

1 Micrarchaeota are covered by a proteinaceous S-Layer

2 Sabrina Gfrerer [1], Dennis Winkler [1], Julia Novion Ducassou [2], Yohann Couté [2],
3 Reinhard Rachel [3] and Johannes Gescher [1,4]

4

5 [1] Department of Applied Biology, Institute for Applied Biosciences, Karlsruhe Institute of Technology,
6 Karlsruhe, Germany

7 [2] Univ. Grenoble Alpes, CEA, Inserm, BIG-BGE, Grenoble, France

8 [3] Center for Electron Microscopy, University of Regensburg, Regensburg, Germany

9 [4] Institute for Biological Interfaces, Karlsruhe Institute of Technology (KIT), Eggenstein-
10 Leopoldshafen, Germany

11

12 **Abstract**

13 In previous publications, it was hypothesized that Micrarchaeota cells are covered by two
14 individual membrane systems. This study proves that at least the recently cultivated
15 “*Candidatus* Micrarchaeum harzensis A_DKE” possesses an S-layer covering its cytoplasmic
16 membrane. The potential S-layer protein was found to be among the proteins with the highest
17 abundance in A_DKE and *in silico* characterization of its primary structure indicated
18 homologies to other known S-layer proteins. Homologs of this protein were found in other
19 Micrarchaeota genomes, which raises the question, whether the ability to form an S-layer is a
20 common trait within this phylum. The S-layer protein seems to be glycosylated and the
21 Micrarchaeum expresses genes for N-glycosylation under cultivation conditions, despite not
22 being able to synthesize carbohydrates. Electron micrographs of freeze-etched samples of a
23 previously described co-culture, containing Micrarchaeum A_DKE and a Thermoplasmatales
24 member as its host organism, verified the hypothesis of an S-layer on the surface of A_DKE.
25 Both organisms are clearly distinguishable by cell size, shape and surface structure.

26 Introduction

27 Archaea exhibit a variety of cell surfaces. The most common cell surface type is a
28 proteinaceous surface layer (S-layer), which is often glycosylated (Sumper *et al.*, 1990; König
29 *et al.*, 2007; Albers and Meyer, 2011). S-layers consist of one or two proteins, which can self-
30 assemble and form a 2D paracrystalline layer, spanning the whole cell (Baumeister and
31 Lembcke, 1992; Veith *et al.*, 2009; among others). The S-layer is often the only cell wall
32 component, whereas in some methanogenic species it occurs in combination with an
33 additional layer of pseudomurein or methanochondroitin. The proteins forming this layer are
34 directly or indirectly anchored to the underlying cytoplasmic membrane (König *et al.*, 2007).
35 The lattice symmetry is either oblique (p1, p2), square (p4) or hexagonal (p3, p6), depending
36 on the number of identical S-layer proteins, which form one morphological subunit (Sára and
37 Sleytr, 1996; Sleytr *et al.*, 2014). The fine structure of the protein subunits, arranged on the
38 lattice, results in regularly spaced pores, which can be observed applying electron microscopic
39 techniques (Taylor *et al.*, 1982; Baumeister *et al.*, 1988; Engelhardt, 1988). After the first
40 discovery of a glycosylated S-layer protein in *Halobacterium salinarum* (Mescher *et al.*, 1974;
41 Sumper *et al.*, 1990), more and more glycosylated S-layers were discovered in Archaea and
42 Gram-positive Bacteria (Messner and Sleytr, 1991). The comparison of several S-layer proteins
43 in different Archaea revealed in some cases a poorly conserved amino acid sequence, but
44 often no sequence similarity at all; in contrast, the lattice type is often shared by prokaryotic
45 families or genera (König *et al.*, 2007).

46 S-layers are found – or at least postulated to be encoded in the genomes – in all four archaeal
47 superphyla, including the DPANN superphylum. Besides the eponymous phyla (Diapherotrites,
48 Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, and Nanohaloarchaeota; Rinke *et al.*,
49 2013), this superphylum comprises organisms of the Woese- and Pacearchaeota (Castelle *et al.*,
50 2015), Huberarchaeota (Probst *et al.*, 2018), Micrarchaeota (Baker *et al.*, 2010),
51 Altiarchaeota (Probst *et al.*, 2014), and Undinarchaeota (Dombrowski *et al.*, 2020), as well as
52 several so far undefined phyla (Castelle and Banfield, 2018; Dombrowski *et al.*, 2019). To date,
53 the obligate symbiont *Nanoarchaeum equitans* is the only known DPANN member with an S-
54 layer (Huber *et al.*, 2002). While the cell surface type of other DPANN members is still
55 unknown, Micrarchaeota, as well as Altiarchaeota were postulated to have an inner and outer
56 membrane (Comolli *et al.*, 2009; Baker *et al.*, 2010; Probst *et al.*, 2014). Micrarchaeota are

57 acidophilic organisms, occurring in various habitats around the world (Chen *et al.*, 2017). It is
58 speculated, that these organisms are dependent on other microorganisms, because of their
59 reduced genomes and lack of several essential genes or even metabolic pathways (Baker *et*
60 *al.*, 2010). In natural habitats they are often associated with members of the
61 Thermoplasmatales (Comolli and Banfield, 2014). In fact, the reduced genome in combination
62 with a small cell size are common characteristics of members of the DPANN superphylum,
63 which is postulated to contain several archaeal symbionts. Just a few DPANN-members were
64 cultivated under laboratory conditions so far (Huber *et al.*, 2002; Wurch *et al.*, 2016; Golyshina
65 *et al.*, 2017; St. John *et al.*, 2018; Hamm *et al.*, 2019). Recently, we described the cultivation
66 of the Micrarchaeota member “*Candidatus* Micrarchaeum harzensis A_DKE” in co-culture
67 with “*Candidatus* Scheffleriplasma hospitalis B_DKE”, a member of the Thermoplasmatales,
68 as potential host (Krause *et al.*, n.d.).

69 During the analysis of the above-mentioned co-culture of A_DKE and B_DKE, we were able to
70 detect an S-layer protein encoded in the closed genome of A_DKE (SLP_{Mh}). Transcriptomic and
71 proteomic data revealed that the gene is highly expressed and the corresponding protein
72 among the most abundant polypeptides in A_DKE. Finally, electron micrographs of freeze-
73 etched samples or of frozen, fully hydrated cells of the culture could verify the existence of an
74 S-layer on the surface of “*Ca.* Micrarchaeum harzensis”.

75

76 **Materials and Methods**

77 Culturing conditions

78 Cultures containing “*Candidatus* Micrarchaeum harzensis A_DKE” and “*Candidatus*
79 *Scheffleriplasma hospitalis* B_DKE”, as well as the pure culture of the latter, were grown as
80 described in Krause *et al.* (2020).

81 *E. coli* Rosetta pRARE cells carrying plasmid pBAD202_*slp_{Mh}*-6x His were cultivated in a shaking
82 flask containing Terrific Broth medium (1.2 % (w/v) tryptone, 2.4 % (w/v) yeast extract,
83 0.5 % (w/v) glycerol, 17 mM KH₂PO₄, 72 mM K₂HPO₄) supplemented with 50 µg mL⁻¹
84 kanamycin and 30 µg mL⁻¹ chloramphenicol at 37 °C and 180 rpm.

85 Mass spectrometry (MS)-based proteomic analyses

86 Two 100 mL replicates of an A_DKE-B_DKE co-culture as well as a B_DKE pure culture were
87 centrifuged for 2 min at 15,500 g and 4 °C. The pellets were resuspended in 500 µL TRIS-HCl
88 buffer (pH 6,8). The cells were lysed using a Branson Digital Sonifier Model 102C (Branson
89 Ultrasonics Co. Ltd., Shanghai) for 2 min (0.5 s pulse, 20 s pause) with 60 % intensity. Cell
90 debris was pelleted via centrifugation for 10 min at 9,000 g and 4 °C. The protein concentration
91 in the supernatant was determined using the Bradford assay (Bradford, 1976). Samples were
92 mixed with Laemmli buffer (Laemmli, 1970) to a final protein concentration of 0.1 µg µL⁻¹,
93 incubated at 95 °C for 10 min, frozen in liquid nitrogen and stored at -80 °C until shipping.

94 Proteins were stacked in the top of a 4-12% NuPAGE gel (Invitrogen), stained with R-250
95 Coomassie blue, before being in-gel digested using trypsin (sequencing grade, Promega) as
96 previously described (Casabona *et al.*, 2013). The resulting peptides were analyzed by online
97 nanoliquid chromatography coupled to tandem MS (Ultimate 3000 RSLCnano and Q-Exactive
98 HF, Thermo Scientific). Peptides were sampled on a 300 µm × 5 mm PepMap C18 precolumn
99 (Thermo Scientific) and separated on a 75 µm × 250 mm C18 column (Reposil-Pur 120 C18-
100 AQ, 1.9 µm, Dr. Maisch) using a 200-min gradient. MS and MS/MS data were acquired using
101 the Xcalibur software (Thermo Scientific). Peptides and proteins were identified using Mascot
102 (version 2.6.0, Matrix Science) through concomitant searches against A_DKE and B_DKE
103 databases, homemade classical contaminant database and the corresponding reversed
104 databases. Trypsin/P was chosen as the enzyme and two missed cleavages were allowed.

