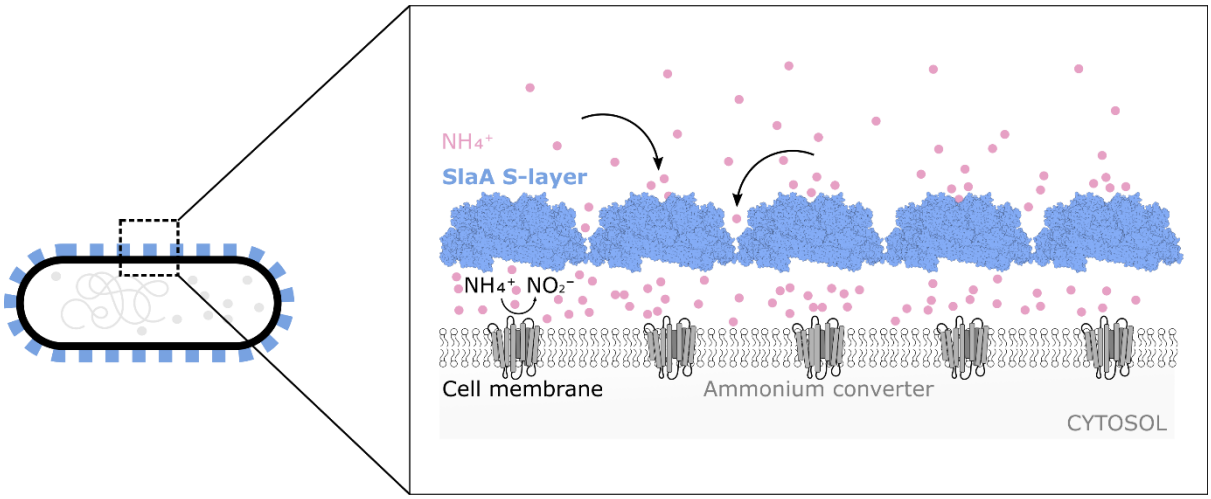


FIGURE 5 Schematic model of *N. maritimus* S-layer as a molecular sieve. Negative charges lining the *N. maritimus* S-layer pores draw in positively charged ammonium ions from the environment, increasing local concentrations of ammonium in the cell periphery. Continuous conversion of ammonium to nitrite by the ammonia monooxygenase enzyme leads to a chemical ammonium gradient over the S-layer, creating a flux of ammonium into the cell periphery.

Results - S2 | **Ig-like domains dominate the *Pyrobaculum arsenaticum* S-layer**

In more extreme environmental conditions, the role of the S-layer in prokaryotes can play a crucial role in the structural support of the cell. *Pyrobaculum arsenaticum* is a thermophilic archaeon and an obligate anaerobe that was first isolated from volcanic hot springs in northern Italy, where it thrives at 95 °C (Huber et al., 2000). *P. arsenaticum* is the first recorded arsenic oxidising species and plays a significant ecological role in heavy metal fixation within its high temperature ecosystem. Its ability to not just survive toxic arsenic concentrations but utilise them as a nutrient source shows *P. arsenaticum* is a highly specialised species, well adapted to its ecological niche. Early characterisations of the *Pyrobaculum organotrophum* S-layer have shown that *Pyrobaculum* subspecies have a hexagonal S-layer that fully tiles the surface of the cell, and demonstrate that even after harsh isolation of S-layer sheets the S-layer maintains its lattice conformation and cellular shape (Phipps et al., 1991). This suggests a highly stable and rigid S-layer structure, contributing to cell shape maintenance at high temperatures, a function that has been previously suggested for prokaryotic S-layers (Sleytr et al., 2014). However, the underlying structure of the *P. arsenaticum* S-layer that mediates this resilience even at near boiling temperatures remains poorly understood. Previously, we have collected cellular tomograms of *P. arsenaticum* cells using cryoET. An S-layer is distinctly visible tiling the outer surface of the elongated rod-shaped cells and inclusions can be seen reaching from the S-layer to the cell membrane **(Fig. 6A)**. Initial subtomogram averages showed that the *P. arsenaticum* S-layer is made up of hexagonal subunits, with a funnel-like structure extending directly into the membrane (data not shown). Here, I aimed to resolve the structure of the *P. arsenaticum* S-layer to higher resolution, in order to understand the structural basis of the extraordinary stability and rigidity of this protein at high temperatures.

To this end, we performed cryoEM on isolated sheets of the *P. arsenaticum* S-layer. Micrographs of isolated S-layer sheets were collected perpendicular to the beam as well as at a 30° tilt and processed with a single particle analysis workflow. Top and bottom views of the S-layer protein were extracted from across the hexagonal lattice and side views along the edge of the sheets **(Fig 6B)**. After 2D classification, hexagonal top and bottom views were found, as well as tilted angles and side views of the S-layer protein **(Fig. 6C)**. Using the previously obtained cellular subtomogram average as an initial model, we reconstructed the *P. arsenaticum* S-layer in 3D. Further refinements with imposed C6 symmetry led us to achieve a 6.7 Å resolution structure of the S-layer protein **(Fig. 6D, S2B)**.

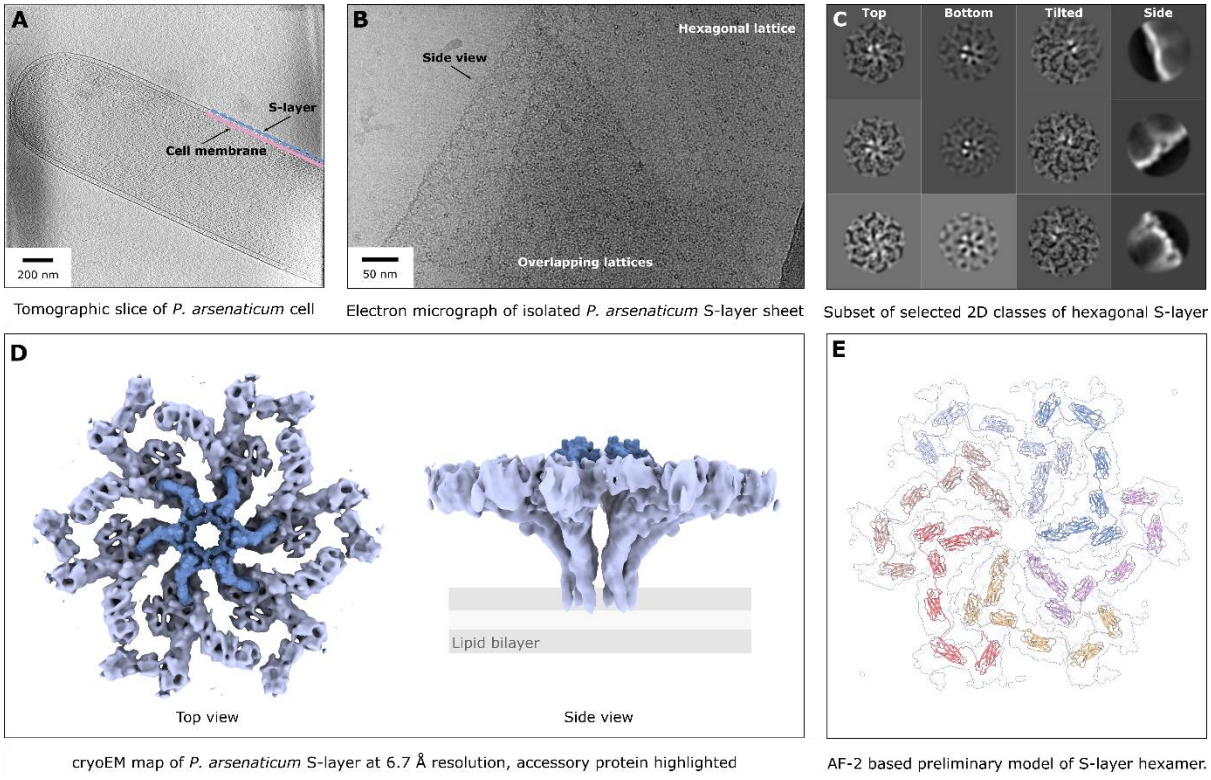


FIGURE 6 Structure of the *P. arsenaticum* Slayer protein. A) Tomographic slice of *P. arsenaticum* cell. Cell is fully encapsulated by regular S-layer lattice. B) cryoEM micrograph of isolated *P. arsenaticum* S-layer sheets. C) Subset of selected 2D class averages, displaying top, bottom, tilted and side views of the S-layer hexameric unit. C) Top and side views of cryoEM map of *P. arsenaticum* S-layer resolved to 6.7 Å resolution. S-layer protein is shown in light blue, accessory protein is shown in dark blue. Lipid bilayer indication based on bilayer seen in subtomogram average. D) Top view of AlphaFold2 based protein model of *P. arsenaticum* S-layer protein. Colours indicate separate hexameric subunits. cryoEM map is shown as outline.

The S-layer protein is a relatively open structure, arranged as a hexameric unit with a central pore. The central pore is organised in a funnel-like structure, with a diameter of 130 Å at the top and 70 Å at the membrane binding side. While the current resolution is insufficient to build the protein structure *de novo*, some secondary structure elements can already be recognized. On the cellular side, the S-layer is anchored into the cell membrane by an alpha helical domain. The S-layer lattice is formed by a set of distinct globular domains. This is reflected in sequence-based structure predictions by AlphaFold2 (Tunyasuvunakool et al., 2021), which suggest the presence of a hydrophobic transmembrane helix as well as seven conserved Ig-like beta sandwich domains. The predicted structure shows the Ig-like domains to be arranged in a beads-on-a-string fashion **(Fig 6E)**. Notably, when fitting the predicted protein structure into the hexagonally arranged cryoEM density, we find an additional density blocking the central pore of the S-layer hexamer that cannot be explained by the S-layer protein sequence. It thus seems the S-layer pore is blocked by an accessory protein of unknown

identity **(Fig. 6D)**. Local resolution mapping of the S-layer structure shows accessory protein is resolved more poorly than the S-layer lattice **(Fig. S2A)**. To discern whether this limited resolution is due in part to local conformational heterogeneity, we performed a focussed classification without additional alignment for the accessory protein. We find that although classification did not lead to significant improvement of overall resolution, distinct conformational states of the accessory protein could be found, ranging from fulling blocking the S-layer pore to a more open conformation **(Fig. 7)**.

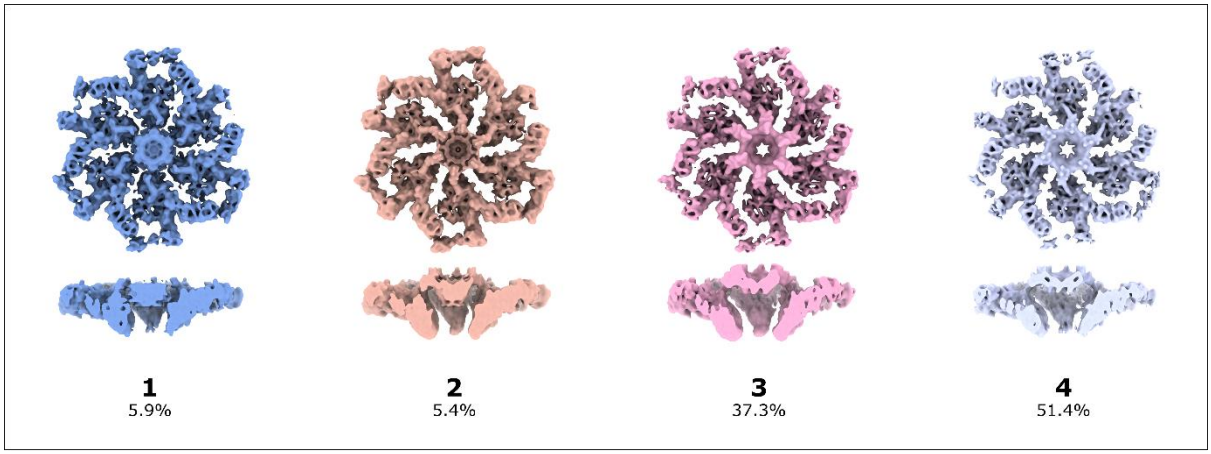


FIGURE 7 3D classification of *P. arsenaticum* S-layer reveals different pore states. Top and side views of S-layer shown in different states of pore blocking by S-layer accessory protein, shown from closed to open, left to right. Class occupation indicated as percentage of particles allocated to class.

