Running head: A novel surface layer protein from Archaea

A novel carbohydrate-binding surface layer protein from the hyperthermophilic archaeon *Pyrococcus horikoshii*

Shuichiro Goda*, Tomoyuki Koga, Kenichiro Yamashita, Ryo Kuriura, and Toshifumi Ueda

Biomolecular Chemistry Laboratory, Graduate School of Engineering, Nagasaki University, Bunkyo-machi 1–14, Nagasaki 852-8521, Japan

*Email: sgoda@nagasaki-u.ac.jp

Abbreviations: GlcNAc, *N*-acetylglucosamine; ITC, isothermal titration calorimetry; S-layer protein, surface-layer protein; TBS, Tris-buffered saline.

Abstract

In Archaea and Bacteria, surface layer (S-layer) proteins form the cell envelope and are involved in cell protection. In the present study, a putative S-layer protein was purified from the crude extract of *Pyrococcus horikoshii* using affinity chromatography. The S-layer gene was cloned and expressed in *Escherichia coli*. Isothermal titration calorimetry analyses showed that the S-layer protein bound *N*-acetylglucosamine and induced agglutination of the gram-positive bacterium *Micrococcus lysodeikticus*. The protein comprised a 21-mer structure, with a molecular mass of 1,340 kDa, as determined using small-angle X-ray scattering. This protein showed high thermal stability, with a midpoint of thermal denaturation of 79°C in dynamic light scattering experiments. This is the first description of the carbohydrate-binding archaeal S-layer protein and its characteristics.

Key words: Archaea, carbohydrate binding, *Pyrococcus horikoshii*, surface-layer protein

Surface layer (S-layer) protein is a component of the cell envelope in Bacteria and Archaea [1,2]. The protein covers the cell surface and acts as protective coat, maintaining the cell shape, trapping molecules and ions, and participating in direct cell division, cell adhesion, and surface recognition [3–5]. The molecular mass of S-layer proteins ranges from approximately 40 to 200 kDa, and thickness of S-layer ranges from 5 to 20 nm [6]. The S-layer lattices can have oblique (p1 or p2), square (p4), or hexagonal (p3 or p6) symmetry with pores (generally in the 2- to 8-nm range) [6]. In Archaea, many species thrive in extreme conditions of temperature, salt concentration, and pH [7]. Especially in thermophiles, S-layer protein is directly exposed to the outer environment, and it is important for cell survival in extreme conditions [8]. Therefore, S-layer proteins from hyperthermophiles have high thermal stability [8]. Furthermore, there is a wide variety of structural features and chemical components in cell envelopes and S-layer proteins of Archaea and Bacteria [9]. Genome analyses have shown that S-layer gene sequences have low homology at the sequence level [5]. Therefore, archaeal S-layer proteins are poorly characterized, and their function and structure remain unknown.

In this study, S-layer protein was purified from crude extracts of the hyperthermophilic archaeon *Pyrococcus horikoshii* using affinity chromatography.

Although S-layer proteins are found in numerous archaeal species, they have mostly been studied in methanogens, halophiles, Sulfolobales and Desulfurococcales [10]. There is less information about S-layer proteins from Thermococcales including *P. horikoshii*. The proteins from *P. horikoshii* are expected to be highly thermostable. We cloned and expressed the gene encoding S-layer protein in *Escherichia coli* and characterized the recombinant protein. The results showed that the novel 21-mer S-layer protein could bind N-acetylglucosamine and induce agglutination of a gram-positive bacterium with high thermal stability. These findings revealed the existence of a carbohydrate-binding S-layer protein in a hyperthermophilic archaeon.

Materials and Methods

Materials. *E. coli* strains JM109 and BL21-CodonPlus(DE3)-RIPL were purchased from Promega (Madison, WI, USA) and Agilent Technologies (Santa Clara, CA, USA), respectively. The plasmid pGFPd was kindly provided by Dr. Drew at Stockholm University [11]. The plasmid pET-3a was purchased from Merck Millipore (Darmstadt, Germany).

Isolation of S-layer proteins from crude extracts of P. horikoshii.