105 Precursor and fragment mass error tolerances were set at 10 and 25 mmu, respectively.
106 Peptide modifications allowed during the search were: Carbamidomethyl (C, fixed), Acetyl
107 (Protein N-term, variable) and Oxidation (M, variable). The Proline software (Bouyssié *et al.*,
108 2020) was used to filter the results: conservation of rank 1 peptides, peptide score ≥ 25 ,
109 peptide length ≥ 6 amino acids, false discovery rate of peptide-spectrum-match identifications
110 $< 1\%$ as calculated on peptide-spectrum-match scores by employing the reverse database
111 strategy, and minimum of 1 specific peptide per identified protein group. Proline was then
112 used to perform a compilation, grouping and MS1 quantification of the protein groups on the
113 basis of razor and specific peptides. For each replicate, intensity-based absolute quantification
114 (iBAQ, Schwanhäusser *et al.*, 2011) values were normalized by the sum of iBAQ values in the
115 analyzed sample. The normalized iBAQ values were then summed to provide the mean iBAQ
116 value of each quantified protein.

117 In silico analysis of protein characteristics

118 The theoretical molecular mass and isoelectric point, as well as the relative amino acid content
119 of SLP_{Mh} were calculated using the CLC Main Workbench 20.0.1 (QIAGEN, Aarhus, Denmark).
120 Conserved domains in the amino acid sequence of SLP_{Mh} were identified based on hidden
121 markov models via the HHPred server (Söding *et al.*, 2005; Hildebrand *et al.*, 2009; Gabler *et al.*,
122 2020) using Pfam-A_v34 (Mistry *et al.*, 2021), TIGRFAMs_v15.0 (Haft *et al.*, 2001) and
123 PDB_mmCIF70_3_Mar (Berman *et al.*, 2000) databases as reference. Putative N- and O-linked
124 glycosylation sites, as well as GPI anchor motifs were determined using the web-servers
125 NetNGlyc 1.0 (Gupta and Brunak, 2002), NetOGlyc 4.0 (Steentoft *et al.*, 2013) and GPI-SOM
126 (Fankhauser and Mäser, 2005), respectively.

127 Identification of SLP_{Mh} homologues

128 Homologues of SLP_{Mh} in other Micrarchaeota genomes were detected via BLASTp search with
129 the following parameters; number of threads: 4, expect: 0.05, word size: 6, matrix:
130 BLOSUM62, gap cost: existence 11, extension 1. A total of 51 available Micrarchaeota
131 genomes (Table S1) were used in this study. BLASTp hits in Micrarchaeota (Table S2) were
132 analysed for known protein domains using the HMMER 3.1b1 algorithm (Finn *et al.*, 2011) in
133 combination with the Pfam database (Bateman *et al.*, 2004; Release 33.1). For visualization,
134 an alignment of SLP_{Mh} with all 52 protein sequences detected via BLASTp search, was created

135 using MUSCLE v3.8.425 (Edgar, 2004) implementation of CLC Main Workbench 20.0.1
136 (QIAGEN, Aarhus, Denmark). The alignment was visualised using the CLC Main Workbench.

137 Antibody production

138 Polyclonal Rabbit antibodies against the potential S-Layer protein of A_DKE were generated
139 by GenScript Biotech B. V. (Leiden, Netherlands). A suitable antigen peptide was designed and
140 evaluated with the proprietary OptimumAntigen™ design tool. The chosen antigen was the
141 peptide region between amino acid position 133 and 147 (NRGVKTDQYGATKT). During peptide
142 synthesis a cysteine was added to the C terminus of the peptide and conjugated to keyhole
143 limpet hemocyanin (KLH). The conjugated peptide antigen was used for rabbit immunization.

144 Cloning and recombinant expression of *slp_{Mh}*-6x His

145 The *slp_{Mh}* gene (Micr_00292) was PCR-amplified from genomic DNA isolated from a "*Ca.*
146 *Micrarchaeum harzensis* A_DKE" and "*Ca.* Scheffleriplasma hospitalis B_DKE" co-culture using
147 the oligonucleotide primers 1 & 2 (see Table S3), which introduced a 6x His-tag encoding
148 sequence to the 3'-end, as well as complementary overlaps to the target vector pBAD202
149 (Invitrogen, Carlsbad, CA, USA). The plasmid pBAD202 was linearized via inverse PCR using
150 primers 3 & 4 (see Table S3). The linearized vector and the *slp_{Mh}*-6x His fragment were gel-
151 purified using the Wizard® SV Gel and PCR Clean-Up System (Promega, Mannheim, Germany)
152 and assembled via isothermal *in vitro* ligation according to Gibson *et al.* (2009). The resulting
153 plasmid (pBAD202_*slp_{Mh}*-6xHis) was transformed into *E. coli* Rosetta pRARE (Merck,
154 Darmstadt, Germany).

155 *E. coli* Rosetta pRARE pBAD202_*slp_{Mh}*-6x His was cultivated as described above. Upon reaching
156 an OD₆₀₀ of 0.8, expression of *slp_{Mh}*-6x His was induced by addition of 1 mM L-(+)-arabinose.
157 After incubation for 18 h at 30 °C, the OD₆₀₀ of the culture was determined using a
158 GENESYS™ 20 spectrophotometer (Thermo Fisher Scientific, Schwerte, Germany) and a
159 sample (1 mL) was centrifuged at 16 000 g for 2 min. The pellet was resuspended in 75 µL of
160 2x SDS loading dye per OD₆₀₀ of 0.2, boiled for 10 min at 95 °C and centrifuged for 5 min at
161 16 000 g prior to loading on the gel.

162

163

164 Preparation of cell lysate from archaeal cells

165 Samples (50 mL each) of a dense "*Ca. Micrarchaeum harzensis* A_DKE" + "*Ca. Schefflerioplasma*
166 *hospitalis* B_DKE" co-culture, and a "*Ca. Schefflerioplasma hospitalis* B_DKE" pure culture were
167 centrifuged at 15,500 g and 4 °C for 15 min. After resuspending the pellets in 1 mL protein
168 buffer (50 mM HEPES, 150 mM NaCl, pH 8.0), OD₆₀₀ was measured using a Nanodrop 2000
169 spectrophotometer (Thermo Fisher Scientific, Schwerte, Germany) and was adjusted to 0.65
170 in both samples. Four volumes of each sample were mixed with one volume of 5x SDS loading
171 dye (600 mM TRIS-HCl pH 6.8, 50 % (v/v) glycerol, 5 % (w/v) sodium dodecyl sulfate, 0.25 %
172 (w/v) Orange G, 250 mM dithiothreitol) respectively, boiled for 10 min at 95 °C and
173 centrifuged at 16,000 g for 5 min, prior to loading on the gel.

174 SDS-PAGE, Western Blot & Protein Staining Methods

175 Cell lysate samples were separated via denaturing SDS-PAGE according to (Laemmli, 1970) in
176 hand cast 8 % TRIS-Glycine gels. Separated proteins were transferred from the acrylamide gel
177 to a nitrocellulose membrane (Roth, Karlsruhe, Germany) via a semi-dry blot with a Trans-
178 Blot® Turbo™ device (Bio-Rad, Munich, Germany) at 1.3 A for 12 min using a continuous
179 blotting buffer system (330 mM TRIS, 267 mM glycine, 15 % (v/v) ethanol, 5 % (v/v) methanol,
180 pH 8.8).

181 PAS staining of glycosylated proteins in acrylamide gels following SDS-PAGE was performed
182 according to (Segrest and Jackson, 1972). The gels were subsequently stained with
183 InstantBlue® Coomassie Protein Stain (Abcam, Cambridge, UK) according to manufacturer's
184 instructions.

185 For immuno-staining the membrane was blocked over night at room temperature with TBST
186 (20 mM TRIS, 500 mM NaCl, 0.05 % (v/v) Tween® 20, pH 7.5) containing 3 % (w/v) skim milk
187 powder and incubated with a rabbit anti-SLP_{Mh 133-147} primary antibody (GenScript, Leiden,
188 Netherlands), diluted 1:200 in TBS (10 mM TRIS, 150 mM NaCl, pH 7.5) containing 3 % (w/v)
189 bovine serum albumin for 1 h. The blot was washed with TBST (4x 5 min) and incubated with
190 a goat anti-rabbit alkaline phosphatase secondary antibody (Sigma-Aldrich, Steinheim,
191 Germany) diluted 1:30,000 in TBST containing 3 % (w/v) skim milk powder for 45 min. After
192 washing with TBST (4x 5 min) and several brief rinses with dH₂O, colorimetric band

193 visualisation was achieved using the AP conjugate substrate kit (Bio-Rad, Munich, Germany)
194 according to the manufacturer's instructions.