We have thus resolved a medium resolution map of the *P. arsenaticum* S-layer protein and propose a preliminary model of this protein based on protein structure predictions. Currently, the limiting factor to the resolution of the cryoEM map remains unclear. The lattice appears relatively stable and exclusion of the more flexible accessory protein and transmembrane domains from alignment have not resulted in higher resolution reconstructions. Data quantity does not seem to be limiting, as the resolution does not decrease after subdivision of the dataset during classification. Potentially, improved per particle CTF corrections could benefit the resolution of this reconstruction, since CTF corrections were not successful for every micrograph. Additionally, we saw equivalent resolution in reconstructions processed with applied C6 symmetry as with C1 symmetry **(Fig. S2)**, suggesting there may be a case of pseudosymmetry. Finding of the correct symmetry and further efforts in the processing of this data will be required to achieve high resolution reconstructions. A higher resolution single particle cryoEM map will in turn allow us to build a definitive model of the *P. arsenaticum* S-layer protein in its hexameric state, and potentially find the identity of the unknown accessory protein by structure-based sequencing (Chojnowski et al., 2022).

The preliminary structure of the S-layer can provide some initial insights into the structure and function of the *P. arsenaticum* S-layer. Notably, the S-layer lattice domain likely consists predominantly of Ig-like domains. This reiterates the diverse suitability of Ig-like folds, in this case serving to be highly resilient to temperature induced motion and maintaining strong binding within the fold. Furthermore, the overall structure of the *P. arsenaticum* S-layer lattice is remarkably open, and Ig-like subunits are linked with short, disordered domains. This implies the S-layer lattice may not be entirely rigid, but allows some flexibility and movement within the lattice, perhaps to accommodate the high curvature of the thin cell body or to prevent lattice breaks by the increased molecular motion at near-boiling temperatures. However, overall rigidity in the lattice is still expected, supporting the role of the S-layer in cell shape maintenance. Furthermore, the pore-blocking accessory protein was found in several conformational states. Potentially, these states represent distinct and regulated states that allow *P. arsenaticum* to control pore permeability. Considering the relatively open structure of the S-layer protein, this pore is presumably molecule specific, as ions and small molecules will be able to freely pass through the S-layer. At the current resolution, the identity and function of the S-layer accessory protein remain unclear. To find whether different states of pore blocking are discrete or continuous and confirm that these states are representative of cellular protein conformations, future work will aim at improving the resolution of the cryoEM map and subtomogram average to study the S-layer and its accessory protein in their cellular environment. This may allow a more comprehensive understanding of the role of the *P. arsenaticum* S-layer in cell shape maintenance and regulation of cell permeability.

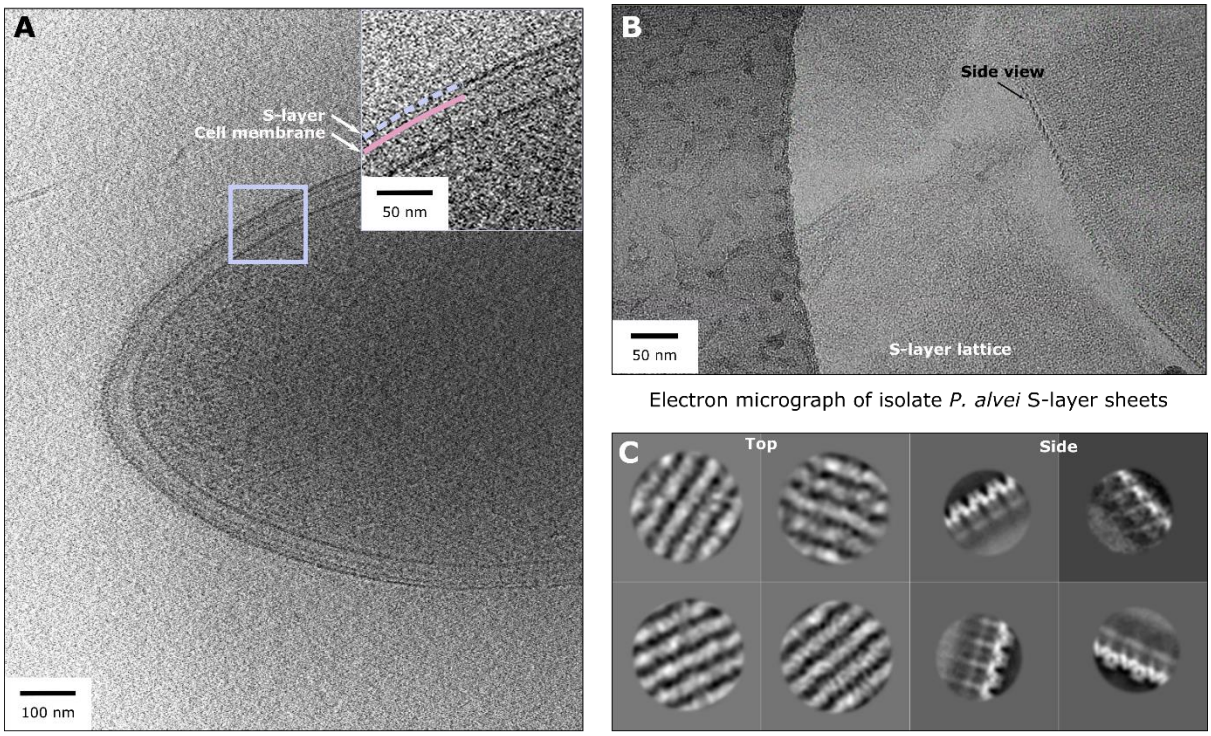
First Version – S3 | ***Paenibacillus alvei* has a p2 S-layer lattice**

In addition to their role as molecular sieves and in cell shape maintenance, S-layers are also known to be involved in pathogenicity. S-layers can serve as a binding site for phages and parasites (Sleytr et al., 2019), but can also play an active role in infection by pathogenic prokaryotes (Ravi & Fioravanti, 2021). The human pathogen *Clostridium difficile* for example, utilizes its S-layer for adherence to the host cell (Merrigan et al., 2013) and loss of the (Spa) S-layer has been shown to neutralize pathogenicity in *Bacillus anthracis* (Fioravanti et al., 2019; Missiakas & Schneewind, 2017). Studying S-layers involved in pathogenicity is thus of keen interest, as improving our molecular understanding of the pathology can highlight potential mechanisms for intervention or treatment. Here, we study the S-layer of the pathogenic Gram-negative bacterium *Paenibacillus alvei*, a secondary pathogen to honeybees, contributing to European Foulbrood disease that threatens honeybee colonies on a global scale (Forsgren, 2010).

The *P. alvei* cell surface is well characterised and a set of genetic tools has been established, allowing functional study of proteins *in vivo* (Zarschler et al., 2009). It has been suggested that the *Paenibacillus* S-layer plays a key role in its pathogenicity by mediating swarming behaviour and biofilm formation, as well as playing a role in host invasion (Janesch et al., 2013; Poppinga et al., 2012). In *P. alvei*, the S-layer is anchored to the secondary cell wall polymers (SCWPs) via an S-layer homology (SLH) domain (Messner et al., 1995), which has been well characterised. The identity of SCWPs has been described and the structure of the SLH-domain in its SCWP bound state has recently been solved by X-ray crystallography (Legg et al., 2022), however, the molecular structure of the S-layer lattice remains elusive. Elucidation of the S-layer structure will further complete our molecular understanding of the *P. alvei* cell surface, which can serve as a model for S-layer mediated pathogenicity and can be used to study S-layer targeting as a form of antibiotic or antivirulence treatment. Here, I take the first steps towards solving the structure of the *P. alvei* S-layer.

First, we used cryoET to produce tomograms of *P. alvei* cells. Cellular tomograms of *P. alvei* cells showed cells were tiled fully by an S-layer, at 9.4 ± 1 nm distance to the outer membrane **(Fig. 8A)**. We then used single particle cryoEM to resolve the structure of the *P. alvei* S-layer. Using a crude purification, we isolated S-layer sheets from *P. alvei* cells **(Fig 8B)**. Mass spectrometry analysis of selected bands from SDS-PAGE analysis confirmed the S-layer glycoprotein SpaA (Uniprot K5ABG3) is the major component of the sample by an order of magnitude excess **(Table S1)**. Initial processing of micrographs of S-layer sheets and 2D classification of top and side views suggest that the *P. alvei* S-layer has a *p2* symmetry rather than the *p6* symmetry most commonly seen in S-layer lattices **(Fig. 8C)**. Structure predictions by AlphaFold2 show with high confidence that SpaA is made up of 8 Ig-like

domains as well as its SLH domain. Further reconstruction of the S-layer protein requires an initial 3D model to align particle projection to. For this, a subtomogram average of the *P. alvei* S-layer must be obtained. Unfortunately, *P. alvei* cells proved to be too thick, and tomograms were not amenable to subtomogram averaging, halting our progress to 3D reconstruction of this protein.



Tomographic slice of P. *alvei* cell

Electron micrograph of isolate *P. alvei* S-layer sheets

Subset of 2D classes of p2 symmetrical S-layer

FIGURE 8 The *P. alvei* S-layer shows oblique p2 symmetry. A) Tomographic slice of *P. alvei cell.* Inset shows magnified view of the cell surface, S-layer and outer cell membrane are marked. B) cryoEM micrograph of isolate *P. alvei* S-layer sheets. C) Subset of obtained 2D class averages of *P. alvei* S-layer, top views and side views are shown.

We thus provide the initial work for reconstruction of the *P. alvei* S-layer structure. Cell thickness forms the limiting factor to this analysis, preventing the production of an adequate subtomogram average required for initial 3D reconstruction. To address this, cryoEM grids with *P. alvei* cells could be subjected to focussed ion beam milling to thin cells and create a sample amenable to high resolution cryoET (Wagner et al., 2020). Based on our initial 2D averages, it becomes apparent the *P. alvei* S-layer does not form a canonical hexagonal lattice, but rather a lower-symmetry lattice with *p2* symmetry. Non-hexagonal S-layers have been often reported. However, the case of *P. alvei*, the unconventional symmetry of the S-layer is of particular interest as it coincides with the presence of an SLH-domain. SLH-domains are highly conserved through evolution and display a trifold

symmetry, which is commonly achieved by bringing together SLH-regions from three subunits in a trimeric protein complex (Kern et al., 2011). Instead, the *P. alvei* SpaA contains three of such SLH-regions within a single S-layer subunit, which is visible in the crystal structure of the SpaA SLH-domain. A high-resolution structure of the S-layer will explain how the mismatch of two-fold symmetry of the S-layer lattice and the three-fold symmetric SLH domain are accommodated in the cell.