71 Hyperthermophilic archaeon *P. horikoshii* OT-3 (JCM 9975) cells were obtained from
72 the Japanese Collection of Microorganisms (JCM), Wako, Saitama, Japan, and were
73 cultured in medium as reported previously [12]. After culture for 24 h, cells were
74 collected by centrifugation at 8,000 × *g* for 15 min at 4°C and suspended in Tris-
75 buffered saline (TBS) containing 10 mM Tris-HCl (pH 7.5), 0.15 M NaCl, and 10 mM
76 $CaCl_2$. Carbohydrate-binding proteins were purified using affinity chromatography with
77 glucose-Cellufine columns. Glucose-Cellufine was prepared using divinyl-sufone as
78 described by Teichberg et al. [13] with slight modification. Cellufine (JNC Corp.,
79 Tokyo, Japan) was used instead of Sepharose 4B and glucose was used as the
80 carbohydrate. The eluates from glucose-Cellufine column were separated by SDS-
81 PAGE and stained with Coomassie Brilliant Blue R-250. The protein bands were cut
82 from the gel and destained with 30% CH_3CN and 25 mM NH_4HCO_3. The destained gels
83 were reduced using 10 mM DTT and 25 mM NH_4HCO_3, and then alkylated using 55
84 mM ICH_2CONH_2 and 25 mM NH_4HCO_3. The protein in the gel was digested using 10
85 μg/mL lysyl endopeptidase in 5 mM Tris-HCl (pH 8.5) and 4 mM NH_4HCO_3. The
86 peptides were extracted from the gel in 50% CH_3CN and 5% CF_3COOH. The extracted
87 peptides were analyzed using mass spectrometry (Voyager DE PRO; Applied
88 Biosystems, Waltham, MA, USA) and peptide mass fingerprinting according to Mascot

89 database search results (Matrix Science, London, UK).

90

91 *Cloning and expression of the gene encoding the S-layer protein*. The genome
92 of *P. horikoshii* was used as a template for amplification of the gene encoding the
93 putative S-layer protein PH1395.1 using polymerase chain reaction (PCR). The first
94 primer (5′-CATATGCCTTCAGTTCCGAAGGA-3′) had an NdeI restriction site
95 overlapping the 5′ initiation codon, whereas the second primer (5′-
96 GGATCCTCAGAGCTTTGAGATGTACTC-3′) had a unique BamHI restriction site
97 proximal to the 3′ end of the termination codon. Because the first 29 residues of the
98 protein were considered a signal peptide by TMHMM [14], we amplified DNA
99 encoding the 30th amino acid (aa) residue onwards. The resulting DNA fragment was
100 cloned into *E. coli* JM109 cells using the pGEM-T vector pGEM-T/PH1395.1
101 (Promega). One NdeI site within the original sequence was mutated using the
102 QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies) with the primers 5′-
103 GTTCCAACCACACATGACGAGTGGAA-3′ and 5′-
104 TTCCACTCGTCATGTGTGGTTGGAAC-3′. The inserted DNA fragment (1.7 kb) was
105 digested with NdeI and BamHI and ligated with the expression vector pET-3a after
106 linearization with NdeI and BamHI to generate pET3a/PH1395.1. To express a green

fluorescent protein (GFP)-conjugated PH1395.1, the following set of oligonucleotide primers was used to PCR-amplify a PH1395.1 gene fragment lacking a stop codon: 5′-CATATGCCTTCAGTTCCGAAGGA-3′, which introduced a NdeI restriction site overlapping the 5′ initiation codon, and 5′-GGATCCGAGCTTTGAGATGTACT-3′, which introduced a unique BamHI restriction site proximal to the 3′ end of the termination codon. The plasmid pGEM-T/PH1395.1 was used as a template for amplification, and the resulting DNA fragment was cloned into *E. coli* JM109 cells using the pGFPd vector pGFPd/PH1395.1. Cells of the *E. coli* strain BL21-CodonPlus(DE3)-RIPL were transformed with pGFPd/PH1395.1 or pET3a/PH1395.1 and were incubated at 27°C in 200 mL of Luria-Bertani (LB) broth containing 50 μg/mL kanamycin to select pGFPd/PH1395.1 transformants or 50 μg/mL ampicillin to select pET3a/PH1395.1 transformants. Subsequently, transgenes were induced in cells at an optical density (OD) of 0.6 at 600 nm by culturing in the presence of 0.4 mM isopropyl-β-D-thiogalactopyranoside for 16 h.