195 Electron microscopic samples

196 For electron microscopic imaging, 3 mL of a dense culture were centrifuged at 10,000 g for
197 10 min. The cell pellet was resuspended the remaining supernatant (~ 15 μ L), transferred onto
198 a gold carrier and rapidly frozen by plunging into liquid nitrogen. After transfer into a high-
199 vacuum freeze-etch device ($p < 10^{-5}$ mbar; CFE-50, Cressington, Watford, UK), samples were
200 freeze-fractured at a sample temperature of $T = -97$ °C. The surface water was removed by
201 sublimation at this temperature for 4 min ('freeze-etching'), and the resulting sample surfaces
202 were coated with 1.5 nm Pt/C (shadowing angle: 45 deg.) and 15 nm C (shadowing angle: 90
203 deg.). The resulting replica were cleaned by floating onto freshly prepared sulfuric acid (70 %
204 (v/v)) for 17 hours and then washed twice with ddH₂O, before being picked up onto
205 hydrophilized copper grids (600 mesh, hex) for imaging.

206 For electron cryo-microscopy of suspensions of the Archaea, about 2 x 1 mL of a grown cell
207 culture were gently concentrated by centrifugation (5,000 x g, 5 min) and resuspended in a
208 minimum amount of supernatant (ca. 10 μ L). 3 μ L were applied onto a Cu grid with a holey
209 carbon film (Quantifoil R 2/2; Quantifoil, Jena, Germany), blotted, and cryo-immobilized in
210 liquid ethane using a Leica EM-GP2 plunge freezer (Leica, Wetzlar, Germany). Grids were
211 stored in liquid nitrogen, before they were inserted into a cryoARM 200 (Z200FSC; JEOL GmbH,
212 Freising, Germany). Samples were kept at $T < 96$ K throughout all imaging steps. The grid was
213 screened at low magnification (30 x) and maps were acquired at medium magnification
214 (8,000 x) when searching for suitable sample areas. Images were taken under strict low-dose
215 conditions using a Rio16 CMOS camera (Ametek-Gatan GmbH, München, Germany), at a
216 nominal magnification of 20,000 x (pixel size: 0.39 nm), using SerialEM version 3.8.6
217 (Mastronarde, 2005).

218 Electron micrographs of freeze-etched samples were recorded on a JEM-2100F transmission
219 electron microscope (JEOL GmbH, Freising, Germany) using a CMOS camera (F416; TVIPS
220 GmbH, Gauting, Germany), using pixel sizes of about 0.25 to 1 nm (relative magnification:
221 10,000 to 40,000 x) using the software packages SerialEM 3.8.6 (Mastronarde, 2005) or EM-
222 MENU 5 (TVIPS GmbH, Gauting, Germany). The corresponding powerspectrum of selected
223 image areas were done with built-in algorithms of both software packages.

224 **Results and Discussion**

225 Transcriptomic and proteomic evidence for the production of a putative S-layer protein by 226 A_DKE

227 Publications in the past had interpreted the cell surface of Micrarchaeota to be composed of
228 two membranes. This hypothesis was based on cryo-electron microscopy and tomographic
229 reconstructions of environmental samples containing Micrarchaeota (Comolli *et al.*, 2009).
230 The latter were distinguished from other community members by cell size, shape and
231 presence of “two surrounding layers” of unknown composition. In this study we used the
232 previously published co-culture of “*Ca. Micrarchaeum harzensis* A_DKE” and its possible host
233 “*Ca. Scheffleriplasma hospitalis* B_DKE” (Krause 2020) to characterize the surface type of the
234 *Micrarchaeum* using a variety of approaches.

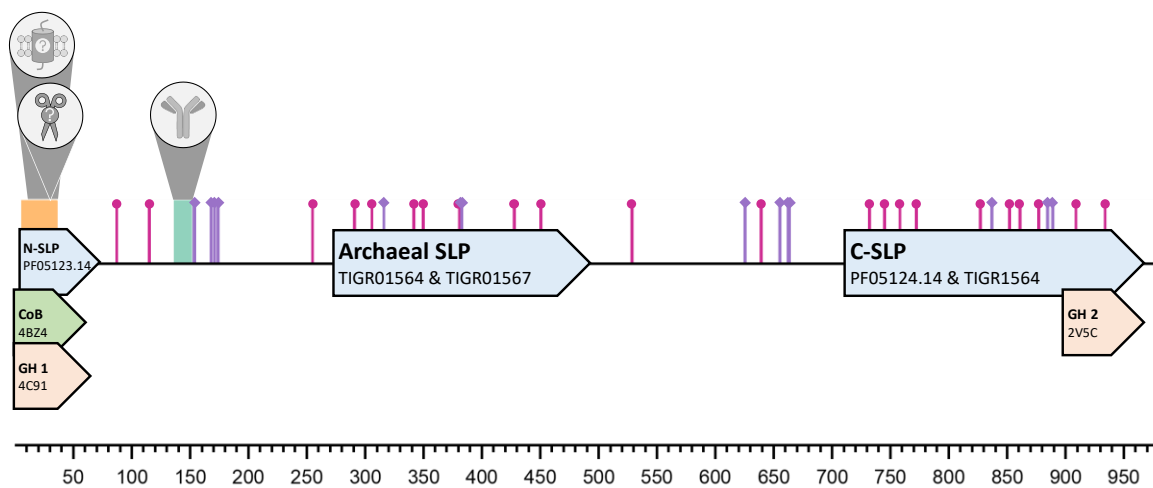
235 Surprisingly, the examination of the previously published closed genome of A_DKE (Krause *et*
236 *al.*, n.d.), revealed a gene (Micr_00292/*slp_{Mh}*) proposed to encode a protein with similarities
237 to S-layer domain PF05123 (Table S4). The genome of B_DKE does not contain such a gene.
238 This is in line with previous results revealing that S-layers are rare among members of the
239 Thermoplasmatales and occur only on cells of the Picrophilaceae family (Golyshina *et al.*,
240 2016). The putative S-layer encoding gene of A_DKE was found to be the third highest
241 expressed gene within the transcriptome (Table S4). A proteomic analysis of the above-
242 mentioned co-culture of A_DKE and B_DKE corroborated this result, as the potential S-layer
243 protein was the A_DKE protein with the highest abundance within the proteome (Table S5).

244

245 *In silico* analysis of SLP_{Mh} reveals homology to S-layer proteins

246 S-layer proteins share only a few traits among their primary structure, which renders their
247 identification on a bioinformatic level difficult. S-layer proteins are described to have a
248 molecular mass between 40 and 200 kDa, an isoelectric point between 4 and 6 and a content
249 of 40 to 60 % hydrophobic, as well as only few sulfur-containing amino acids (Sára and Sleytr,
250 2000; Sleytr *et al.*, 2014). With a predicted molecular mass and an isoelectric point of 101 kDa
251 and 5.55, respectively, as well as being comprised of about 50 % hydrophobic and almost no
252 sulfur-containing amino acids (Table S4), SLP_{Mh} matches these characteristics quite well. In

253 addition, numerous putative N- and O-glycosylation sites located throughout the whole
254 sequence (Figure 1) indicate strong glycosylation of SLP_{Mh}, which is a common trait of many
255 known archaeal and bacterial S-layer proteins (Sumper *et al.*, 1990; Kandiba and Eichler,
256 2014). Lastly, conserved S-layer domains at the N- and C-terminus of the protein, as well as a
257 single domain in between, with strong homology to S-layer proteins from several Archaea such
258 as *Pyrococcus horikoshii*, *Methanococcus jannaschii* (both TIGR01564) and *Methanosarcina*
259 species (TIGR01567) could be detected. Contrary to SLP_{Mh}, this specific domain is tandemly
260 duplicated in *Methanosarcina* species (Arbing, 2012). Overall, the presented data strongly
261 implies that SLP_{Mh} is indeed an S-layer protein.



262
263 Figure 1: Putative domain architecture of SLP_{Mh}. Sequence motifs such as the N-terminal region containing a
264 putative transmembrane domain or signal peptide is shown in yellow. Putative N- and O-glycosylation sites are
265 highlighted by pink circles and purple diamonds respectively. The peptide-antigen of the SLP_{Mh} 133-147 IgG is
266 highlighted in teal. Putative domains with homology to S-layer proteins (N/C-SLP; SLP), copper binding proteins
267 (CoB) and glycoside hydrolases (GH) are depicted as arrows and are coloured according to their respective
268 prediction probability: blue - > 90 %; green ≈ 50 %; orange < 30 %. The respective data base accession numbers
269 of the putative domains are given as well. For more detailed information on the predicted domains see Table S6.