Results – S4 | **HPI and SlpA organise the *Deinococcus radiodurans* cell envelope**

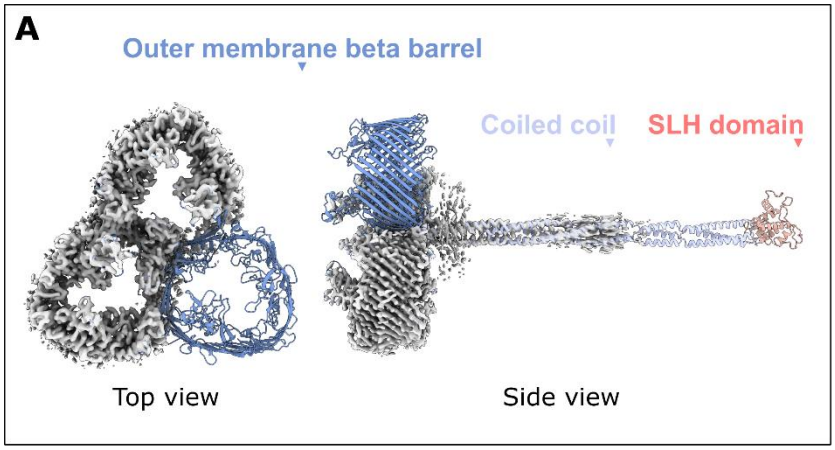
One of the most remarkable functions of S-layers, is their ability to support prokaryotic life in extreme conditions. *Deinococcus radiodurans* is an evolutionarily deep-branching bacterium with exceptional resilience to harsh environments (Davis et al., 1963). *D. radiodurans* can resist fluctuations in temperature and pressure as well as high doses of ionising UV and gamma radiation (Yuan et al., 2009). It has even been shown to be viable after several months in outer space (Ott et al., 2020). This remarkable resilience against ionising radiation is attributed mostly to the presence of extensive DNA repair machinery (Cox & Battista, 2005) and a unique and complex cell surface (Baumeister et al., 1981). The *D. radiodurans* cell surface stains Gram positive, due to a thick peptidoglycan layer, but upon closer inspection, electron micrographs show the cell surface is arranged as in Gram negative (diderm) bacteria, with an outer membrane covering the peptidoglycan and interstitial layer, and an S-layer with a dense carbohydrate coat covering coating the cell surface in patches (Sexton et al., 2022). Studies from 1970s and 1980s have first demonstrated the hyperstability of the hexagonal S-layer in *D. radiodurans*. S-layer sheets were shown to remain intact after isolation from cells at boiling temperatures and even after treatment with harsh denaturants such as sodium dodecyl sulfate (SDS) and urea (Baumeister et al., 1982).

Over the years, two conflicting models of the *D. radiodurans* S-layer have been proposed. Pioneering electron microscopy studies by Baumeister et al. (1986) and Rothfuss et al. (2006) suggest that the hexagonally packed interlayer (HPI) protein is the sole component of the S-layer. This model is supported by electron micrographs of *Deinococcus* deletion mutants, that show *hpi* deletion leads to disruption of the S-layer, while the remaining cell surface stays intact (Rothfuss et al., 2006). More recently, Farci et al. (2021) have proposed the *D. radiodurans* S-layer is in fact composed of three different proteins, together forming the hexagonal S-layer that can be found on the cell surface. An 11 Å resolution structure of a highly abundant cell surface protein SlpA was resolved by single particle cryoEM, demonstrating that it is a pin-like protein, with a trimeric head and an extended tail (Farci et al., 2021). The authors propose that SlpA, a T4P-like protein and an unknown dihedral complex tile together to form the S-layer. An initial average of isolated S-layer sheets was presented and annotated to accommodate these proteins, but the map was highly anisotropic and structural details were not well resolved (Farci et al., 2022). A high-resolution structure of the S-layer thus remains elusive. In order to settle the ongoing debate on the protein identity of the hyperstable *D. radiodurans* cell surface, we resolved the structure of HPI protein and SlpA protein using single particle cryoEM and study the *D. radiodurans* cell surface using cryoET.

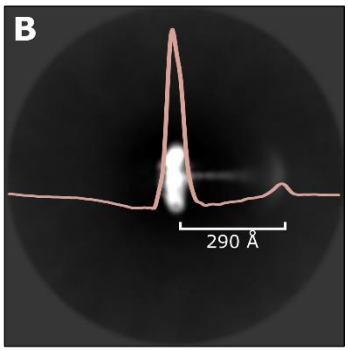
First, we studied the structure of SlpA protein aiming to uncover its role in the *D. radiodurans* cell envelope. We purified SlpA from *D. radiodurans* cell surface using detergent solubilization. We then applied a conventional single particle cryoEM workflow to obtain 2D classes and reconstructed the cryoEM map to a 2.9-3.3 Å resolution, which allowed us to build the atomic model for this protein **(Fig. 9A)**. SlpA is a homotrimer, with a large pore-like beta barrel head and an elongated coiled coil stalk protruding from the trimeric centre, that was poorly resolved in the cryoEM map due to its flexibility. The beta-barrel domains were embedded within a detergent micelle, indicating that SlpA is in fact a membrane protein. Furthermore, using AlphaFold-Multimer for protein prediction, we extended the model to include the residues 20-216, which were not well resolved in the cryoEM map. This revealed that a coiled coil of alpha helices extending from the trimeric centre of the beta barrels are joined by a short, disordered linker to an N-terminal SLH-domain **(Fig. 9A)**. It is thus apparent that the SlpA protein is an outer membrane protein and thus cannot constitute part of the S-layer, which is distinct from the outer membrane.

Next, we applied single particle cryoEM to solve the structure of the *D. radiodurans* S-layer. Using an protocol (Baumeister et al., 1981), S-layer sheets were isolated from the cell surface **(Fig. 9C)**. cryoEM single particle analysis of these sheets then allowed reconstruction of the S-layer structure to a 2.5 Å resolution and the building of the corresponding atomic model **(Fig. 9D)**. This unambiguously demonstrated the S-layer is formed by the HPI protein, without additional accessory proteins. The structure of HPI protein consists of a series of six Ig-like domains that form a hexagonal lattice. Notably, subunits of the hexamer are highly interconnected, reaching from one hexameric unit into the next **(Fig. 9D)**. This extensive interconnection between neighbouring hexameric units presumably has a significant role in the stability of the *D. radiodurans* S-layer.

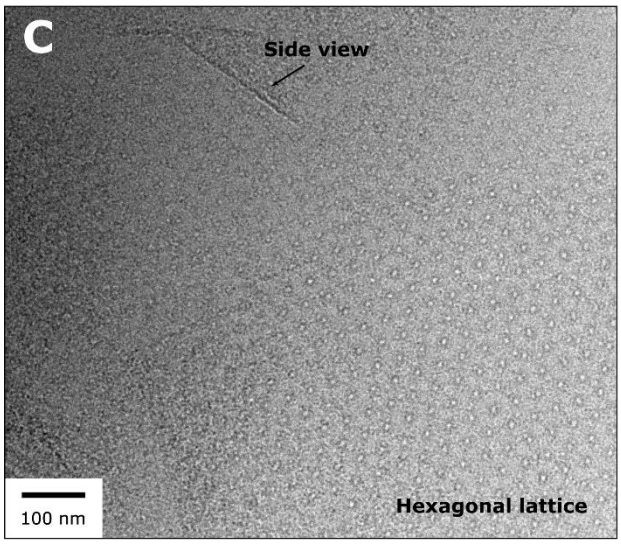
In order to understand the role of HPI and SlpA in the organisation of the cell surface, we used cryoET to study the native *D. radiodurans* cell periphery. Tomograms confirm the known organisation of the cell surface, consisting of an inner membrane, thick periplasmic layer, outer membrane and S-layer **(Fig. 10A)**. The peptidoglycan layer in the periplasm is seen as a darker density, with a fuzzy edge. By averaging the cell surface in the Z-direction of the tomogram, we define the edge of the peptidoglycan layer more clearly **(Fig. 10B)**. The tomograms and Z-average show that the distance from outer membrane to peptidoglycan (30 ± 3 nm) corresponds to the predicted length of the SlpA



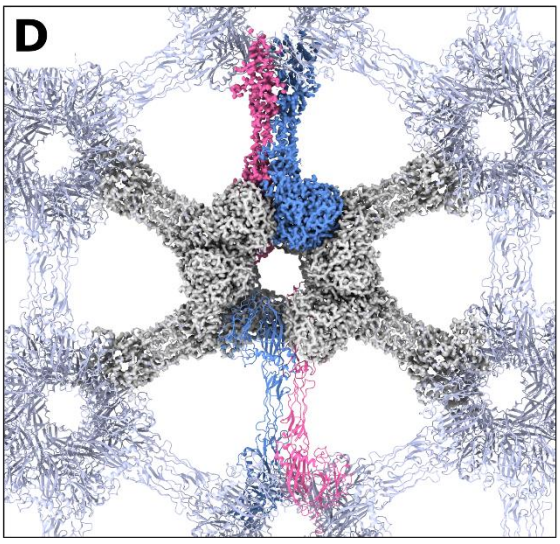
cryoEM map and model of SlpA



2D class of SlpA sideview



Electron micrograph of isolated HPI lattice



cryoEM map and model of HPI lattice

FIGURE 9 Structures of *D. radiodurans* cell surface proteins SlpA and HPI. A) Top view and side view of cryoEM map and atomic model of SlpA. CryoEM map is shown in grey, atomic model of the outer membrane beta barrel is shown blue. AlphaFold structure predictions of the coiled-coil domain and SLH-domain of SlpA are depicted in light blue and pink respectively. B) 2D class average of side view of SlpA protein. A profile plot is overlaid, peaks correspond to the density for outer membrane beta barrel and SLH-domain respectively. Measured distance between these peaks is 290 Å. C) Cryo-electron micrograph of isolated *D. radiodurans* S-layer lattice. D) cryoEM map and model of *D. radiodurans* S-layer protein HPI. cryoEM density is shown in grey, atomic model is shown in light blue. Two subunits are highlighted in blue and pink, symmetrically interconnecting adjoining hexameric centres.

coiled coil domain (28 nm) and the length of the blurred density corresponding to the coiled coil and SLH-domain in 2D class averages (~29 nm) **(Fig. 9B)**. Considering that the SlpA beta barrel must be embedded in a membrane, and SLH-domains are highly conserved and known to bind peptidoglycan and SCWPs (Blackler et al., 2018), we thus conclude SlpA is an outer membrane protein, with a linker domain extending towards the cell, embedded in the peptidoglycan with its SLH-domain.

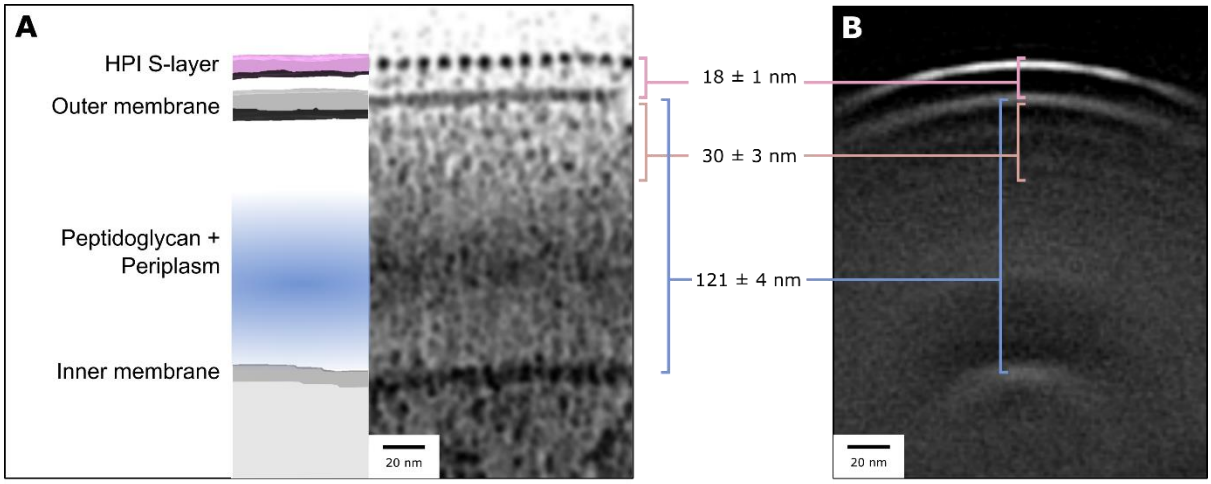


FIGURE 10 **Tomographic analysis of the *D. radiodurans* cell envelope.** A) Segmentation and tomographic slice of *D. radiodurans cell surface*. Density shown in black. B) Z-average of cell surface tomograms. Density shown in white.

This leads us to propose a renewed structure-based model for the cell surface organisation of *D. radiodurans* **(Fig. 11)**. In line with previous models by Baumeister et al. (1986) and Rothfuss et al. (2006), we demonstrate HPI to be the sole S-layer protein, forming the distinct hexagonal lattice. SlpA on the other hand, does not constitute part of the S-layer, but instead forms an outer membrane beta barrel with a long linker domain that anchors the outer membrane to the cell surface by binding peptidoglycan with its conserved SLH-domain. SlpA mutants have previously been shown to lead to major disruption of the cell surface (Rothfuss et al., 2006) and peptidoglycan binding by the SlpA SLH domain has been experimentally confirmed (Misra et al., 2015). It thus seems SlpA plays a critical role in cell surface integrity and may function directly in anchoring the outer membrane and with it the S-layer to the cell wall. Being an evolutionarily deep-branching bacterium, the key role of SlpA in tethering the outer membrane to the cell wall may indicate an ancestral role for PG binding outer membrane proteins in the evolution of monoderm and diderm bacteria.