Purification of the S-layer protein. Collected *E. coli* cells were suspended in TBS containing 10 mM $CaCl_2$ and were disrupted by sonication (Vibra-Cell Processor VC505, Sonics & Materials, Newtown, CT, USA). The sonication time was 40 sec for

125 pellets with 2 grams of wet weight. Crude extracts of cell lysates were obtained by
126 centrifugation at 15,000 × *g* for 20 min at 4°C. The supernatant solution containing the
127 GFP-conjugated protein with a C-terminal His-tag was purified using the COSMOGEL
128 His-Accept column (Nacalai Tesque, Kyoto, Japan). *E. coli* supernatants containing
129 pET-3a/PH1395.1 were applied to an *N*-acetylglucosamine (GlcNAc)-Cellufine column
130 equilibrated with TBS [13]. The column was washed with the same buffer, and the
131 adsorbed proteins were eluted with TBS containing 50 mM GlcNAc (GlcNAc-TBS).
132
133 *Protein determination*. Protein concentrations were determined using the
134 Bradford method with bovine serum albumin as a standard [15].
135
136 *Carbohydrate-binding assays using isothermal titration calorimetry (ITC).*
137 Carbohydrate-binding parameters of the S-layer protein and GFP-conjugated S-layer
138 protein (GFP/S-layer protein) were determined using ITC (iTC200; GE Healthcare,
139 Buckinghamshire, UK) at 30°C in the presence of GlcNAc, *N*-acetylmuramic acid
140 (MurNAc), glucose, N-acetylgalactosamine (GalNAc), glucosamine, and cellobiose.
141 Briefly, protein solutions were dialyzed with 0.1 M 4-(2-hydroxyethyl)-1-
142 piperazineethanesulfonic acid (HEPES) buffer (pH 7.5) containing 10 mM $CaCl_2$, and

143 binding isotherms were fitted to a one-site binding model using the ORIGIN version 7.0
144 software (OriginLab, Northampton, MA, USA).
145
146 *Bacterial agglutination assays*. Bacterial agglutination assays were performed
147 using microscopy (Carton NSZT, Tokyo, Japan) and by assessing changes in OD. In
148 these experiments, the gram-positive bacterium *Micrococcus lysodeikticus* was
149 suspended in TBS solution at a concentration of 0.01 mg/mL, and the S-layer protein
150 was added to 500-µL bacterial suspensions to final concentrations of 2.5, 0.25, 0.025,
151 and 0.0025 mg/mL. The mixtures were incubated at 25°C for approximately 1 h, and
152 cells were observed using a bright-field microscope (Carton NSZT, Tokyo, Japan) at a
153 magnification of 400X.
154 To investigate inhibition of agglutination by carbohydrates, changes in OD at 550
155 nm were determined in *M. lysodeikticus* cell suspensions containing the S-layer protein
156 in the presence and absence of 10 mM GlcNAc using previously reported procedures
157 [16]. Absorbance was measured using a UV-1640 UV/visible spectrophotometer
158 (Shimadzu, Kyoto, Japan).
159
160 *Small-angle X-ray scattering (SAXS) analyses.* SAXS measurements were