270 Further conserved motifs in the SLP_{Mh} primary structure allow speculation about how exactly
271 the protein could be attached to the A_DKE cell membrane. Unfortunately, there are
272 conflicting predictions whether there is a signal peptide (Phobius & SignalP) or
273 transmembrane domain (TMHMM) located at the N-terminus of SLP_{Mh} and the compiled
274 proteomic data does not suffice to verify the cleavage of the putative signal peptide in the
275 mature protein (data not shown). In case of a signal peptide, which is predicted to be cleft
276 between Ala26 and Gly27 upon secretion, resulting in the loss of the transmembrane domain,
277 SLP_{Mh}-anchoring might depend on a second SLP with a transmembrane domain acting as a
278 stalk, as described in several *Sulfolobales* species (Veith *et al.*, 2009). Other than
279 transmembrane domains, archaeal S-layer proteins can be anchored by either a lipid

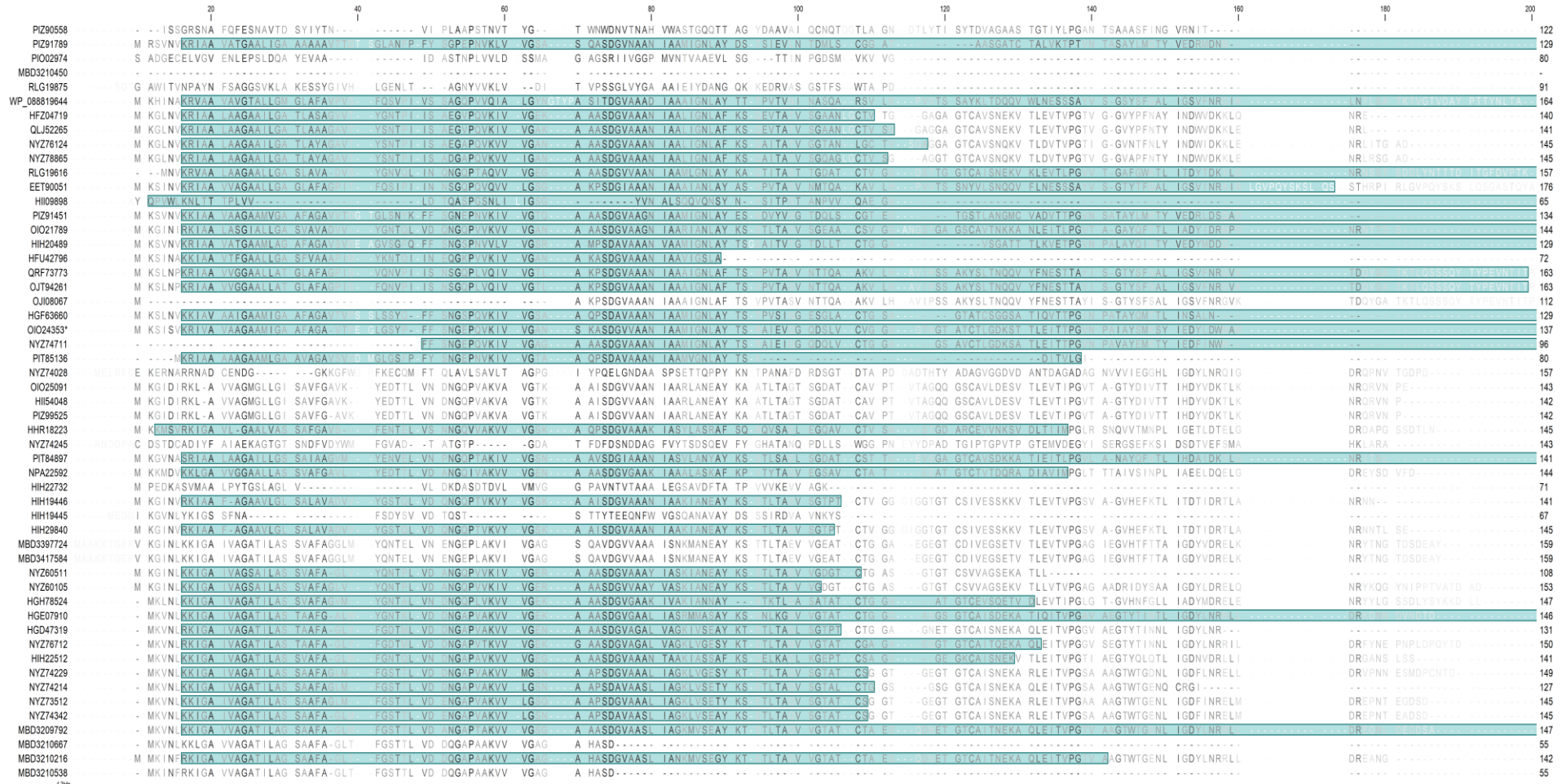
280 modification or interaction with cell wall polymers (reviewed in Sleytr *et al.*, 2014). As a
281 corresponding signal motif could not be detected, SLP_{Mh} does not seem to comprise a
282 glycosylphosphatidylinositol anchor. Still, other sorts of lipid modifications to anchor the
283 protein in the cell membrane are possible, i.e. as described for *Haloferax volcanii* (Kandiba *et al.*
284 *et al.*, 2013). Further experiments with isolated SLP_{Mh} are necessary to answer this question,
285 however. Since there is no indication for synthesis of archaeal cell wall polymers in the A_DKE
286 genome (Krause *et al.*, n.d.), anchoring of the S-layer via cell wall interaction as described for
287 *Methanothermus spec.* (Albers and Meyer, 2011) is not likely. At its very N- and C-terminus
288 SLP_{Mh} shows weak homology to glycoside hydrolases (Figure 1). These domains could
289 potentially allow glycan binding and therefore might facilitate membrane anchoring in a
290 lipopolysaccharide matrix or physical interaction with B_DKE cells. Since the former has so far
291 been exclusively described in Gram-negative Bacteria (see examples in Engelhardt, 2007;
292 Sleytr *et al.*, 2014), such as *Caulobacter crescentus* (von Kugelgen *et al.*, 2020) and A_DKE also
293 lacks crucial enzymes for the synthesis of lipopolysaccharides (see below), the latter seems to
294 be more likely. Due to homology to a copper binding protein, the N-terminal domain could
295 also allow binding of divalent cations (i.e. Ca²⁺), which are known to stabilize S-layer structures
296 in many bacterial and archaeal cells, i.e. in *Haloferax volcanii* (Cohen *et al.*, 1991; Engelhardt,
297 2007; Sleytr *et al.*, 2014).

298

299 Formation of S-Layers might be a common characteristic of Micrarchaeota members

300 A BLASTp database search for similarities to SLP_{Mh} revealed 52 homologues in 38 out of 49
301 examined Micrarchaeota genomes. Whereas some of these homologs were already
302 annotated as S-layer proteins, others are so far annotated as hypothetical or uncharacterized
303 proteins. An alignment of these proteins revealed conserved regions throughout the whole
304 sequence, but especially at the N-terminal end containing the S-layer domain. Figure 2 shows
305 the first 200 positions of this alignment which cover the S-layer domain of SLP_{Mh}, revealing a
306 high conservation on amino acid level. Cross-referencing the BLAST hits with the PFAM protein
307 domain database revealed that 38 of the proteins show homologies to N-terminal S-layer
308 domains and one to a C-terminal S-layer domain (Table S2). This finding suggests that also
309 other Micrarchaeota can form an S-layer, which would make the proteinaceous surface layer
310 a common characteristic of the phylum. In fact, so far, we could not find a nearly complete

311 genome of a Micrarchaeota member that does not seem to possess a corresponding gene for
312 an S-layer protein. Whether or not these genes are expressed under natural conditions cannot
313 be revealed in all cases, due to missing data. Still, the homologue of SLP_{Mh} in “*Ca.*
314 *Micrarchaeum acidiphilum* ARMAN-2” (EET90051.1) was found to be the most abundant
315 protein assigned to ARMAN-2 in proteomic data of acid mine drainage (AMD) biofilms (Baker
316 *et al.*, 2010).



317

318 Figure 2: MUSCLE-Alignment of SLP_{Mh} homologues. Depicted are the first 200 alignment positions of each sequence, with the NCBI accession numbers on the left and a sequence
 319 logo on the bottom. Amino acids shown in the corresponding sequence logo are coloured according to their polarity (black: hydrophobic, hydrophilic: green, anionic: red, cationic:
 320 blue). Faint blue background colour displays predicted S-layer domains. Conservation of amino acids at each position is indicated through a colour gradient from white (0 %) to
 321 black (100 %).

322 Immuno-staining reveals potential glycosylation of the mature protein

323 To detect the S-layer protein SLP_{Mh}, we constructed polyclonal rabbit antibodies against a 14
324 amino acid peptide of the protein (Figure 1). The functionality of the antibody was verified via
325 SDS-PAGE and Western Blot, followed by an immuno-detection (Figure 3a & b). Lysed cells of
326 the co-culture containing A_DKE and B_DKE were compared with the B_DKE pure culture.
327 Whereas no signal was obtained with the cell lysate of the pure culture, the cell lysate of the
328 co-culture showed one signal corresponding to a protein with a molecular mass higher than
329 130 kDa. Interestingly, this is roughly 30 kDa more than what was expected regarding the
330 molecular mass deduced from the protein sequence. The 130 kDa signal was also distinctly
331 visible in Coomassie stained co-culture cell lysates (Figure 3a), which corroborates a high
332 abundance of this protein in the culture.