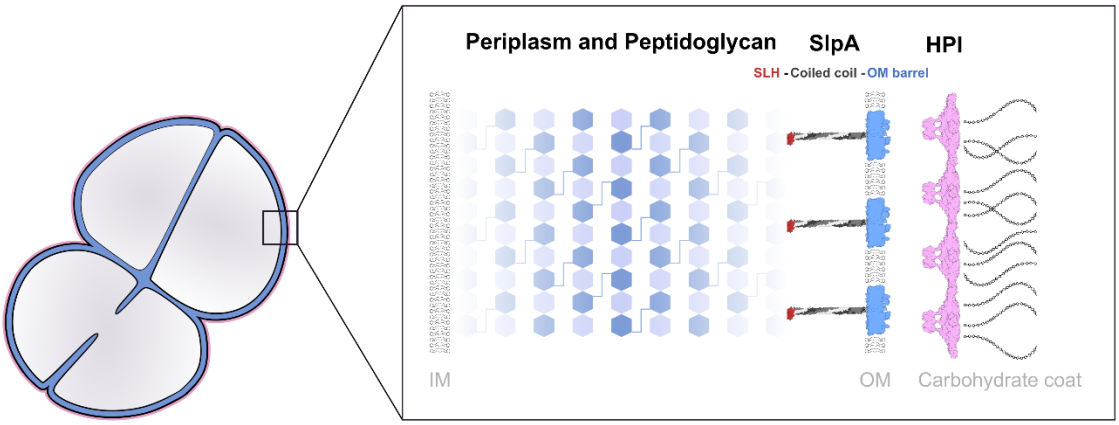


FIGURE 11 **Schematic representation of *D. radiodurans* cell surface organisation.** A thick peptidoglycan layer covers the inner membrane. SlpA binds the peptidoglycan with its SLH-domain and tethers the outer membrane to the cell wall. The carbohydrate coated HPI protein forms the S-layer and the outer boundary of the cell.

Discussion

Here, we have established a reproducible approach to studying prokaryotic S-layer structural biology that can be applied to a wide range of organisms, even when they come with biological challenges and limitations. Our results exhibit the power of state-of-the-art cryoEM and cryoET not only in finding the molecular architecture of S-layer proteins but also understanding them within a functional biological context.

S-layers in the face of environmental challenges

Survival in hostile conditions is dependent on extensive adaptation to the physical challenges of the environment. We have shown that the *N. maritimus* S-layer acts as a molecular sieve, allowing the cells to survive in nutrient deprived conditions. Notably, this coincides with a closed S-layer architecture, where only the central pore and some interfaces between subunits appear to allow free passage of molecules. The size of these pores suggest extensive regulation of molecular diffusion occurs at the S-layer level. The *N. maritimus* S-layer shows a remarkable level of structural similarity with the *H. volcanii* S-layer (von Kügelgen et al., 2021), both S-layers consisting of hexameric units with small pores, in a distinct cup-like shape. It thus appears this closed S-layer organisation is an adaptation to life in marine environments, where regulation of permeation to ions is critical to survival. *N. maritimus* environmental adaptivity is also seen in the changing cell size we observed in response to varying ammonium concentrations. This may reflect adaptation to fluctuating nutrient concentrations in the deep-sea dependent on seasonal changes and ocean currents.

The various intermediate states of opening and closing the S-layer pore seen in the *P. arsenaticum* structure suggest that here too, the S-layer plays a role in regulation of cell surface permeation. This is expected to be especially important in a high temperature environment, as increased molecular motion may lower energy thresholds that commonly limit diffusion of molecules in and out of the cell.

Aside from its role in regulating cell permeability, our results confirm that the architecture of S-layer is important for its role as a structural component to the cell. The *P. arsenaticum* S-layer, as well as the *D. radiodurans* S-layer must withstand severe physical stresses of heat and radiation respectively. Structural similarities between these Ig-like domain containing S-layers again underline similar adaptation to environmental factors, despite their evolutionary distance. Indeed, both S-layers show an open architecture with large pores, which may contribute to maintaining structural integrity of the S-layer under stress. A more open structure may allow the S-layer to accommodate tension and physical stress without leading to lattice breaks, thereby contributing to the maintenance of cell

shape. Additional structural strength in *D. radiodurans* is likely derived from the highly interconnected lattice structure. Our results thus show that S-layers structures reflect extensive adaptation of prokaryotes to their niche environments.

Structural themes in S-layer biology

Despite wide variation in sequences of (predicted) S-layer proteins (Bharat et al., 2021), with increasing number of recorded S-layer models, common structural themes are starting to rise to the surface. Here, we have studied four S-layer structures, three of which unambiguously show to have a hexagonal symmetry. At a medium resolution level, hexagonal symmetry can readily be identified, and a large number of S-layers has been demonstrated to be arranged as a hexagonal array over the last five decades (Pum et al., 2021). Hexagonal symmetry thus seems a shared solution to the need to tile a surface. It has been shown that introducing symmetry spontaneously drives the formation of supramolecular structures (Garcia-Seisdedos et al., 2017). This applies to a certain extent to lower order symmetries, as can be seen in the two-fold symmetry we see in initial 2D classes of the *P. alvei* S-layer and recorded cases of three-fold and four-fold symmetry S-layer architectures. However, higher order symmetry, as the hexagonal organisation seen in most S-layers, may simply be the most energy efficient evolutionary solution to the formation of a S-layer of protein.

At a closer look, the similarity in domain folds between S-layers of widely varying organisms is striking. Ig-like domains are ubiquitous in S-layer structures, across domains and in evolutionarily deep branching species (von Kügelgen et al., 2021). In all S-layers presented here, as well as in several previously reported S-layer structures, Ig-like beta-sandwich folds dominate the protein structure (Arbing et al., 2012; Baranova et al., 2012; Fioravanti et al., 2019; von Kügelgen et al., 2021). Sequence based fold predictions have also suggested that Ig-like domains are widespread in uncharacterised archaeal S-layer proteins (von Kügelgen et al., 2021). These Ig-like domains arise from highly variable protein sequences, and form multimeric assemblies with different organisations, but appear to have a similar structural organisation in archaea, Gram-positive and Gram-negative bacteria. This suggest Ig-like domains are highly stable in the face of various environmental stresses, and likely important in the conserved functions of S-layer proteins, including cell-cell recognition. Considering that several characterised S-layers contain Ig-like domains from phylogenetically diverse species, we can speculate that the Ig-like S-layer domain was perhaps involved in the evolution of more complex recognition systems seen in the eukaryotic immune system, where Ig folds are commonly found.

High-throughput S-layer structural studies and applications in biotechnology

We have demonstrated here that a cryoEM and cryoET based approach to structure determination of prokaryotic S-layer sheets is replicable in various organisms and across domains. Sample preparation methods and data analysis methods established here can be applied to S-layers from species that are difficult to study by conventional methods, due limited sample availability or due to a lack of genetic tools needed for overexpression or molecular tagging of proteins. With this, we provide a strong base for high-throughput S-layer structure determination in the future. This will allow a shift of focus towards the biological context and towards potential biotechnological applications.

Due to their straightforward isolation, self-assembling properties and high copy numbers, S-layers form very interesting candidates for applications in biotechnology. S-layers in well characterised and genetically tractable systems, such as *P. alvei* and *C. crescentus* can be utilised as a lattice for surface display of proteins and tags of interest. These systems can be applied for example to study surface protein interactions or to express enzymes for catalysis in an organised array. The thermostability and radiation resistance found in S-layers of extremophilic prokaryotes such as *P. arsenaticum* and *D. radiodurans* are of major interest for the development of biobased materials and medical applications. Several examples of biotechnological applications of S-layers have already been described (Misra et al., 2015; Raff et al., 2016; Sára et al., 2005; Schuster & Sleytr, 2021), and major developments are expected in this area in the near future.

Summary

Taken together, we have present molecular models at medium to high resolution of several cell surface molecules and thereby contributing to the growing diversity of S-layer structures and our understanding of prokaryotic cell surface biology. Our cryoEM and cryoET based methodology has been reproducible in various of organisms and can be applied to a wide range of prokaryotes in the future. We have shown S-layers must be highly adapted to the challenges and constraints of the environmental niche of the prokaryote. This is seen in the role of S-layers in regulating cell permeation and formation of highly resilient structures that maintain cell shape and protect against physical stresses. We begin to uncover common structural themes in S-layer biology such as a predisposition to hexagonal symmetry and the ubiquitous presence of Ig-like domains. With only a limited set of S-layer protein structures solved, this is only the tip of the iceberg. Future work focussing on expanding the record of S-layer structures and a focus on how they mediate environmental interaction may provide a greater understanding of the underlying principles of prokaryotic cell surface biology in the future.

Methods

Much of this work has been based on previous efforts from our lab or collaboration with external colleagues. For the sake of clarity and brevity, only methods I have personally applied within this project are listed here. For methods of remaining experiments, please refer to our published works. For further information, please contact SvD or TAMB.

CELL CULTURE

Nitrosopumilus maritimus (strain SCM1) cells were cultured in Modified Synthetic Crenarcheaota Medium (SCM) **(Table S3)** in the dark at 28 °C. Cell growth was monitored indirectly by evaluating NO_2^- concentration with a spectrophotometric Griess assay. NO_2^- concentrations were determined by addition of 100 µL Griess reagent **(Table S3)** to 100 µL culture and measurement of absorbance at 538 nm using a plate reader (SpectraMax M5, Molecular Devices). Cultures were serially diluted 1:1, 1:2, 1:4 and 1:8 (v/v) in SCM to ensure accurate measurement in the linear range of the Griess assay. Cultures were propagated by 5 (v-)% inoculation in SCM approximately every 7 days or when NO_2^- concentrations reached 300-500 µM at mid-log phase.

For cell growth assays of NO_2^- concentrations in 10 mL *N. maritimus* cultures were recorded with a Griess assay approximately every 24 hours for 14 days after 5 (v-)% inoculation. Cultures were grown in SCM containing 15 mM, 10 mM, 7.5 mM, 2.5 mM, 5 mM, 2.5 mM, 1 mM, 0.75 mM, 0.5 mM, 0.25 mM and 0.1 mM ammonium. Cultures were grown *in triplo* and the experiment was duplicated with inoculation from two independent cultures. Results were plotted using RStudio[1], data points were interpolated with local polynomial regression.

Paenibacillus alvei (strain CCM2051T) was cultured in Luria Bertani Broth (LB) **(Table S3)** at 30 °C and shaking at 180 rpm. Cultures were inoculated by 5 (v-)% from cells in late-log phase, when cultures reached an optical density at 600 nm (OD_{600}) of approximately 1.0.

SAMPLE PREPARATION FOR CRYO-EM

Protein preparation

For the preparation of *N. maritimus* SlaA sheets, wild type S-layer sheets were isolated from *N. maritimus* cells grown in SCM with 1 mM or 2.5 mM ammonium. Cultures were grown to mid-log, signified by an NO_2^- concentration of 300-450 µM, and spun down from 500 mL culture by centrifugation (15 000x g, 30 min, 4 °C). Cell pellets were resuspended in 1 mL SCM and harvested by centrifugation (16 100x g, 30 min, 4 °C). Cell pellets were then resuspended in 200 µL NSL Buffer **(Table**

[1] http://www.rstudio.com/

S3) and incubated on ice for 10 min. The cell suspension was then resuspended in 5.8 mL NSL Buffer and sonicated 5 times (2 sec, 10 kHz, 30 sec). S-layer sheets were harvested by ultracentrifugation (80 000x g, 1h, 4 °C). SlaA S-layers pellets were resuspended in 200 µL NSR Buffer **(Table S3)**. Samples were stored at 4 °C until use in cryoEM grid preparation.