161 performed using a Beamline 10C at the Photon Factory, Tsukuba, Japan. The
162 wavelength (λ) and distance between sample detectors were 1.488 Å and approximately
163 90 cm, respectively. Raw data from $Q = 1.3 \times 10^{-1}$ to 4.1×10^{-1} Å^{-1} (the Bragg spacing
164 equivalent to $dB = 483 - 15.3$ Å; $Q = 4\pi\sin\theta/\lambda$, where 2θ is the scattering angle) were
165 measured using a 2-dimensional imaging detector (PILATUS3-300KW; Dectris,
166 Banden-Daettwil, Switzerland). SAXS measurements were performed as reported
167 previously [17], and molecular weights were determined with reference to forward
168 scattered intensities normalized to protein concentrations (*J(0)/C*) of ovalbumin, bovine
169 serum albumin, catalase, and thyroglobulin. Measurements were conducted with 40-μL
170 aliquots of 3.2 mg/mL protein solutions in TBS, and 10 mM GlcNAc was added to
171 determine the effects on the S-layer protein structure.
172
173 *Dynamic light scattering (DLS) analyses*. DLS measurements were performed using
174 the Zetasizer Nano ZS (Malvern Instruments, Malvern, UK) instrument. Z-average size
175 values were measured from 40°C to 90°C, using trend measurements to determine
176 protein denaturation points at a protein concentration of 1.0 mg/mL.
177
178 **Results and Discussion**

Isolation of the S-layer protein from crude extracts of P. horikoshii

We found the putative S-layer protein in crude extracts of *P. horikoshii* when we tried to find thermostable carbohydrate-binding proteins using affinity chromatography. The S-layer protein was eluted with 20 mM ethylenediaminetetra acetate (EDTA)-TBS from the glucose-Cellufine column and peptide mass fingerprinting analyses were performed for the column eluates after separation by SDS-PAGE and in-gel digestion by lysyl endopeptidase (Fig. 1A and 1B). Multiple proteins were observed from the eluates. The protein was identified as open reading frame (ORF) ID PH1394 from the mascot search (Fig. 1C). Whereas, other proteins except the PH1394 were not identified by the mascot search. The top score of PH1394 was 41 and the significance threshold ($p < 0.05$) was 66. This indicated that the identification of PH1394 was not significant. However, the ORFs PH1394 and PH1395 were annotated as N– and C–terminals of this S-layer protein, respectively, at that time. Subsequently, re-annotation of the entire genome of *P. horikoshii* by the National Institute of Technology and Evaluation (NITE) in Tokyo was performed in March 2007, and PH1394 and PH1395 were redesignated as a single protein with the ORF ID PH1395.1. This difference of annotation may have caused the low significance in the mascot search. In contrast, peptides originating from PH1395 were not detected in this mass spectrometry analysis. The amino acid sequence

197 of PH1395.1 had homology with S-layer proteins from *Methanocaldococcus voltae*
198 (45% homology with amino acid positions 12–88 and 36% homology with amino acid
199 498–601) [18] and *Methanocaldococcu jannaschii* (43% homology with aa 1–98 and
200 34% homology with aa 452–604). Additionally, the amino acid sequence of PH1395.1
201 showed high homology to other putative or uncharacterized S-layer proteins from
202 *Pyrococcus abyssi* (75%), *Pyrococcus kukulkanii* (59%), and *Pyrococcus furiosus*
203 (57%). Therefore, PH1395.1 protein may have characteristics that differ from those of
204 S-layer proteins from Methanococci. The existence of the S-layer protein in *P.*
205 *horikoshii* was already reported, as the cell envelope of *P. horikoshii* consists of a
206 complete S-layer enclosing a periplasmic space around the cytoplasmic membrane [19].
207 Therefore, it is reasonable to suppose that this carbohydrate protein designated as
208 PH1395.1 is the S-layer protein of *P. horikoshii*.
209
210 *Expression and purification of the S-layer protein*
211 To measure the amount of produced protein in *E. coli* cells and purify the protein
212 using metal chelate affinity chromatography, the S-layer protein was expressed as a
213 GFP-conjugated and His-tagged protein using the pGFPd vector in *E. coli*. However, the
214 recombinant GFP-conjugated PH1395.1 (GFP/PH1395.1) protein did not bind the

glucose-Cellufine column. Thus, we determined the carbohydrate-binding specificity of the GFP/PH1395.1 protein using ITC analyses of the partially purified protein. These analyses showed that the GFP/PH1395.1 protein bound GlcNAc (Fig. 2). In further experiments, the PH1395.1 protein without GFP was expressed in *E. coli* using the same procedure, purified using a GlcNAc-Cellufine column, and eluted with GlcNAc-TBS. The purity of PH1395.1 was confirmed using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Fig. 3) [20].