333 Possible explanations for a higher apparent molecular mass are post-translational
334 modifications. Many S-layer proteins are described to be glycosylated in the literature
335 (Sumper *et al.*, 1990; Kandiba and Eichler, 2014). Moreover, the bioinformatic analysis of
336 SLP_{Mh} also revealed numerous putative glycosylation sites (Figure 1). Hence, this potential
337 reason for the higher molecular mass was analysed in further experiments.

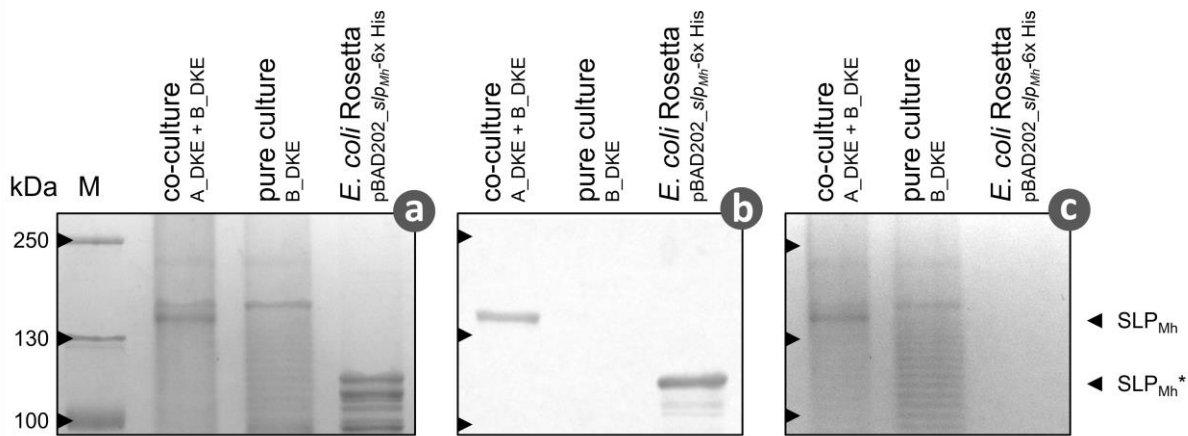
338

339 SLP_{Mh} is a glycosylated protein

340 To investigate the hypothesis of a glycosylated protein, we conducted a PAS stain of the cell
341 lysate from co-culture and pure culture and correlated it to the immuno-detection with the
342 developed antibody (Figure 3c). PAS staining revealed a corresponding band in the co-culture,
343 which correlates to the SLP_{Mh} signal after immuno-detection. This point together with the
344 absence of the signal in the pure culture strongly suggests that the higher molecular mass is
345 due to glycosylation. The additional signals stained in pure and co-culture seem to be
346 glycosylated proteins of the Thermoplasmatales relative.

347 In order to provide further evidence for the posttranslational modification of SLP_{Mh} we
348 conducted heterologous expression in *E. coli*. We hypothesized that the lacking
349 posttranslational modification should lead to a protein matching the theoretical molecular
350 mass of 100 kDa. As expected, immuno-staining showed a corresponding band at around

351 100 kDa, while the PAS stain was negative. Thus, the higher apparent molecular mass of SLP_{Mh}
352 is very likely a result of post-translational glycosylation.



353
354 Figure 3: Detection of S-Layer protein & glycosylation. SDS-PAGE (8 % T resolving gel) of cell lysate samples from
355 a co-culture, containing “*Ca. Micrarchaeum harzensis* A_DKE” and “*Ca. Scheffleriplasma hospitalis* B_DKE”, a
356 pure culture, containing only “*Ca. Scheffleriplasma hospitalis* B_DKE”, as well as a culture of *E. coli* Rosetta pRARE
357 pBAD202_slp_{Mh}-6x His. Gels were either PAS- [c] and subsequently Coomassie-stained [a], or subjected to
358 Western Blotting followed by immuno-detection, using an anti-SLP_{Mh} 133-147 primary antibody [b]. The bands
359 corresponding to SLP_{Mh} and recombinant SLP_{Mh} (*) are highlighted.

360 Interestingly, bioinformatic analysis could not provide evidence for the synthesis of activated
361 carbohydrates by A_DKE. Still, the organism seems to possess several parts of a canonical N-
362 glycosylation machinery. Table S7 summarizes the results of the bioinformatic search for
363 corresponding genes needed for glycosylation reactions. Detected glycosyltransferases
364 typically transfer carbohydrates from UDP-glucose, UDP-N-acetyl-galactosamine, GDP-
365 mannose or CDP-abequose. Besides the latter, all other carbohydrates can potentially be
366 produced by the host organism B_DKE. Further analysis will have to reveal how these
367 molecules are exchanged by the two organisms. In this regard, a study by Krause *et al.*
368 provides evidence that the two organisms might interact similarly as *Ignicoccus hospitalis* and
369 *N. equitans*. For these cells a cytoplasmic bridge was proven to exist by electron tomography
370 (Heimerl *et al.*, 2017), which would render the exchange of even complex metabolites
371 possible.

372

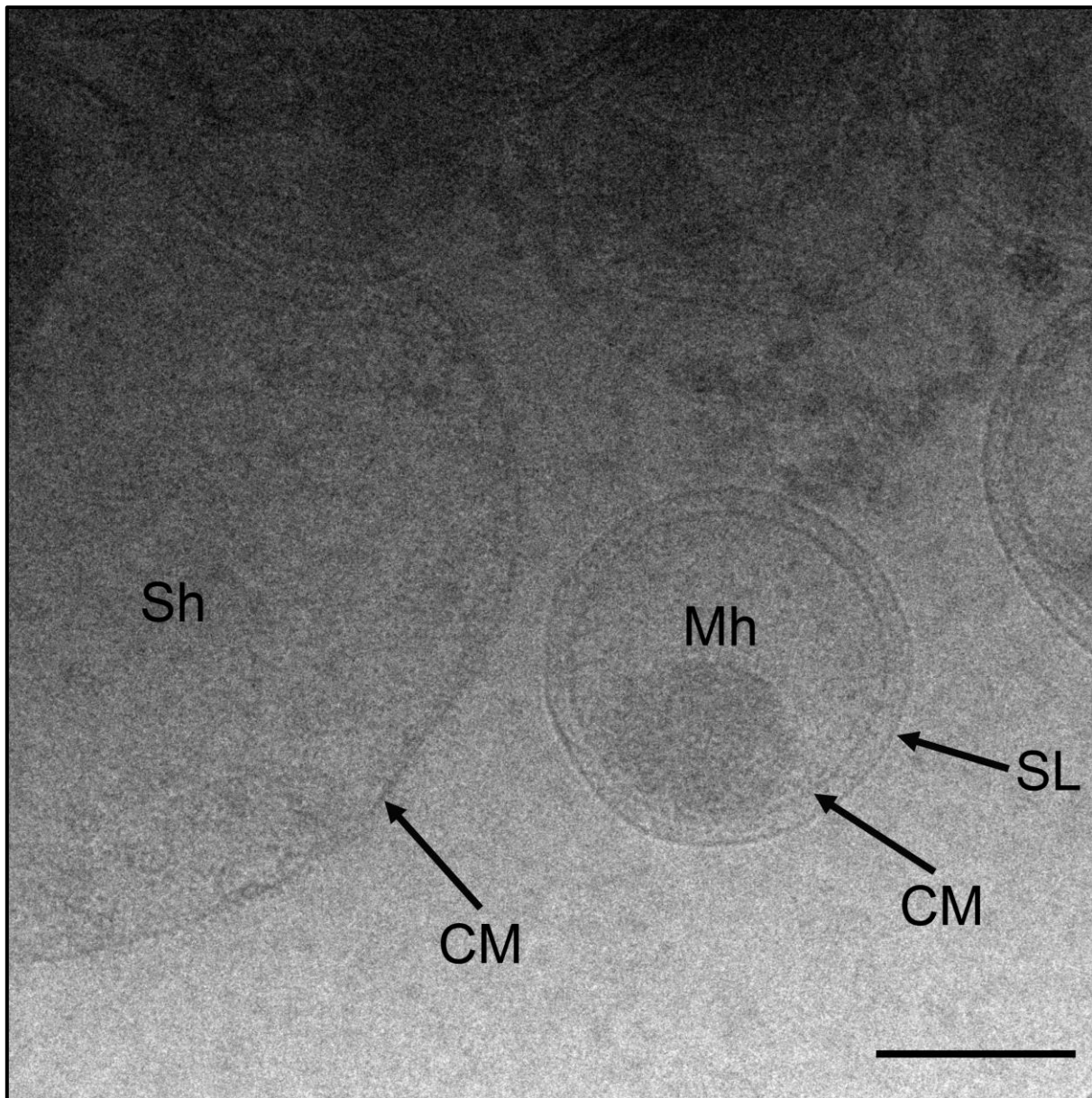
373

374 Electron microscopy

375 As a final prove for the existence of an S-layer covering A_DKE, the surface structure of cells
376 in co-culture samples containing A_DKE and B_DKE was examined via electron microscopy.