For the crude purification of ribosomes, *N. maritimus* cells were grown to mid-log, signified by an NO_2^- concentration of 300-450 µM. Cells were harvested from 4 L culture by centrifugation (15 000x g, 30 min, 4 °C). Pellets were resuspended in 1.5 mL culture media and pelleted collectively (16 100x g, 30 min, 4 °C). Cell pellets were snap frozen and stored at -80 °C. Cells were then thawed on ice, resuspended in 6 mL NRL Buffer **(Table S3)** and lysed by homogenisation (2000 bar, 1.5 min, 4 °C). Cell debris was removed by centrifugation (16 100x g, 15 min, 4 °C). Ribosomes were then extracted from the supernatant fraction by ultracentrifugation (100 000x g, 1.5h, 4 °C). Supernatant was discarded and ribosome pellets were resuspended in 10 µL NRR Buffer **(Table S3)**. Ribosome samples were kept on ice and used for cryoEM grid preparation directly.

To extract *P. alvei* SpaA sheets, wild type S-layers were isolated from native *P. alvei* cell surfaces. *P. alvei* cells were grown to an OD_{600} of approximately 1.0 in 1 L LB. Cells were harvested by centrifugation (6 000x g, 30 min, 4 °C) and pellets were stored at -80 °C. Cell pellets were then thawed on ice, resuspended in 40 mL PSL Buffer **(Table S3)**, and incubated slowly rotating for 30 min at 4 °C. Cells were then further lysed by 5 rounds of sonication (10 kHz, 5 sec, 30 sec recovery). Cell debris was spun down (3 000x g, 15 min, 4 °C) and soluble components were collected by centrifugation of the supernatant fraction (35 000x g, 30 min, 4 °C). This pellet was resuspended in 1 mL PSW Buffer **(Table S3)**. SpaA S-layer sheets were separated from soluble components by centrifugation (16 100x g, 10 min, 4 °C). The S-layer pellet was washed again in 1 mL PSW Buffer and remaining cell debris was spun down by low-speed centrifugation (3 000x g, 30 min, 4 °C). SpaA sheets were harvested from the supernatant by a final round of centrifugation (16 100x g, 10 min, 4 °C) and resuspended in 50 µL PSW Buffer. Purified SpaA sheets were kept on ice and used for cryoEM grid preparation immediately. SDS-PAGE analysis was performed to check sample quality.

For the preparation of *D. radiodurans* HPI sheets, wild-type S-layer sheets were isolated from native cell surfaces as previously described (Baumeister et al., 1982). *D. radiodurans* cells were cultured in 1L TGY medium **(Table S3)** to mid-log phase, when the OD_{600} reached approximately 1.0. Cells were then harvested by centrifugation (5 000x g, 30 min 4 °C), resuspended in 50 mL TGY medium, and stored at -80 °C. Cell pellets were then thawed briefly on ice and then in a water bath at RT, and cells were harvested by centrifugation (4 000xg, 15 min, 25 °C). The cell pellet was resuspended in 40 mL DSE Buffer **(Table S3)** and incubated slowly rotating at RT for 2.5 hours. Stripped cells were

removed by centrifugation (3 000x g, 20 min, 25 °C). S-layers were harvested from the supernatant by centrifugation (35 000x g, 30 min, 25 °C). The S-layer pellet was washed in 1.5 mL DSW Buffer **(Table S3)** and pelleted (16 100xg, 15 min, 4 °C) 4 times. The final HPI pellet was then resuspended in 100 µL DSW Buffer and stored at 4 °C until further use in cryoEM grid preparation. SDS-PAGE analysis was performed to check sample quality.

SDS-PAGE analyses were performed with the NuPage system (Invitrogen). NuPage LDS Sample Buffer (4x, Invitrogen) and NuPage Sample Reducing Agent (10x, Invitrogen) were added to 10 µL sample according to manufacturer's protocol and incubated at 95 °C for 15 min. For HPI, an additional 10 µL aliquot of protein sample was incubated with 1 µL EDTA to separate protein fractions. Samples were then loaded onto a NuPage protein gel with Precision Plus Protein™ All Blue Prestained Protein Standards (Bio-Rad) and run in MES buffer (200V, 400 mA, 35 min). Gels were stained with Coomassie Brilliant Blue dye and stained for 30 min and destained in dH_2O for approximately 24 hours.

Protein identity of *P. alvei* SpaA was confirmed with peptide fingerprinting tandem mass spectrometry performed at the MRC LMB mass spectrometry facility. Protein bands of interest were excised from an SDS-PAGE gel and sent for MS-MS. Identified proteins were sorted and proteins below 70 kDa and contaminant proteins from other species were excluded.

Cell preparation

For the preparation of *N. maritimus* cells for cellular tomography, cultures were grown in SCM with 0.25 mM, 1 mM or 7.5 mM ammonium, after 5 (v-)% inoculation from 1 mM background cultures. 50 mL cultures were grown until mid-log, signified by an NO_2^- concentration of 300-450 µM, harvested by centrifugation (15 000x g, 30 min, 4 °C) and resuspended in 30 µL SCM. Cells were kept at RT and immediately taken for cryoEM grid preparation.

For the preparation of *P. alvei* cells for cellular tomography, 25 mL cultures were grown to mid-log phase signified by an OD_{600} of approximately 1.0. A 30 µL aliquot of culture was kept at RT and immediately taken for cryoEM grid preparation.

Grid preparation

For the preparation of cryoEM grids, 2,5 µL protein or cell sample was applied to freshly glow discharged Quantifoil grids (R2/2 Cu/Rhu 200 mesh), adsorbed for 60 sec and blotted 2-5 sec before plunge freezing into liquid ethane. Grid preparation was performed with a Vitrobot Mark IV (ThermoFisher), the chamber was maintained at 100% humidity and 4 °C or 10 °C for protein or cellular

samples respectively. 10 nm protein-A gold (CMC Utrecht) was added to cellular tomography samples (1:10) directly prior to grid preparation. Vitrified grids were clipped into autogrid rings and stored in liquid nitrogen.

CRYO ELECTRON MICROSCOPY AND TOMOGRAPHY

All grids were imaged at a 300 kV FEI Titan Krios TEM (ThermoFisher Scientific) with K3 summit direct electron detector (Gatan), equipped with a post-column energy filter (20 keV slit width). Single Particle data were collected with EPU software (ThermoFisher Scientific) using a sampling pixel size of 0.55 Å in counting super-resolution mode. Movies of *N. maritimus* ribosomes were collected with 3.5 s exposure at a total dose of 42.296 e^-/Å over 40 frames. Movies of *P. alvei* S-layer sheets were collected with 3.5 s exposure at a total dose of 42.269 e^-/Å over 40 frames. Cellular tomography was performed with SerialEM software at a pixel size of 3.55 Å (Mastronarde, 2005). Tilt series of ± 60° degrees tilt with a 2° tilt increment were collected of *N. maritimus* cells with a 1 s exposure time and 142.78 e^-/Å total dose using a bidirectional collection scheme. Medium magnification maps of grid squares were recorded with a pixel size of 79.34 Å for quantification of *N. maritimus* cell size.

DATA ANALYSIS

cryoEM single particle analysis of *N. maritimus ribosome*

Single particle analysis of the *N. maritimus* ribosome was performed in RELION4.0 (beta) (Kimanius et al., 2021). Movies collected over the course of 24 hours were grouped into optics groups using a k-means algorithm[2]. MotionCor2 was used for motion correction, dose weighting and Fourier cropping of imported movies (Zheng et al., 2017). CTFfind4 was used to perform CTF estimations on all micrographs (Rohou & Grigorieff, 2015). Based on CTF estimations, 5527 good micrographs were selected with a resolution better than 6 Å, further manual evaluation led to a selection of 4969 micrographs. From a subset of 200 micrographs 1112 ribosome particles were manually picked and used to train a TOPAZ neural network with standard RESNET8 architecture (Bepler et al., 2020). This neural network was then applied to all micrographs to pick 121877 particles which were extracted in a 126x128 pixel box (bin4, pixel size 4.368 Å). After two rounds of reference-free 2D-classification and manual class selection, classes of 70S ribosomes were used to train a new TOPAZ network and pick 359598 particles from all good micrographs. Particles were subjected to a final round of 2D classification, and classes depicting the ribosome were selected. Particles were then classified into 5

[2] https://github.com/DustinMorado/EPU_group_AFIS

3D classes by reference free 3D classification. 2/5 classes showing the 70S ribosome were selected for further auto refinement, resulting in a final set of 30703 particles. Particles were re-extracted in full sampling size (bin1, box size 512x512, pixel size 1.092 Å), and subjected to another 3D auto refinement with local searches, using the previously obtained map as a reference model. Micrographs were then subjected to anisotropic magnification correction, per micrograph astigmatism estimation, per particle defocus estimation and trefoil corrections. After another round of 3D auto refinement, micrographs were corrected for higher order aberrations and beam tilt. Particles were then subjected to 3D refinement and Bayesian polishing with default parameters. Postprocessing with a tight soft-edged mask resulted in a global resolution of 3.4 Å according to the gold standard Fourier Shell Correlation (FSC) criterium (Scheres, 2012). The full ribosome was subjected to a focussed multi-body refinement, with masks focussing refinement on the large and small subunit respectively. Additional refinement and postprocessing of these two maps yielded resolutions of 3.1 Å and 3.8 Å for the 50S and 30S subunits respectively.

cryoEM single particle analysis of *P. arsenaticum* S-layer

Single particle analysis of 2D sheets of the hexameric *P. arsenaticum* S-layer protein was performed in RELION4.0 (beta). Analyses of movies from regular and 30° tilted data collection were initially processed independently. Movies were clustered into optics groups using a k-means algorithm. Imported movies were then motion corrected, dose-weighted and Fourier cropped with MotionCor2 implemented in RELION4.0 (beta). CTF estimation of motion-corrected micrographs was performed using CTFfind4. Micrographs with resolution better than 7 Å were selected based on CTF estimations for further processing. First, the helical picking tool in RELION was used to manually pick side views along the sheet edges, and top and bottom views in line with the hexagonal lattice of the sheets. 4x binned particles were extracted in a 128x128 pixel box with a helical rise of 64-100 Å. Extracted particles were classified into 2D classes using RELION4.0 reference-free 2D classification. Classes representing top and bottom views of the hexameric centre were combined and used for reference-based picking on the full micrograph set. Picked hexameric centres were then extracted (bin4, pixel size 4.368 Å) and subjected to two rounds of 2D classification and class selection. In parallel, initial classes of sideview particles derived from manual picking were combined and to train TOPAZ integrated in RELION4.0 with a CONV127 neural network architecture. The trained network was then used to pick particles in the full micrograph set. Side views were obtained from the TOPAZ picked particle set by several rounds of reference-free 2D classification and selection. Selected top, bottom and side views picked were combined, and regular and tilted datasets were merged. Particles were then re-extracted (bin4, box size 128x128 pixels, pixel size 4.368 Å) and duplicate picking positions

were removed. Remaining particles were subjected to 3D auto refinement with C6 symmetry imposed, aligned to a previously obtained medium resolution subtomogram average of the native structure from cell surface tomograms which served as reference. 2D projections of this initial 3D model were then used for referenced based picking in all micrographs. Good classes from 2D-classification were selected. These particles were merged with previously obtained good particles and duplicate particles were removed, resulting in a final set of 296237 particles. All particles were extracted (bin4, box size 128x128, pixel size 4.368 Å) and subjected to a 3D auto refinement. Refined particles were re-extracted in bin2 (box size 256x256, pixel size 2.184 Å) and subjected to a second round of focused 3D auto refinement, with the previously obtained map as reference and a mask focused on the central, stable domains of the hexameric unit. These particles were re-extracted in full size (bin1, 512x512, 1.092 Å). Another focussed 3D refinement was performed, after which corrections for beam tilt and trefoil, anisotropic magnification and per particle defocus were performed sequentially. A final round of 3D auto refinement yielded and postprocessing with a soft-edged mask focussing on the central lattice and excluding flexible membrane binding domains resulted in a resolution of 6.7 Å according to the gold standard FSC criterion of 0.143. Classification for different pore states was performed on the final map with a tight puck-shaped soft-edged mask focussed on the central lattice forming domain.

cryoEM single particle analysis of *P. alvei* S-layer

Single particle analysis of *P. alvei* S-layer sheets was performed in RELION4.0 (beta). Movies were clustered into optics groups and, using MotionCor2, motion corrected, dose-weighted and Fourier cropped. CTFfind4 was used for CTF estimations. Micrographs with a resolution better than 6 Å were selected for analysis. Using the helical picking tool in RELION, particles were picked along the sheet edges to obtain side views, and along the regular lattice to obtain top and bottom views. After extraction with a helical rise of 50 Å and reference-free 2D classification, classes displaying either side views or top and bottom views were selected respectively. Side views were used to train a CONV127 TOPAZ network and pick particles in the full dataset. After 2D classification, 511181 particles were selected as side views. Selected top view classes from manual picking were used for reference-based automated particle picking on all micrographs. After reference-free 2D classification, 25045 particles were selected as top or bottom views.