Carbohydrate-binding affinity of PH1395.1

ITC measurements of GFP/PH1395.1 showed that the S-layer protein bound to GlcNAc. Therefore, the S-layer protein was purified using a GlcNAc-Cellufine column. The carbohydrate specificity of this protein was analyzed using ITC with the carbohydrates GlcNAc, MurNAc, glucose, GalNAc, glucosamine, and cellobiose. Although the S-layer protein was purified from a crude extract of *P. horikoshii* using glucose affinity chromatography, thermal change was detected only between the protein and GlcNAc and not with the other carbohydrates, including glucose (Fig. 4). The results of affinity chromatography using raffinose cellufine resin showed that the band patterns of SDS-PAGE were similar to those for glucose (data not shown). In addition,

233 the support matrix of Cellufine is cellulose and a small amount of crystalline cellulose
234 remains in Cellufine. Therefore, the S-layer protein may bind to this crystalline
235 cellulose by hydrophobic interactions during purification. The stoichiometry (N)
236 between the S-layer protein and GlcNAc was 1.04 ± 0.101, indicating that one protein
237 molecule bound to one GlcNAc molecule. In addition, the association constant (K_a) and
238 enthalpy change were $1.24 \times 10^5 \pm 5.21 \times 10^4\ M^{-1}$ and −45 ± 7.52 kJ/mol, respectively,
239 and were similar to those reported for lectin and carbohydrate [21]. Some S-layer
240 proteins of gram-positive bacteria bind to peptidoglycan, which contains GlcNAc and
241 MurNAc [22]. Moreover, the S-layer protein of the thermophilic bacterium
242 *Thermoanaerobacter kivui* binds to peptidoglycan targets via its surface layer homology
243 domain [23]. Archaea have various types of cell wall including pseudomurein [24].
244 However, the cell envelope of *P. horikoshii* resembles that of *P. furiosus* [19], which
245 has no pseudomurein layer and consists of only S-layer protein in the cell envelope [24].
246 Thus, the S-layer protein from *P. horikoshii* is supposed to bind to GlcNAc that exists
247 out of the cell. During the exponential phase of growth, cells of *P. horikoshii* are often
248 seen in pairs, and they aggregate increasingly as the culture ages [19]. Therefore, S-
249 layer protein may bind to the carbohydrated S-layer protein of other cells, and this leads
250 to the aggregation of the cells.

251

252 *Bacterial agglutination assays*

253 Because the S-layer protein showed binding affinity for GlcNAc, it may useful
254 for the detection of bacteria by a highly thermostable protein. Therefore, we
255 investigated associations between the S-layer protein and the gram-positive bacterium
256 *M. lysodeikticus*. In bright-field microscopy experiments (Fig. 5), the S-layer protein
257 agglutinated *M. lysodeikticus* in a concentration-dependent manner. In addition, to
258 determine whether the S-layer protein recognized GlcNAc on the surface of *M.*
259 *lysodeikticus*, we measured the inhibition of agglutination using a spectrophotometer. In
260 these experiments, the OD of the mixture of the S-layer protein and *M. lysodeikticus*
261 gradually decreased by 10% over 150 min (Fig. 6). Moreover, in the presence of 10 mM
262 GlcNAc, the OD of the mixture was similar to that in the absence of the protein (Fig. 6),
263 suggesting that the S-layer protein bound to GlcNAc on the surface of *M. lysodeikticus*.
264 However, multiple binding sites were assumed to be required for agglutination,
265 suggesting that the S-layer protein formed oligomers.