377 Cryo-EM images show cells with a larger diameter (~0.8 to 1.5 μm) and one surrounding layer,
378 as well as cells with a smaller diameter (~0.4 μm) and two layers, the latter resembling the
379 cells imaged during the study of Comolli *et al.* (2009) (Figure 4). Deduced from published
380 information on the surface of Thermoplasmatales and ultra-thin sections (Darland *et al.*, 1970;
381 Yasuda *et al.*, 1995; Golyshina *et al.*, 2009; Golyshina *et al.*, 2016) we identified the larger cells
382 as the Thermoplasmatales relative B_DKE, rendering the observed layer a cytoplasmic
383 membrane. We hypothesized that the smaller cells are in fact “Ca. Micrarchaeum harzensis
384 A_DKE” cells and suggest based on our previous findings the inner layer to be a cytoplasmic
385 membrane and the outer one the S-layer.

386 We were able to measure 20 A_DKE cells and calculated an inner diameter of 400 nm and an
387 outer diameter of 440 nm, leaving a pseudoperiplasmic space of 20-25 nm (Table S8).

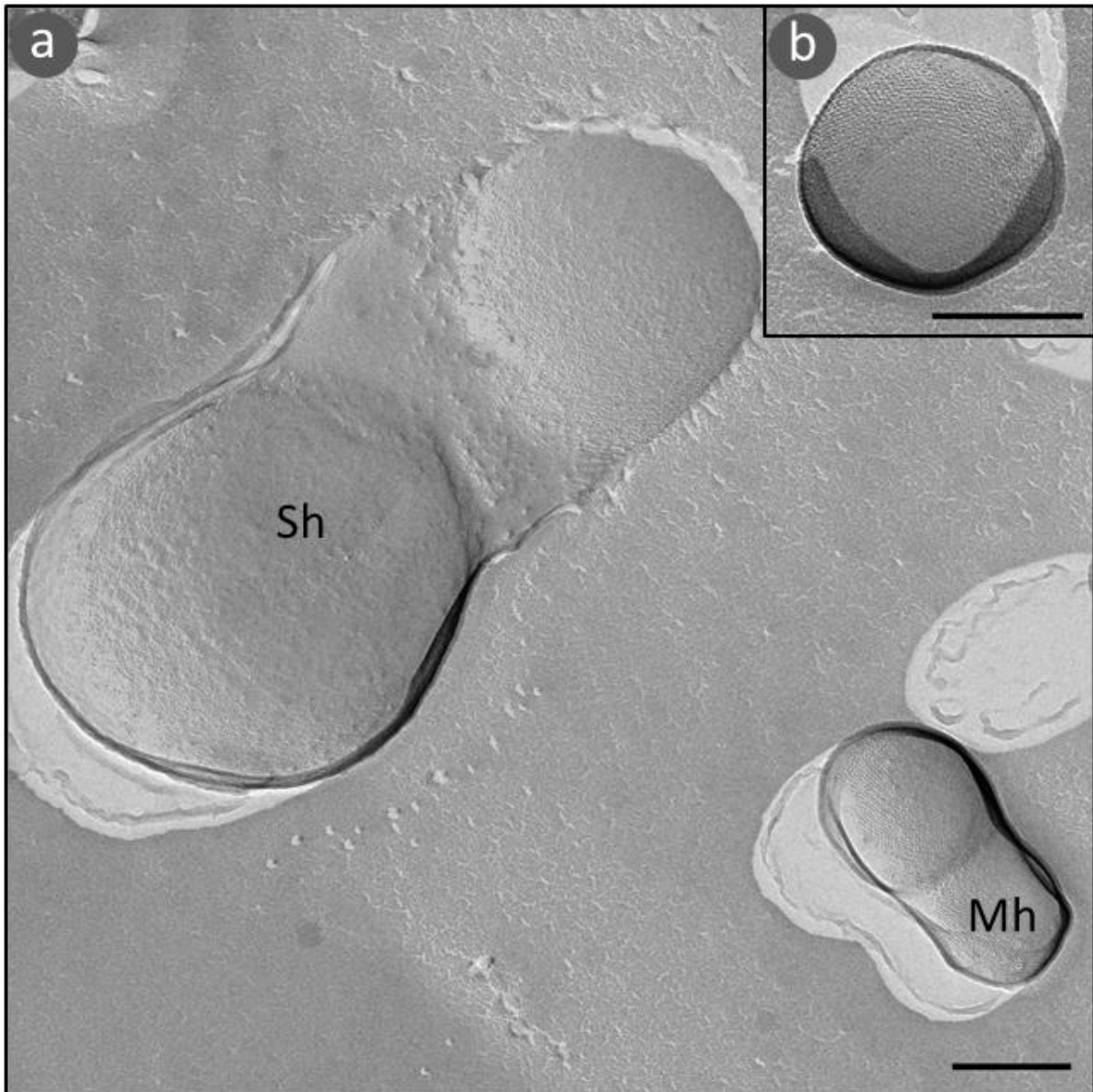


388

389 Figure 4. Cryo-EM micrograph of fully hydrated co-culture cells, frozen in liquid ethane. The image shows a “*Ca.*
390 *Scheffleriplasma hospitalis B_DKE*” cell (Sh) next to a “*Ca. Micrarchaeum harzensis A_DKE*” (Mh) cell. Arrows
391 indicate cytoplasmic membranes (CM) and S-layer (SL). Scale bar, 200 nm.

392

393 A fast and easy way to unequivocally visualize an S-layer on the surface of cells is the electron
394 microscopic imaging of metal-shadowed, freeze-etched cultures. Hence, we conducted this
395 experiment with a sample of a co-culture, containing the DPANN member and its host B_DKE.
396 As a control, we also examined the pure culture of B_DKE.



397

398 Figure 5: Electron micrographs of freeze-etched, Platinum-Carbon shadowed cells. The images show surface and
399 cell morphology of dividing "*Ca. Schefflerioplasma hospitalis* B_DKE" (Sh) and "*Ca. Micrarchaeum harzensis*
400 A_DKE" (Mh) cells (a) and a close-up of a single A_DKE cell (b). Scale bars, 200 nm.

401 As expected from cryo-electron microscopy, electron micrographs of the metal-shadowed co-
402 culture showed two cell types, distinguishable by cell size and surface morphology (Figure 5).
403 The diameters of the observed cells matched those determined from cryo-EM images. The
404 cells with larger diameter, pleomorphic shape and a rough surface are most likely the host
405 organism B_DKE. A_DKE cells have a smaller diameter and are evenly round. These cells clearly
406 displayed an S-layer on their surface.

407 The quantitative determination of the lattice, i.e. the distance of the molecular complexes, is
408 notoriously difficult on small cells. The small diameter inevitably results in a highly bent lattice,
409 the regularity in projection images is reduced and nice reflexes in the powerspectrum are
410 hardly visible. Nevertheless, we have conducted measurements on few cells with a visible S-
411 layer pattern and determined the distance to be at around 16 ± 1 nm (Figure S1). Most likely,
412 the complexes are arranged on a hexagonal lattice; we cannot exclude, however, that the
413 complexes are arranged on a lattice with lower symmetry (p3 or even p2). Note that width of
414 periplasm and centre to centre distance of S-layer complexes in A_DKE corresponds to
415 characteristics of *N. equitans* cells (Huber *et al.*, 2003). While there is no significant sequence
416 conservation between the *N. equitans* S-layer protein (NEQ300) and SLP_{Mh}, the putative
417 domain architecture of both proteins determined via HHPred is quite similar (Table S9).

418 Hence, cells assigned to “*Ca. Scheffleriplasma hospitalis* B_DKE” possess one membrane only
419 with no additional surface polymer, while Micrarchaeota cells exhibit one membrane covered
420 by an S-layer. Judging from our results we postulate that the published structural data by
421 Comolli and colleagues potentially show a Micrarchaeum with a surface layer, similar to
422 A_DKE and *N. equitans* (Huber *et al.*, 2002).

423 **Conclusion**

424 During the last decades, cultivation-independent approaches expanded our knowledge about
425 the diversity and evolution of microorganisms. However, laboratory cultures remain essential
426 for detailed characterization of an organism's genomic potential. Cultivation of the
427 *Micrarchaeum* A_DKE proved to be a critical point in order to investigate its cell morphology,
428 since differentiation between cells in electron micrographs is rather difficult and the
429 pleomorphic character and heterogeneous diameter of *Thermoplasmatales* cells, which are
430 often associated with *Micrarchaeota* cells, renders a discrimination even more complex. Via
431 electron micrographs of metal-shadowed, freeze-etched cells of the co-culture containing
432 A_DKE and its putative host B_DKE, we were able to identify the DPANN member as round
433 cells with one membrane and a proteinaceous S-layer. This discovery was supported by the
434 detection of an S-layer protein encoding gene in the closed genome of A_DKE. The
435 identification of homologous proteins in other *Micrarchaeota* genomes suggests a distribution
436 in the whole phylum.

437 Therefore, we propose a similar surface type to *Nanoarchaeum equitans* with one membrane
438 and a surface-spanning S-layer for members of the phylum *Micrarchaeota*. Glycosylation of
439 the A_DKE S-layer protein seems to be dependent on a cooperation of A_DKE and B_DKE.
440 While glycosylation of the S-layer protein could be catalysed by A_DKE alone, which expresses
441 a functional N-glycosylation system, activated carbohydrates for chain elongation must be
442 provided by B_DKE.