Visualisation and analysis of cryoEM maps was performed using UCSF Chimera, UCSF ChimeraX and COOT (Emsley et al., 2010; Pettersen et al., 2004, 2021).

Protein structure prediction

Structure predictions based on nucleotide sequence were performed using AlphaFold2.1.1 (Tunyasuvunakool et al., 2021). Sequences were extracted from UniProt database, with identifier K5ABG3 for *P. alvei* SpaA and identifier A0A7L4PCH8 for the *P. arsenaticum* S-layer protein. AlphaFold2 was run for the monomeric form of these proteins, with default parameters and models with the highest confidence are reported. An initial model of hexameric conformation of the *P. arsenaticum* S-layer was manually built into the cryoEM density from its AlphaFold2 prediction using UCSF ChimeraX.

Analysis of cellular microscopy

Collected tilt series were aligned using gold fiducials and reconstructed into tomograms using weighted back projection and a SIRT-like filter in Etomo from the IMOD package (Kremer et al., 1996). For tomograms of *D. radiodurans* cell envelopes, tensor based membrane detection was performed with TomosegmemTV2020 (Martinez-Sanchez et al., 2014), and segmentations were refined and visualised using UCSF ChimeraX. Distances between *D. radiodurans* cell envelope components were determined in at several positions across the cell surface in tomograms of the *D. radiodurans* cell surface.

For the analysis of *N. maritimus* cell size, *N. maritimus* cells were measured in medium magnification maps of full grid squares along their long axis using 3Dmod from the IMOD package. Cell sizes were compared with a Welch two-sample t-test and plotted using RStudio.

Acknowledgements

I would like to sincerely thank Dr. Tanmay Bharat for the opportunity to work on these incredible projects, taking me into the group and lending me his knowledge and advice. I am also very thankful to Andriko von Kügelgen, for his great supervision and entrusting me with his projects. The work I could build on performed previously by Zo Ford and Lennart Pagani was indispensable, and I am very grateful for their efforts. Likewise, I would like to thank all lab members for their help, discussions, patience and for welcoming me warmly into the lab. I would like to acknowledge Vikram Alva for his insightful contributions and continuous collaboration. Finally, I would like to thank my friends at NOOC for their friendship and support, and for making Oxford feel like home.

Project contributions

*Contributions to this project were made by **Sofie van Dorst (SvD)**, Andriko von Kügelgen (AvK), Tanmay A.M. Bharat (TAMB), Zo Ford (ZF), Lennart Pagani (LP) and Vikram Alva (VA).*

N. maritimus S-layer:
Project design: AvK, TAMB
Cell growth assays: SvD
EM/ET sample preparation: AvK, SvD
EM data collection and processing: AvK, SvD
Model building: AvK, TAMB
ET data collection and processing: SvD, AvK

N. maritimus ribosome:
Project design: AvK, TAMB
EM/ET sample preparation: AvK, SvD, LP
EM data collection and processing: SvD, AvK, LP
ET data collection and processing: AvK, LP
Model building: ZF, LP, TAMB

P.arsenaticum S-layer
Project design: ZF, TB
EM/ET sample preparation: ZF, TB
EM data collection and processing: SvD, TB, ZF, AvK, JB
ET data collection and processing: ZF, TB
Initial model building: SvD

P. alvei S-layer
Project design: SvD, TAMB
EM/ET sample preparation: SvD
EM data collection and processing: SvD, AvK

D. radiodurans SlpA
Project design: AvK, TAMB
EM/ET sample preparation: AvK
EM data collection and processing: AvK
ET data collection and processing: AvK, SvD
Model building: AvK, TAMB

D. radiodurans HPI
Project design: AvK, TB
EM/ET sample preparation: AvK, SvD
EM data collection and processing: AvK
Model building: AvK, TAMB

Results presented in this thesis have contributed to the following scientific works:

Andriko von Kügelgen, Sofie van Dorst, Vikram Alva, and Tanmay A. M. Bharat. 'A Multi-Domain Connector Links the Outer Membrane and Cell Wall in Evolutionary Deep-Branching Bacteria'. *bioRxiv*, 18 May 2022. https://doi.org/10.1101/2022.05.18.492506 and *PNAS (in press)*

Andriko von Kügelgen, Sofie van Dorst, Keitaro Yamashita, Danielle Sexton, Elitza Tocheva, Garib Murshudov, Vikram Alva and Tanmay A. M. Bharat. 'Interdigitated immunoglobulin arrays form the hyperstable surface layer of *Deinococcus radiodurans* bacteria' *Under preparation.*

Supplementary data

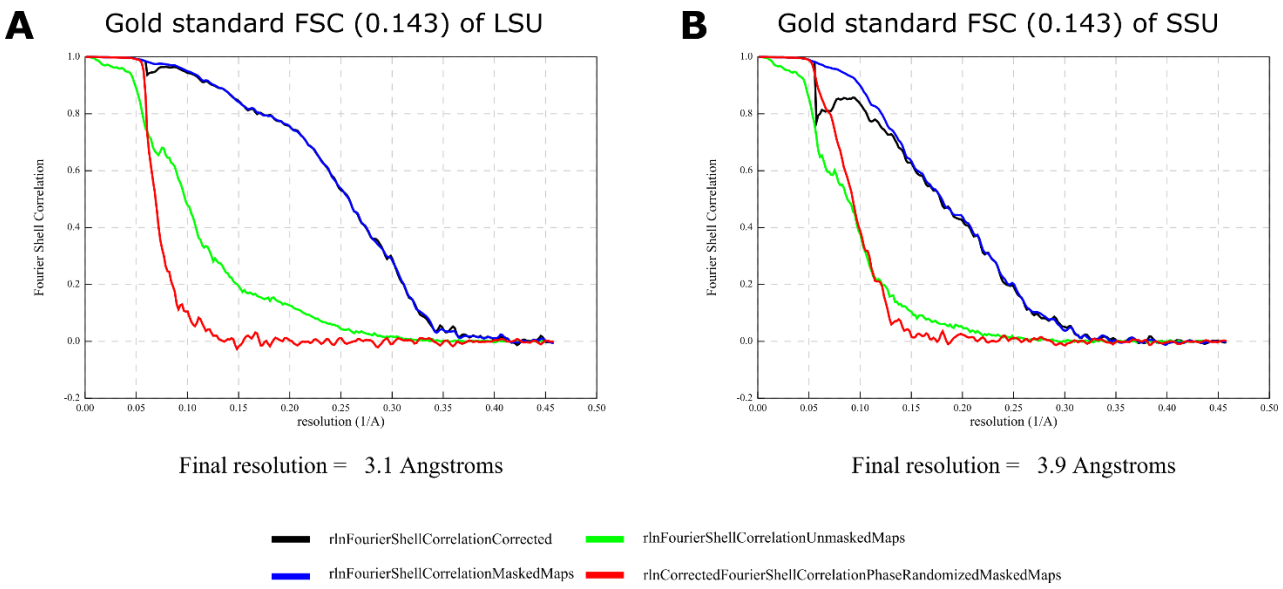


FIGURE S1 **Gold standard FSC (0.143) resolution determination for *N. maritimus* large ribosomal A) ribosomal subunit (LSU) and B) small ribosomal subunit (SSU).**

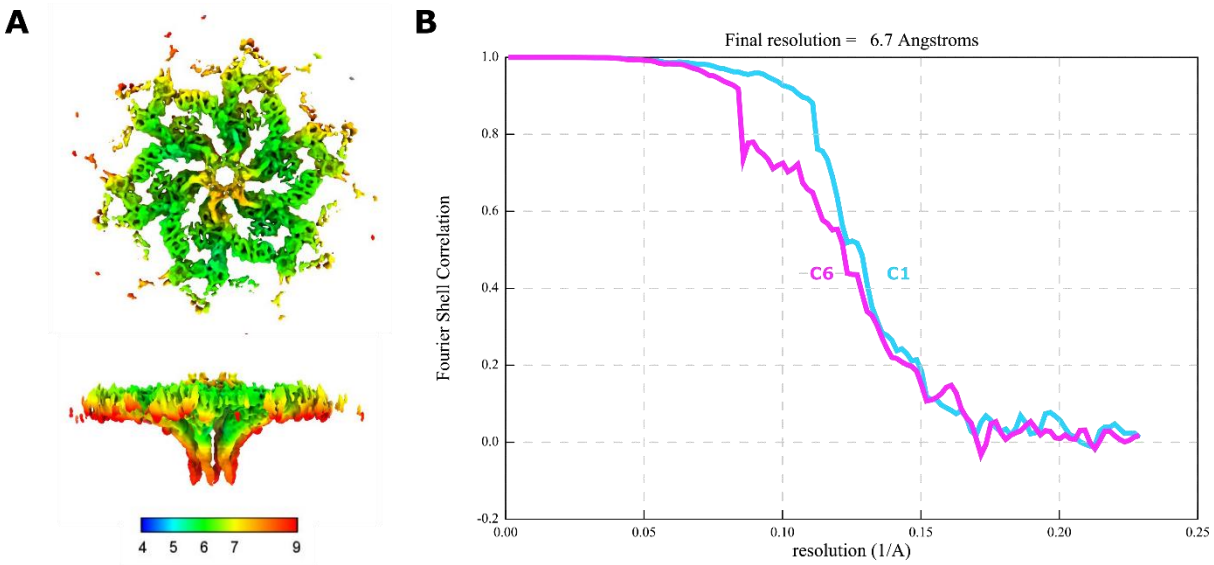


FIGURE S2 **Local resolution and FSC curves of *P. arsenaticum* S-layer map.** A) Local resolution plot shows reduced resolution at extended lattice, transmembrane domain and pore-blocking accessory protein. Scalebar shown in Å. B) Gold standard FSC curves (cut-off 0.143) of *P. arsenaticum* S-layer reconstructed with C1 or C6 symmetry applied. Both result in 6.7 Å resolution.

TABLE S1 **Exclusive unique peptide counts found by tandem mass spectrometry analysis of excised higher and lower molecular weight bands of *P. alvei* S-layer preparation.** S-layer protein SpaA is most abundant by tenfold.