266

267 *SAXS*

268 SAXS analyses of the S-layer protein revealed radius of gyration (*Rg)* and

maximum particle (D_{max}) values of 67.8 ± 0.45 and 237 ± 9.0 Å, respectively, as indicated by Guinier plots and $p(r)$ functions, which were Hankel transformations of the scattering curve (Fig. 7A) [25,26]. The molecular mass of the S-layer protein was estimated at 1,344 ± 17 kDa, according to forward scattering intensities ($J(0)$) normalized to protein concentrations C ($J(0)/C$) [26,27]. The molecular mass of the monomeric recombinant S-layer protein is 63.5 kDa. These data indicated that the S-layer protein formed a 21-mer. The ensuing Kratky plot of the S-layer protein showed three peaks (Fig. 7B), typical of oligomeric proteins [17,28]. Moreover, these analyses were impervious to the presence of GlcNAc. S-layer proteins form oligomeric structures with oblique (p1 or p2), square (p4), or hexagonal (p3 or p6) lattice symmetry [6], and hexagonal symmetry is predominant among archaeal S-layer proteins [29]. SAXS analyses of the S-layer protein showed the formation of a 21-mer, suggesting that the small unit of the S-layer protein was further assembled into larger 21-mer oligomers.

Thermal stability of the S-layer protein

To determine the thermal stability of the S-layer protein from *P. horikoshii*, we performed DLS experiments between 40°C and 90°C. Two peak sizes in the radius of the S-layer protein were observed at 40°C. Scattering was mainly from 7-nm particles

287 and a few particles were observed around 100 nm. After heating, the amount of
288 scattering from 7-nm particles decreased and the other peak was observed around 200
289 nm (data not shown). To analyze the change in the number and size of the S-layer
290 protein, Z-average size was plotted (Fig. 8). The Z-average size is an overall average
291 size of all the intensities based on a specific fit to the raw correlation function data. This
292 value aids in determination of the average particle size of a population and is insensitive
293 to experimental noise. The Z-average size of the S-layer protein increased between
294 75°C and 85°C, reflecting the denaturation of the protein (Fig. 8). As a result, the
295 middle point temperature of thermal denaturation was 79 ± 0.5°C, indicating that the S-
296 layer protein from *P. horikoshii* was a highly thermostable S-layer protein. However,
297 79°C is lower than the optimum growth temperature (98°C) of *P. horikoshii*. S-layer
298 proteins form the cell envelope and interact with the cell surface. The interaction
299 between the S-layer protein and cell surface may strengthen thermal stability. In
300 addition, the protein produced in *E. coli* has no glycosylation and an extra methionine
301 remains on the N-terminus of the protein. These aspects differ from those of the native
302 protein. Some S-layer proteins from archaea are known as glycosylated proteins [30].
303 The S-layer protein may be glycosylated in its native form. Glycosylation and the
304 additional methionine at the N-terminus position may affect the thermal stability of this

protein.

In conclusion, we identified and characterized the highly thermostable carbohydrate-binding protein PH1395.1 from *P. horikoshii*. PH1395.1 was annotated as a putative S-layer protein in whole genome analyses after isolation from crude extracts of *P. horikoshii* using a glucose-Cellufine column. Moreover, ITC analyses showed that the S-layer protein bound GlcNAc with a stoichiometry of 1.04 ± 0.101, indicating one-to-one binding. We demonstrated that the S-layer protein formed 21-mers with multiple carbohydrate-binding sites and caused agglutination of gram-positive bacteria. However, it is unclear what binds to the S-layer protein in nature.

S-layer proteins have regular lattices with pores of identical morphology on solid supports and liquid-surface interfaces [6]. Therefore, they are used in biotechnological applications such as the production of isoporous ultrafiltration membranes, the construction of supporting structures for the controlled immobilization, and incorporation of functional molecules (antigens, antibodies, ligand, and enzymes) required for biosensors [31]. Thermostable and GlcNAc-binding S-layer protein may be a useful tool for biotechnology and biomedical applications.

Author contributions

S.G. designed research; T.K., K.Y., R.K. and T.U. performed the experiments. S.G.

analyzed data; S.G. wrote the article.

Acknowledgements

This study was performed under the approval of the Photon Factory Advisory Committee (Proposal 2013G027, 2013G706, 2015G015, and 2015G613). We would like to thank Editage (www.editage.jp) for English language editing.