443 Further advanced examination techniques like subtomogram averaging can reveal details of
444 the S-layer ultrastructure at near-atomic resolution including the anchoring mechanism of
445 SLP_{Mh} in the membrane (Bharat *et al.*, 2017; von K ugelgen *et al.*, 2020).

446 Information about the surface of A_DKE and B_DKE will most probably be essential to
447 understand the fundamental basis of the direct physical contact between these two *Archaea*
448 (Krause *et al.*, n.d.; Junglas *et al.*, 2008; Giannone *et al.*, 2011). Other studies showed that S-
449 layers can have a role in recognition and interaction between prokaryotic cells in general and
450 especially between different *Haloferax* species (Shalev *et al.*, 2018). S-layer proteins have been
451 discussed to stabilize the cell as a kind of exoskeleton (Engelhardt, 2007) and influence or
452 stabilize the proteins in the underlying membrane, like for example the proteins of the
453 archaeellum complex (Banerjee *et al.*, 2015). At the site of direct interaction however, between
454 *N. equitans* and *I. hospitalis* the S-layer of *N. equitans* is seen to be decomposed specifically at

455 the attachment site (Heimerl *et al.*, 2017). Interaction studies with other membrane proteins
456 of A_DKE and B_DKE on the one hand, as well as transplantation of the S-layer into other
457 organisms to observe its potential physical interaction, might help to reveal a putative
458 function of SLP_{Mh} during the interaction.

459 **Acknowledgements**

460 We highly appreciate the use of the JEOL JEM-2100F in the Institute of Molecular and Cellular
461 Anatomy, Prof. Dr. Ralph Witzgall and of the JEOL Z200-FSC (cryoARM 200) of the Faculty of
462 Biology and Preclinical Medicine, University of Regensburg.

463 We thank Helena Hoang for generation of genetically modified strain *E. coli* Rosetta pRARE
464 pBAD202_*slp_{Mh}*-6x His during the course of her bachelor thesis.

465 Proteomic experiments were partly supported by the French National Agency for Research,
466 grant number ANR-10-INBS-08-01 (Proteomics French Infrastructure).

467 **References**

- 468 Albers, S.-V., and Meyer, B.H. (2011) The archaeal cell envelope. *Nat Rev Microbiol* **9**: 414–
469 426.
- 470 Arbing, M.A. (2012) Structure of the surface layer of the methanogenic archaean
471 *Methanosarcina acetivorans*. *Proc Natl Acad Sci USA* **109**.
- 472 Baker, B.J., Comolli, L.R., Dick, G.J., Hauser, L.J., Hyatt, D., Dill, B.D., *et al.* (2010) Enigmatic,
473 ultrasmall, uncultivated Archaea. *Proc Natl Acad Sci U S A* **107**: 8806–8811.
- 474 Banerjee, A., Tsai, C.-L., Chaudhury, P., Tripp, P., Arvai, A.S.S., Ishida, J.P.P., *et al.* (2015) FlaF Is
475 a β -sandwich protein that anchors the archaeellum in the archaeal cell envelope by binding the
476 S-layer protein. *Structure* **23**: 863–872.
- 477 Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., *et al.* (2004) The
478 Pfam protein families database. *Nucleic Acids Res* **32**: D138–D141.
- 479 Baumeister, W., and Lembcke, G. (1992) Structural features of archaebacterial cell envelopes.
480 *J Bioenerg Biomembr* **24**: 567–575.
- 481 Baumeister, W., Wildhaber, I., and Engelhardt, H. (1988) Bacterial surface proteins. Some
482 structural, functional and evolutionary aspects. *Biophys Chem* **29**: 39–49.
- 483 Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., *et al.* (2000) The
484 Protein Data Bank. *Nucleic Acids Res* **28**: 235–242.
- 485 Bharat, T.A.M., Kureisaite-Ciziene, D., Hardy, G.G., Yu, E.W., Devant, J.M., Hagen, W.J.H., *et al.*
486 (2017) Structure of the hexagonal surface layer on *Caulobacter crescentus* cells. *Nat Microbiol*
487 **2**: 1–6.
- 488 Bouyssié, D., Hesse, A.-M., Mouton-Barbosa, E., Rompais, M., Macron, C., Carapito, C., *et al.*
489 (2020) Proline: an efficient and user-friendly software suite for large-scale proteomics.
490 *Bioinformatics* **36**: 3148–3155.
- 491 Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram
492 quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254.
- 493 Casabona, M.G., Vandenbrouck, Y., Attree, I., and Couté, Y. (2013) Proteomic characterization
494 of *Pseudomonas aeruginosa* PAO1 inner membrane. *Proteomics* **13**: 2419–2423.
- 495 Castelle, C.J., and Banfield, J.F. (2018) Major new microbial groups expand diversity and alter
496 our understanding of the tree of life. *Cell* **172**: 1181–1197.
- 497 Castelle, C.J., Wrighton, K.C., Thomas, B.C., Hug, L.A., Brown, C.T., Wilkins, M.J., *et al.* (2015)
498 Genomic expansion of domain Archaea highlights roles for organisms from new phyla in

- 499 anaerobic carbon cycling. *Curr Biol* **25**: 690–701.
- 500 Chen, L.-X., Méndez-García, C., Dombrowski, N., Servín-Garcidueñas, L.E., Eloë-Fadrosch, E.A.,
501 Fang, B.-Z., *et al.* (2017) Metabolic versatility of small Archaea Micrarchaeota and
502 Parvarchaeota. *ISME J* **1**.
- 503 Cohen, S., Shilo, M., and Kessel, M. (1991) Nature of the salt dependence of the envelope of
504 a Dead Sea Archaeobacterium, *Haloferax volcanii*. *Arch Microbiol* **156**: 198–203.
- 505 Comolli, L.R., Baker, B.J., Downing, K.H., Siegerist, C.E., and Banfield, J.F. (2009) Three-
506 dimensional analysis of the structure and ecology of a novel, ultra-small archaeon. *ISME J* **3**:
507 159–167.
- 508 Comolli, L.R., and Banfield, J.F. (2014) Inter-species interconnections in acid mine drainage
509 microbial communities. *Front Microbiol* **5**: 367.
- 510 Darland, G., Brock, T.D., Samsonoff, W., and Conti, S.F. (1970) A thermophilic, acidophilic
511 mycoplasma isolated from a coal refuse pile. *Science (80-)* **170**: 1416–1418.
- 512 Dombrowski, N., Lee, J.-H., Williams, T.A., Offre, P., and Spang, A. (2019) Genomic diversity,
513 lifestyles and evolutionary origins of DPANN archaea. *FEMS Microbiol Lett* **366**.
- 514 Dombrowski, N., Williams, T.A., Sun, J., Woodcroft, B.J., Lee, J.-H., Minh, B.Q., *et al.* (2020)
515 Undinarchaeota illuminate DPANN phylogeny and the impact of gene transfer on archaeal
516 evolution. *Nat Commun* **11**: 3939.
- 517 Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high
518 throughput. *Nucleic Acids Res* **32**: 1792–1797.
- 519 Engelhardt, H. (1988) 16 correlation averaging and 3-D reconstruction of 2-D crystalline
520 membranes and macromolecules. Elsevier, pp. 357–413.
- 521 Engelhardt, H. (2007) Are S-layers exoskeletons? The basic function of protein surface layers
522 revisited. *J Struct Biol* **160**: 115–124.
- 523 Fankhauser, N., and Mäser, P. (2005) Identification of GPI anchor attachment signals by a
524 Kohonen self-organizing map. *Bioinformatics* **21**: 1846–1852.
- 525 Finn, R.D., Clements, J., and Eddy, S.R. (2011) HMMER web server: interactive sequence
526 similarity searching. *Nucleic Acids Res* **39**: W29–W37.
- 527 Gabler, F., Nam, S.-Z., Till, S., Mirdita, M., Steinegger, M., Söding, J., *et al.* (2020) Protein
528 sequence analysis using the MPI bioinformatics toolkit. *Curr Protoc Bioinforma* **72**: e108.
- 529 Giannone, R.J., Huber, H., Karpinets, T., Heimerl, T., Küper, U., Rachel, R., *et al.* (2011)
530 Proteomic characterization of cellular and molecular processes that enable the