Identified Proteins	Accession Number	Molecular Weight	Exclusive Unique peptide count - low MW	Exclusive Unique peptide count - high MW	Location in cell
Surface (S-) layer glycoprotein SpaA	K5ABG3_PAEAL	112 kDa	102	112	*not experimentally confirmed*
S-layer domain protein	K4ZPY4_PAEAL	127 kDa	26	3	*not experimentally confirmed*
DNA-directed RNA polymerase subunit beta	K4ZGE1_PAEAL	135 kDa	18	10	*not experimentally confirmed*
Minor extracellular protease vpr	K4ZSZ1_PAEAL	180 kDa	15	0	cytosol
Putative S-layer protein	K4ZL98_PAEAL	120 kDa	14	27	*not experimentally confirmed*
Nonribosomal peptide synthetase subunit	K4ZA11_PAEAL	123 kDa	14	0	*not experimentally confirmed*
Copper amine oxidase domain-containing protein	K4ZT64_PAEAL	112 kDa	13	0	*not experimentally confirmed*
DNA-directed RNA polymerase subunit beta	K4ZPR2_PAEAL	133 kDa	9	12	*not experimentally confirmed*
Wall-associated protein WapA	K4Z7Z4_PAEAL	199 kDa	9	0	cell surfae
NADH dehydrogenase-like protein YjlD	K5ABM8_PAEAL	70 kDa	6	0	cytosol
Swarming motility protein SwrC	K4Z8Q7_PAEAL	105 kDa	4	5	membrane
Putative family 2 glycosyltransferase	D6QW62_PAEAL	147 kDa	4	3	*not experimentally confirmed*
Polyketide synthase subunit	K4ZHS7_PAEAL	183 kDa	4	0	*not experimentally confirmed*
Glutamine--fructose-6-phosphate aminotransferase [isomerizing]	K4ZSD5_PAEAL	67 kDa	2	3	cytosol
Quinol oxidase subunit 1	K4ZQB0_PAEAL	74 kDa	2	2	membrane
Polyketide synthase subunit	K4ZCX0_PAEAL	78 kDa	2	1	*not experimentally confirmed*
Nonribosomal peptide synthetase subunit	K5A5Y9_PAEAL	343 kDa	2	0	*not experimentally confirmed*

TABLE S2 **Prokaryotic strains and origin.**

Species	Strain	Source
Nitrosopumilus maritimus	SCM1	Dr. Felix Elling
Paenibacillus alvei	CCM2051T	Prof. Dr. Christina Schäffer
Deinococcus radiodurans	R1	American Type Culture Collection
Pyrobaculum arsenaticum	PZ6	Dr. Mart Krupovic

TABLE 3 **Media, Buffers and Reagent composition.**

Media

Modiefied Synthetic Crenarchaeota Medium (SCM)	NaCl	444.9 mM
	$MgSO_4 \cdot 7H_2O$	20.3 mM
	$MgCl_2 \cdot 6H_2O$	24.6 mM
	$CaCl_2 \cdot 2H_2O$	10.2 mM
	KBr	0.84 mM
	Sodium Bicarbonate	1 mM
	HEPES/NaOH pH 7.0	10 mM
	KH_2PO_4	15 µM
	FeNaEDTA	7.5 µM
	NH_4Cl	1 mM
	Trace Elements	
	HCl-25%	100 µM
	H_3BO_3	0.5 µM
	$MnCl_2 \cdot 4H_2O$	0.5 µM
	$CoCl_2 \cdot 6H_2O$	0.8 µM
	$NiCl_2 \cdot 6H_2O$	0.1 µM
	$CuCl_2 \cdot 2H_2O$	0.01 µM
	$ZnSO_4 \cdot 7H_2O$	0.5 µM
	$Na_2MoO_4 \cdot 2H_2O$	0.15 µM
Luria Bertani (LB) Medium	Tryptone	1 % (w/v)
	Yeast Extract	0.5 % (w/v)
	NaCl	1 % (w/v)
Modified Tryptone-Glucose-Yeast Extract (TGY) Medium	Tryptone	0.5 % (w/v)
	Yeast Extract	0.25 % (w/v)
	Dextrose	0.1 % (w/v)
	$MnCl_2$	5 µM

Buffers

N. maritmus S-layer Lysis (NSL) Buffer	HEPES/NaOH pH 7.5	50 mM
	NaCl	500 mM
	MgCl2	50 mM
	CaCl2	10 mM
	CHAPS zwitterionic detergent	1 % (w/v)
	Benzonase	1 unit
	cOmplete™ Protease Inhibitor (Sigma Aldrich)	2 units
	NH4Cl	1 or 2.5 mM
N. maritmus S-layer Reconstitution (NRR) Buffer	HEPES/NaOH pH 7.5	50 mM
	NaCl	500 mM
	MgCl2	50 mM
	CaCl2	10 mM
	cOmplete™ Protease Inhibitor (Sigma Aldrich)	2 units
	NH4Cl	1 or 2.5 mM
P. alvei S-layer Lysis (PSL) Buffer	HEPES/NaOH pH 7.9	50 mM
	NaCl	500 mM
	MgCl2	25 mM
	CaCl2	5 mM
	CHAPS zwitterionic detergent	1 % (w/v)
	cOmplete ™ Protease Inhibitor (Sigma Aldrich)	1 unit
	Benzonase	1 unit
	DNAase I	50 µg/mL
P. alvei S-layer Wash (PSW) Buffer	HEPES/NaOH pH 7.9	50 mM
	NaCl	500 mM
	MgCl2	25 mM
	CaCl2	5 mM
	cOmplete™ Protease Inhibitor (Sigma Aldrich)	1 unit
D. radiodurans S-Layer Extraction (DSE) Buffer	Sodium Dodecyl Sulfate (SDS)	2 % (w/v)
	cOmplete™ Protease Inhibitor (Sigma Aldrich)	2 units
	Benzonase	1 unit

D. radiodurans S-layer Wash (DSW) Buffer	HEPES/NaOH pH 7.5	50 mM
	NaCl	150 mM
	$CaCl_2$	1 mM
N. maritimus Ribosome Lysis (NRL) Buffer	NH_4Cl	100 mM
	KOAc	100 mM
	Tris/HCl pH 8.5	20 mM
	$MgCl_2$	40 mM
	$CaCl_2$	0.1 mM
	Dithiothreitol (DTT)	2 mM
	cOmplete™ Protease Inhibitor (Sigma Aldrich)	2 units
	DNAase I	50 µg/mL
N. maritimus Ribosome Resuspesion (NRR) Buffer	NH_4Cl	100 mM
	KOAc	100 mM
	Tris/HCl pH 8.5	20 mM
	$MgCl_2$	40 mM
	$CaCl_2$	0.1 mM
	Dithiothreitol (DTT)	2 mM
	cOmplete™ Protease Inhibitor (Sigma Aldrich)	2 units

Reagents

Griess Reagent	N-(1-Naphthyl)ethylenediamine	5.1 mM
	Sulfanilamide	58.1 mM
	Orthophosphoric acid (85%)	11.4 % (v/v)

References

Arbing, M. A., Chan, S., Shin, A., Phan, T., Ahn, C. J., Rohlin, L., & Gunsalus, R. P. (2012). Structure of the surface layer of the methanogenic archaean Methanosarcina acetivorans. *Proceedings of the National Academy of Sciences*, *109*(29), 11812–11817. https://doi.org/10.1073/pnas.1120595109

Baranova, E., Fronzes, R., Garcia-Pino, A., Van Gerven, N., Papapostolou, D., Péhau-Arnaudet, G., Pardon, E., Steyaert, J., Howorka, S., & Remaut, H. (2012). SbsB structure and lattice reconstruction unveil Ca2+ triggered S-layer assembly. *Nature*, *487*(7405), 119–122. https://doi.org/10.1038/nature11155

Baumeister, W., Barth, M., Hegerl, R., Guckenberger, R., Hahn, M., & Saxton, W. O. (1986). Three-dimensional structure of the regular surface layer (HPI layer) of Deinococcus radiodurans. *Journal of Molecular Biology*, *187*(2), 241–250. https://doi.org/10.1016/0022-2836(86)90231-7

Baumeister, W., Karrenberg, F., Rachel, R., Engel, A., Ten Heggeler, B., & Saxton, W. O. (1982). The Major Cell Envelope Protein of Micrococcus radiodurans (R1). *European Journal of Biochemistry*, *125*(3), 535–544. https://doi.org/10.1111/j.1432-1033.1982.tb06715.x

Baumeister, W., Kübler, O., & Zingsheim, H. P. (1981). The structure of the cell envelope of Micrococcus radiodurans as revealed by metal shadowing and decoration. *Journal of Ultrastructure Research*, *75*(1), 60–71. https://doi.org/10.1016/s0022-5320(81)80100-1

Bepler, T., Kelley, K., Noble, A. J., & Berger, B. (2020). Topaz-Denoise: General deep denoising models for cryoEM and cryoET. *Nature Communications*, *11*(1), 5208. https://doi.org/10.1038/s41467-020-18952-1

Bharat, T. A. M., Kureisaite-Ciziene, D., Hardy, G. G., Yu, E. W., Devant, J. M., Hagen, W. J. H., Brun, Y. V., Briggs, J. A. G., & Löwe, J. (2017). Structure of the hexagonal surface layer on Caulobacter

crescentus cells. *Nature Microbiology*, *2*(7), 1–6. https://doi.org/10.1038/nmicrobiol.2017.59

Bharat, T. A. M., von Kügelgen, A., & Alva, V. (2021). Molecular Logic of Prokaryotic Surface Layer Structures. *Trends in Microbiology*, *29*(5), 405–415. https://doi.org/10.1016/j.tim.2020.09.009

Blackler, R. J., López-Guzmán, A., Hager, F. F., Janesch, B., Martinz, G., Gagnon, S. M. L., Haji-Ghassemi, O., Kosma, P., Messner, P., Schäffer, C., & Evans, S. V. (2018). Structural basis of cell wall anchoring by SLH domains in Paenibacillus alvei. *Nature Communications*, *9*(1), 1–11. https://doi.org/10.1038/s41467-018-05471-3

Chojnowski, G., Simpkin, A. J., Leonardo, D. A., Seifert-Davila, W., Vivas-Ruiz, D. E., Keegan, R. M., & Rigden, D. J. (2022). findMySequence: A neural-network-based approach for identification of unknown proteins in X-ray crystallography and cryo-EM. *IUCrJ*, *9*(1), Article 1. https://doi.org/10.1107/S2052252521011088

Cox, M. M., & Battista, J. R. (2005). Deinococcus radiodurans—The consummate survivor. *Nature Reviews Microbiology*, *3*(11), 882–892. https://doi.org/10.1038/nrmicro1264

Davis, N. S., Silverman, G. J., & Masurovsky, E. B. (1963). RADIATION-RESISTANT, PIGMENTED COCCUS ISOLATED FROM HADDOCK TISSUE1. *Journal of Bacteriology*, *86*(2), 294–298.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica. Section D, Biological Crystallography*, *66*(Pt 4), 486–501. https://doi.org/10.1107/S0907444910007493

Fagan, R. P., & Fairweather, N. F. (2014). Biogenesis and functions of bacterial S-layers. *Nature Reviews Microbiology*, *12*(3), 211–222. https://doi.org/10.1038/nrmicro3213

Farci, D., Haniewicz, P., & Piano, D. (2022). Mesoscale organization in the cell envelope of Deinococcus radiodurans. *BioRxiv*, 2022.01.29.478271. https://doi.org/10.1101/2022.01.29.478271

Farci, D., Kereïche, S., Pangeni, S., Haniewicz, P., Bodrenko, I. V., Ceccarelli, M., Winterhalter, M., & Piano, D. (2021). Structural analysis of the architecture and in situ localization of the main S-layer complex in Deinococcus radiodurans. *Structure*, *29*(11), 1279-1285.e3. https://doi.org/10.1016/j.str.2021.06.014

Fioravanti, A., Van Hauwermeiren, F., Van der Verren, S. E., Jonckheere, W., Goncalves, A., Pardon, E., Steyaert, J., De Greve, H., Lamkanfi, M., & Remaut, H. (2019). Structure of S-layer protein Sap reveals a mechanism for therapeutic intervention in anthrax. *Nature Microbiology*, *4*(11), 1805–1814. https://doi.org/10.1038/s41564-019-0499-1

Forsgren, E. (2010). European foulbrood in honey bees. *Journal of Invertebrate Pathology*, *103 Suppl 1*, S5-9. https://doi.org/10.1016/j.jip.2009.06.016

Garcia-Seisdedos, H., Empereur-Mot, C., Elad, N., & Levy, E. D. (2017). Proteins evolve on the edge of supramolecular self-assembly. *Nature*, *548*(7666), 244–247. https://doi.org/10.1038/nature23320

Huber, R., Sacher, M., Vollmann, A., Huber, H., & Rose, D. (2000). Respiration of arsenate and selenate by hyperthermophilic archaea. *Systematic and Applied Microbiology*, *23*(3), 305–314. https://doi.org/10.1016/S0723-2020(00)80058-2