Funding

This work was supported by the Japan Society for the Promotion of Science under a Grant-in-aid for Scientific Research [(C) 25450133] to S. G.

References

[1] Messner P, Sleytr UB. Crystalline bacterial cell-surface layers. Adv Microbiol Physiol. 1992;33:213–275.

[2] Sleytr UB, Messner P. Crystalline surface layers on bacteria. Annu Rev Microbiol. 1983;37:311–339.

[3] Beveridge TJ, Pouwels PH, Sára M, et al. Functions of S-layers. FEMS Microbiol Rev. 1997;20:99–149.

[4] Sleytr UB, Beveridge TJ. Bacterial S-layers. Trends Microbiol. 1999;7:253–260.

343 [5] Sara M, Sleytr UB. S-Layer proteins. J Bacteriol. 2000;182:859–868.

344 [6] Sleytr UB, Huber C, Ilk N, et al. S-layers as a tool kit for nanobiotechnological

345 applications. FEMS Microbiol Lett. 2007; 267:131–144.

346 [7] Cowan DA. Biotechnology of archaea. Trends Biotechnol. 1992;10:315–323.

347 [8] Eichler J. Facing extremes: archaeal surface-layer (plyco)proteins. Microbiology.

348 2003;149:3347–3351.

349 [9] Beveridge TJ. Ultrastructure, chemistry, and function of the bacterial wall. Int Rev

350 Cytol. 1981;12:229–317.

351 [10] Rodrigues-Oliveira T, Belmok A, Vasconcellos D, et al. Archaeal S-layers:

352 Overview and current state of the art. Front Microbiol. 2017;8:2597.

353 [11] Drew D, Lerch M, Kunji E, et al. Optimization of membrane protein

354 overexpression and purification using GFP fusions. Nat Methods. 2006;4:303–313.

355 [12] Kawakami R, Sakuraba H, Kamohara S, et al. Oxidative stress response in an

356 anaerobic hyperthermophilic archaeon: presence of a functional peroxiredoxin in

357 *Pyrococcus horikoshii*. J Biochem. 2004;136:541–547.

358 [13] Teichberg VI, Aberdam D, Erez U, et al. Affinity-repulsion chromatography.

359 Principle and application to lectins. J Biol Chem. 1988;263:14086–14092.

360 [14] Krogh A, Larsson B, von Heijne G, et al. Prediction transmembrane protein

361 topology with a hidden Markov model: application to complete genomes. J Mol Biol.
362 2001;305:567–580.
363 [15] Bradford MM. A rapid and sensitive method for the quantitation of microgram
364 quantities of protein utilizing the principle of protein-dye binding. Anal Biochem.
365 1976;72:248–254.
366 [16] Shimotahira N, Oogai Y, Kawada-Matsuo M, et al. The surface layer of *Tannerella*
367 *forsythia* contributes to serum resistance and oral bacterial coaggregation. Infect
368 Immun. 2013;81:1198–1206.
369 [17] Hiragi Y, Seki Y, Ichimura K, et al. Direct detection of the protein quaternary
370 structure and denatured entity by small-angle scattering: guanidine hydrochloride
371 denaturation of chaperonin protein GroEL. J Appl Crystallogr. 2002;35:1–7.
372 [18] Konisky J, Lynn D, Hoppert M, et al. Identification of the *Methanococcus voltae*
373 S-layer structural gene. J Bacteriol. 1994;176:1790–1792.
374 [19] Gonzalez JM, Masuchi Y, Robb FT, et al. *Pyrococcus horikoshii* sp. nov., a
375 hyperthermophilic archaeon isolated from a hydrothermal vent at the Okinawa Trough.
376 Extremophiles. 1998;2:123–130.
377 [20] Laemmli UK. Cleavage of structural proteins during the assembly of the
378 bacteriophage T4. Nature. 1970;227:680–685.