- 531 *Nanoarchaeum equitans-Ignicoccus hospitalis* relationship. *PLoS One* **6**: e22942.
- 532 Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O. (2009)
533 Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**: 343–
534 345.
- 535 Golyshina, O. V., Yakimov, M.M., Lünsdorf, H., Ferrer, M., Nimtz, M., Timmis, K.N., *et al.* (2009)
536 *Acidiplasma aeolicum* gen. nov., sp. nov., a euryarchaeon of the family Ferroplasmaceae
537 isolated from a hydrothermal pool, and transfer of *Ferroplasma cupricumulans* to *Acidiplasma*
538 *cupricumulans* comb. nov. *Int J Syst Evol Microbiol* .
- 539 Golyshina, O. V, Lünsdorf, H., Kublanov, I. V, Goldenstein, N.I., Hinrichs, K.-U., and Golyshin,
540 P.N. (2016) The novel extremely acidophilic, cell-wall-deficient archaeon *Cuniculiplasma*
541 *divulgatum* gen. nov., sp. nov. represents a new family, Cuniculiplasmataceae fam. nov., of
542 the order Thermoplasmatales. *Int J Syst Evol Microbiol* **66**: 332–340.
- 543 Golyshina, O. V, Toshchakov, S. V, Makarova, K.S., Gavrillov, S.N., Korzhenkov, A., Cono, V. La,
544 *et al.* (2017) Mysterious “ARMAN” archaea depend on association with euryarchaeal host in
545 culture in situ. *Nat Commun* 1–11.
- 546 Gupta, R., and Brunak, S. (2002) Prediction of glycosylation across the human proteome and
547 the correlation to protein function. *Pacific Symp Biocomput* 310–22.
- 548 Haft, D.H., Loftus, B.J., Richardson, D.L., Yang, F., Eisen, J.A., Paulsen, I.T., and White, O. (2001)
549 TIGRFAMs: a protein family resource for the functional identification of proteins. *Nucleic Acids*
550 *Res* **29**: 41–43.
- 551 Hamm, J.N., Erdmann, S., Eloë-Fadrosch, E.A., Angeloni, A., Zhong, L., Brownlee, C., *et al.* (2019)
552 Unexpected host dependency of Antarctic Nanohaloarchaeota. *Proc Natl Acad Sci* **116**: 14661
553 LP – 14670.
- 554 Heimerl, T., Flechsler, J., Pickl, C., Heinz, V., Salecker, B., Zweck, J., *et al.* (2017) A complex
555 endomembrane system in the archaeon *Ignicoccus hospitalis* tapped by *Nanoarchaeum*
556 *equitans*. *Front Microbiol* **8**: 1072.
- 557 Hildebrand, A., Remmert, M., Biegert, A., and Söding, J. (2009) Fast and accurate automatic
558 structure prediction with HHpred. *Proteins Struct Funct Bioinforma* **77**: 128–132.
- 559 Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C., and Stetter, K.O. (2002) A new
560 phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**: 63–
561 67.
- 562 Huber, H., Hohn, M.J., Stetter, K.O., and Rachel, R. (2003) The phylum Nanoarchaeota: Present
563 knowledge and future perspectives of a unique form of life. *Res Microbiol* **154**: 165–171.

- 564 John, E. St., Liu, Y., Podar, M., Stott, M.B., Meneghin, J., Chen, Z., *et al.* (2018) A new symbiotic
565 nanoarchaeote (*Candidatus Nanoclepta minutus*) and its host (*Zestosphaera tikiterensis* gen.
566 nov., sp. nov.) from a New Zealand hot spring. *Syst Appl Microbiol* **42**: 94–106.
- 567 Junglas, B., Briegel, A., Burghardt, T., Walther, P., Wirth, R., Huber, H., and Rachel, R. (2008)
568 *Ignicoccus hospitalis* and *Nanoarchaeum equitans*: ultrastructure, cell–cell interaction, and 3D
569 reconstruction from serial sections of freeze-substituted cells and by electron
570 cryotomography. *Arch Microbiol* **190**: 395.
- 571 Kandiba, L., and Eichler, J. (2014) Archaeal S-layer glycoproteins: post-translational
572 modification in the face of extremes. *Front Microbiol* **5**: 661.
- 573 Kandiba, L., Guan, Z., and Eichler, J. (2013) Lipid modification gives rise to two distinct
574 *Haloferax volcanii* S-layer glycoprotein populations. *Biochim Biophys Acta - Biomembr* **1828**:
575 938–943.
- 576 König, H., Rachel, R., and Claus, H. (2007) Proteinaceous surface layers of Archaea:
577 Ultrastructure and biochemistry. *Archaea* 315–340.
- 578 Krause, S., Gfrerer, S., Reuse, C., Dombrowski, N., Villanueva, L., Bunk, B., *et al.* Unraveling the
579 critical growth factors for stable cultivation of (nano-sized) Micrarchaeota. *Unpublished*.
- 580 Kügelgen, A. von, Tang, H., Hardy, G.G., Kureisaite-Ciziene, D., Brun, Y. V, Stansfeld, P.J., *et al.*
581 (2020) *In situ* structure of an intact lipopolysaccharide-bound bacterial surface layer. *Cell* **180**:
582 348-358.e15.
- 583 Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of
584 bacteriophage T4. *Nature* **227**: 680–685.
- 585 Mastronarde, D.N. (2005) Automated electron microscope tomography using robust
586 prediction of specimen movements. *J Struct Biol* **152**: 36–51.
- 587 Mescher, M.F., Strominger, J.L., and Watson, S.W. (1974) Protein and Carbohydrate
588 Composition of the Cell Envelope of *Halobacterium salinarium*. *J Bacteriol* **120**: 945 LP – 954.
- 589 Messner, P., and Sleytr, U.B. (1991) Bacterial surface layer glycoproteins. *Glycobiology* **1**: 545–
590 551.
- 591 Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., *et al.*
592 (2021) Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**: D412–D419.
- 593 Probst, A.J., Ladd, B., Jarett, J.K., Geller-McGrath, D.E., Sieber, C.M.K., Emerson, J.B., *et al.*
594 (2018) Differential depth distribution of microbial function and putative symbionts through
595 sediment-hosted aquifers in the deep terrestrial subsurface. *Nat Microbiol* **3**: 328–336.
- 596 Probst, A.J., Weinmaier, T., Raymann, K., Perras, A., Emerson, J.B., Rattei, T., *et al.* (2014)

- 597 Biology of a widespread uncultivated archaeon that contributes to carbon fixation in the
598 subsurface. *Nat Commun* **5**: 5497.
- 599 Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., *et al.* (2013)
600 Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: 431–
601 437.
- 602 Sára, M., and Sleytr, U.B. (1996) Crystalline bacterial cell surface layers (S-layers): From cell
603 structure to biomimetics. *Prog Biophys Mol Biol* **65**: 83–111.
- 604 Sára, M., and Sleytr, U.B. (2000) S-layer proteins. *J Bacteriol* **182**: 859 LP – 868.
- 605 Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., *et al.* (2011) Global
606 quantification of mammalian gene expression control. *Nature* **473**: 337–342.
- 607 Segrest, J.P., and Jackson, R.L.B.T.-M. in E. (1972) [5] Molecular weight determination of
608 glycoproteins by polyacrylamide gel electrophoresis in sodium dodecyl sulfate. In *Complex*
609 *Carbohydrates Part B*. Academic Press, pp. 54–63.
- 610 Shalev, Y., Soucy, S.M., Papke, R.T., Gogarten, J.P., Eichler, J., and Gophna, U. (2018)
611 Comparative analysis of surface layer glycoproteins and genes involved in protein
612 glycosylation in the genus *Haloferax*. *Genes (Basel)* **9**: 172.
- 613 Sleytr, U.B., Schuster, B., Egelseer, E.-M., and Pum, D. (2014) S-layers: principles and
614 applications. *FEMS Microbiol Rev* **38**: 823–864.
- 615 Söding, J., Biegert, A., and Lupas, A.N. (2005) The HHpred interactive server for protein
616 homology detection and structure prediction. *Nucleic Acids Res* **33**: W244–W248.
- 617 Steentoft, C., Vakhrushev, S.Y., Joshi, H.J., Kong, Y., Vester-Christensen, M.B., Schjoldager,
618 K.T.-B.G., *et al.* (2013) Precision mapping of the human O-GalNAc glycoproteome through
619 SimpleCell technology. *EMBO J* **32**: 1478–1488.
- 620 Sumper, M., Berg, E., Mengele, R., and Strobel, I. (1990) Primary structure and glycosylation
621 of the S-layer protein of *Haloferax volcanii*. *J Bacteriol* **172**: 7111–7118.
- 622 Taylor, K.A., Deatherage, J.F., and Amos, L.A. (1982) Structure of the S-layer of *Sulfolobus*
623 *acidocaldarius*. *Nature* **299**: 840–842.
- 624 Veith, A., Klingl, A., Zolghadr, B., Lauber, K., Mentele, R., Lottspeich, F., *et al.* (2009) *Acidianus*,
625 *Sulfolobus* and *Metallosphaera* surface layers: structure, composition and gene expression.
626 *Mol Microbiol* **73**: 58–72.
- 627 Wurch, L., Giannone, R.J., Belisle, B.S., Swift, C., Utturkar, S., Hettich, R.L., *et al.* (2016)
628 Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from
629 a terrestrial geothermal environment. *Nat Commun* **7**: 1–10.

630 Yasuda, M., Oyaizu, H., Yamagishi, A., and Oshima, T. (1995) Morphological variation of new
631 *Thermoplasma acidophilum* isolates from Japanese hot springs. *Appl Environ Microbiol* **61**:
632 3482 LP – 3485.

633