Janesch, B., Koerdt, A., Messner, P., & Schäffer, C. (2013). The S-Layer Homology Domain-Containing Protein SlhA from Paenibacillus alvei CCM 2051T Is Important for Swarming and Biofilm Formation. *PLOS ONE*, *8*(9), e76566. https://doi.org/10.1371/journal.pone.0076566

Kern, J., Wilton, R., Zhang, R., Binkowski, T. A., Joachimiak, A., & Schneewind, O. (2011). Structure of Surface Layer Homology (SLH) Domains from Bacillus anthracis Surface Array Protein. *The Journal of Biological Chemistry*, *286*(29), 26042–26049. https://doi.org/10.1074/jbc.M111.248070

Kimanius, D., Dong, L., Sharov, G., Nakane, T., & Scheres, S. H. W. (2021). New tools for automated cryo-EM single-particle analysis in RELION-4.0. *The Biochemical Journal*, *478*(24), 4169–4185. https://doi.org/10.1042/BCJ20210708

Könneke, M., Bernhard, A. E., de la Torre, J. R., Walker, C. B., Waterbury, J. B., & Stahl, D. A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature*, *437*(7058), 543–546. https://doi.org/10.1038/nature03911

Kremer, J. R., Mastronarde, D. N., & McIntosh, J. R. (1996). Computer Visualization of Three-Dimensional Image Data Using IMOD. *Journal of Structural Biology*, *116*(1), 71–76. https://doi.org/10.1006/jsbi.1996.0013

Lanzoni-Mangutchi, P., Banerji, O., Wilson, J., Barwinska-Sendra, A., Kirk, J. A., Vaz, F., O'Beirne, S., Baslé, A., El Omari, K., Wagner, A., Fairweather, N. F., Douce, G. R., Bullough, P. A., Fagan, R. P., & Salgado, P. S. (2022). Structure and assembly of the S-layer in C. difficile. *Nature Communications*, *13*(1), 970. https://doi.org/10.1038/s41467-022-28196-w

Legg, M. S. G., Hager-Mair, F. F., Krauter, S., Gagnon, S. M. L., Lòpez-Guzmán, A., Lim, C., Blaukopf, M., Kosma, P., Schäffer, C., & Evans, S. V. (2022). The S-layer homology domains of Paenibacillus alvei surface protein SpaA bind to cell wall polysaccharide through the terminal monosaccharide residue. *Journal of Biological Chemistry*, *298*(4), 101745. https://doi.org/10.1016/j.jbc.2022.101745

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G. W., Prosser, J. I., Schuster, S. C., & Schleper, C. (2006). Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature*, *442*(7104), 806–809. https://doi.org/10.1038/nature04983

Li, P.-N., Herrmann, J., Tolar, B. B., Poitevin, F., Ramdasi, R., Bargar, J. R., Stahl, D. A., Jensen, G. J., Francis, C. A., Wakatsuki, S., & van den Bedem, H. (2018). Nutrient transport suggests an evolutionary basis for charged archaeal surface layer proteins. *The ISME Journal*, *12*(10), 2389–2402. https://doi.org/10.1038/s41396-018-0191-0

Martinez-Sanchez, A., Garcia, I., Asano, S., Lucic, V., & Fernandez, J.-J. (2014). Robust membrane detection based on tensor voting for electron tomography. *Journal of Structural Biology*, *186*(1), 49–61. https://doi.org/10.1016/j.jsb.2014.02.015

Mastronarde, D. N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology*, *152*(1), 36–51. https://doi.org/10.1016/j.jsb.2005.07.007

Merrigan, M. M., Venugopal, A., Roxas, J. L., Anwar, F., Mallozzi, M. J., Roxas, B. A. P., Gerding, D. N., Viswanathan, V. K., & Vedantam, G. (2013). Surface-Layer Protein A (SlpA) Is a Major Contributor to Host-Cell Adherence of Clostridium difficile. *PLOS ONE*, *8*(11), e78404. https://doi.org/10.1371/journal.pone.0078404

Messner, P., Christian, R., Neuninger, C., & Schulz, G. (1995). Similarity of 'core' structures in two different glycans of tyrosine-linked eubacterial S-layer glycoproteins. *Journal of Bacteriology*, *177*(8), 2188–2193. https://doi.org/10.1128/jb.177.8.2188-2193.1995

Messner, P., & Sleytr, U. B. (1988). Separation and purification of S-Iayers from Gram-positive and Gram-negative bacteria. In *Bacterial cell surface techniques*. J. Wiley & Sons, Chichester.

Misra, C. S., Basu, B., & Apte, S. K. (2015). Surface (S)-layer proteins of Deinococcus radiodurans and their utility as vehicles for surface localization of functional proteins. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, *1848*(12), 3181–3187. https://doi.org/10.1016/j.bbamem.2015.09.021

Missiakas, D., & Schneewind, O. (2017). Assembly and Function of the Bacillus anthracis S-Layer. *Annual Review of Microbiology*, *71*, 79–98. https://doi.org/10.1146/annurev-micro-090816-093512

Ott, E., Kawaguchi, Y., Kölbl, D., Rabbow, E., Rettberg, P., Mora, M., Moissl-Eichinger, C., Weckwerth, W., Yamagishi, A., & Milojevic, T. (2020). Molecular repertoire of Deinococcus radiodurans after 1 year of exposure outside the International Space Station within the Tanpopo mission. *Microbiome*, *8*, 150. https://doi.org/10.1186/s40168-020-00927-5

Park, B.-J., Park, S.-J., Yoon, D.-N., Schouten, S., Sinninghe Damsté, J. S., & Rhee, S.-K. (2010). Cultivation of Autotrophic Ammonia-Oxidizing Archaea from Marine Sediments in Coculture

with Sulfur-Oxidizing Bacteria. *Applied and Environmental Microbiology*, *76*(22), 7575–7587. https://doi.org/10.1128/AEM.01478-10

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science: A Publication of the Protein Society*, *30*(1), 70–82. https://doi.org/10.1002/pro.3943

Phipps, B. M., Huber, R., & Baumeister, W. (1991). The cell envelope of the hyperthermophilic archaebacterium Pyrobaculum organotrphum consists of two regularly arrayed protein layers: Three-dimensional structure of the outer layer. *Molecular Microbiology*, *5*(2), 253–265. https://doi.org/10.1111/j.1365-2958.1991.tb02106.x

Poppinga, L., Janesch, B., Fünfhaus, A., Sekot, G., Garcia-Gonzalez, E., Hertlein, G., Hedtke, K., Schäffer, C., & Genersch, E. (2012). Identification and Functional Analysis of the S-Layer Protein SplA of Paenibacillus larvae, the Causative Agent of American Foulbrood of Honey Bees. *PLOS Pathogens*, *8*(5), e1002716. https://doi.org/10.1371/journal.ppat.1002716

Pum, D., Breitwieser, A., & Sleytr, U. B. (2021). Patterns in Nature—S-Layer Lattices of Bacterial and Archaeal Cells. *Crystals*, *11*(8), 869. https://doi.org/10.3390/cryst11080869

Raff, J., Matys, S., Suhr, M., Vogel, M., Günther, T., & Pollmann, K. (2016). S-Layer-Based Nanocomposites for Industrial Applications. *Advances in Experimental Medicine and Biology*, *940*, 245–279. https://doi.org/10.1007/978-3-319-39196-0_11

Ravi, J., & Fioravanti, A. (2021). S-layers: The Proteinaceous Multifunctional Armors of Gram-Positive Pathogens. *Frontiers in Microbiology*, *12*. https://www.frontiersin.org/articles/10.3389/fmicb.2021.663468

Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, *192*(2), 216–221. https://doi.org/10.1016/j.jsb.2015.08.008

Rothfuss, H., Lara, J. C., Schmid, A. K., & Lidstrom, M. E. (2006). Involvement of the S-layer proteins Hpi and SlpA in the maintenance of cell envelope integrity in Deinococcus radiodurans R1. *Microbiology (Reading, England)*, *152*(Pt 9), 2779–2787. https://doi.org/10.1099/mic.0.28971-0

Sára, M., Pum, D., Schuster, B., & Sleytr, U. B. (2005). S-layers as patterning elements for application in nanobiotechnology. *Journal of Nanoscience and Nanotechnology*, *5*(12), 1939–1953. https://doi.org/10.1166/jnn.2005.502

Scheres, S. H. W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

Schuster, B., & Sleytr, U. B. (2021). S-Layer Ultrafiltration Membranes. *Membranes*, *11*(4), 275. https://doi.org/10.3390/membranes11040275

Sexton, D. L., Burgold, S., Schertel, A., & Tocheva, E. I. (2022). Super-resolution confocal cryo-CLEM with cryo-FIB milling for in situ imaging of Deinococcus radiodurans. *Current Research in Structural Biology*, *4*, 1–9. https://doi.org/10.1016/j.crstbi.2021.12.001

Sleytr, U. B., Breitwieser, A., & Pum, D. (2019). Crystalline Cell Surface Layers (S-Layers)☆. In T. M. Schmidt (Ed.), *Encyclopedia of Microbiology (Fourth Edition)* (pp. 783–792). Academic Press. https://doi.org/10.1016/B978-0-12-801238-3.11109-2

Sleytr, U. B., Schuster, B., Egelseer, E.-M., & Pum, D. (2014). S-layers: Principles and applications. *FEMS Microbiology Reviews*, *38*(5), 823–864. https://doi.org/10.1111/1574-6976.12063

Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel, A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S. A. A., … Hassabis, D. (2021). Highly accurate

protein structure prediction for the human proteome. *Nature*, *596*(7873), 590–596. https://doi.org/10.1038/s41586-021-03828-1

Turk, M., & Baumeister, W. (2020). The promise and the challenges of cryo-electron tomography. *FEBS Letters*, *594*(20), 3243–3261. https://doi.org/10.1002/1873-3468.13948

von Kügelgen, A., Alva, V., & Bharat, T. A. M. (2021). Complete atomic structure of a native archaeal cell surface. *Cell Reports*, *37*(8), 110052. https://doi.org/10.1016/j.celrep.2021.110052

von Kügelgen, A., Tang, H., Hardy, G. G., Kureisaite-Ciziene, D., Brun, Y. V., Stansfeld, P. J., Robinson, C. V., & Bharat, T. A. M. (2020). In Situ Structure of an Intact Lipopolysaccharide-Bound Bacterial Surface Layer. *Cell*, *180*(2), 348-358.e15. https://doi.org/10.1016/j.cell.2019.12.006

Wagner, F. R., Watanabe, R., Schampers, R., Singh, D., Persoon, H., Schaffer, M., Fruhstorfer, P., Plitzko, J., & Villa, E. (2020). Preparing samples from whole cells using focused-ion-beam milling for cryo-electron tomography. *Nature Protocols*, *15*(6), 2041–2070. https://doi.org/10.1038/s41596-020-0320-x

Wuchter, C., Abbas, B., Coolen, M. J. L., Herfort, L., van Bleijswijk, J., Timmers, P., Strous, M., Teira, E., Herndl, G. J., Middelburg, J. J., Schouten, S., & Sinninghe Damsté, J. S. (2006). Archaeal nitrification in the ocean. *Proceedings of the National Academy of Sciences*, *103*(33), 12317–12322. https://doi.org/10.1073/pnas.0600756103

Yuan, M., Zhang, W., Dai, S., Wu, J., Wang, Y., Tao, T., Chen, M., & Lin, M. 2009. (2009). Deinococcus gobiensis sp. Nov., an extremely radiation-resistant bacterium. *International Journal of Systematic and Evolutionary Microbiology*, *59*(6), 1513–1517. https://doi.org/10.1099/ijs.0.004523-0

Zarschler, K., Janesch, B., Zayni, S., Schäffer, C., & Messner, P. (2009). Construction of a Gene Knockout System for Application in Paenibacillus alvei CCM 2051T, Exemplified by the S-Layer Glycan Biosynthesis Initiation Enzyme WsfP. *Applied and Environmental Microbiology*, *75*(10), 3077–3085. https://doi.org/10.1128/AEM.00087-09

Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y., & Agard, D. A. (2017). MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods*, *14*(4), 331–332. https://doi.org/10.1038/nmeth.4193