379 [21] Dam TK, Brewer CF. Thermodynamic studies of lectin-carbohydrate interactions
380 by isothermal titration calorimetry. Chem Rev. 2002;102:387–429.
381 [22] Engelhardt H, Peters J. Structural research on surface layers: a focus on stability,
382 surface layer homology domains, and surface layer-cell wall interactions. J Struct Biol.
383 1998;124:276–302.
384 [23] Lupas A, Engelhardt H, Peters J, et al. Domain structure of the *Acetogenium kivui*
385 surface layer revealed by electron crystallography and sequence analysis. J Bacteriol.
386 1994;176:1224–1233.
387 [24] Klingl A. S-layer and cytoplasmic membrane – exceptions from the typical
388 archaeal cell wall with a focus on double membranes. Front. Microbiol. 2014;5:624.
389 [25] Guinier A, Fournet G. Small-angle scattering of X-rays. New York: Chapman &
390 Hall; 1955.
391 [26] Kratky O, Porod G, Kahovec L. Einige Neuerungen in der Technik und
392 Auswertung von Röntgen-Kleinwinkelmessungen. Z Elektrochem. 1951;55:53–59.
393 [27] Glatter O, Kratky O. Small-angle X-ray scattering. New York: Academic Press;
394 1982.
395 [28] Higurashi T, Hiragi Y, Ichimura K, et al. Structural stability and solution structure
396 of chaperonin GroES heptamer studied by synchrotron small-angle X-ray scattering. J

397 Mol Biol. 2003;333:605–620.

398 [29] Howorka S. Rationally engineering natural protein assemblies in

399 nanobiotechnology. Curr Opin Biotechnol. 2011;22:485–491.

400 [30] Yurist-Doutsch S, Chaban B, VanDyke DJ, et al. Sweet to the extreme: protein

401 glycosylation in archaea. Mol Microbiol. 2008;68:1079–1084.

402 [31] Raff J, Matys S, Suhr M, et al. S-layer-based nanocomposites for industrial

403 applications. Adv Exp Med Biol. 2016;940:245–279.

404 **Figure Legends**

405 Fig. 1. Isolation and identification of the S-layer protein from the crude extracts of

406 *Pyrococcus horikoshii*.

407 (A) SDS-PAGE of the two fractions at the peak position of eluates from the glucose-

408 Cellufine column. The arrow indicates the protein band used for peptide mass

409 fingerprinting. (B) Peptide mass spectra of glucose column eluates were digested with

410 lysyl endopeptidase. (C) Amino acid sequence of PH1394; matched peptides are shown

411 in bold and are underlined. Triangles represent cleavage sites observed by the mass

412 spectrometry.

413 Fig. 2. Isothermal titration calorimetry of green fluorescent protein (GFP)-labeled

414 surface (S)-layer protein in the presence of *N*-acetylglucosamine (GlcNAc).

415 S-layer protein was conjugated with GFP and titrated with GlcNAc at 25°C.

416 Titration kinetics and integrated binding isotherms are presented in upper and lower

417 panels, respectively.

418 Fig. 3. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of recombinant

419 surface (S)-layer protein after *N*-acetylglucosamine (GlcNAc)-Cellufine affinity

420 chromatography.

421 Lanes: 1, molecular weight marker; 2, flow-through of the column; 3–8, eluates

422 following addition of 50 mM GlcNAc-Tris-buffered saline (TBS).

423 Fig. 4. Isothermal titration calorimetry of S-layer protein in the presence of *N*-

424 acetylglucosamine (GlcNAc).

425 Surface (S)-layer protein was titrated with GlcNAc at 25°C. Titration kinetics and

426 integrated binding isotherms are presented in upper and lower panels, respectively. The

427 molar ratio refers to the protein monomer. Changes in binding enthalpy (ΔH) and

428 association constants (*K*a) were determined using non-linear regression analyses of

429 integrated data with a one-site binding model.

430 Fig. 5. Agglutination of *Micrococcus lysodeikticus* by surface (S)-layer protein.

431 Bacterial agglutination with indicated S-layer protein concentrations was visualized

432 using bright-light microscopy at a magnification of 400X.

433 Fig. 6. Agglutination of surface (S)-layer protein with *Micrococcus lysodeikticus*.

434 *M. lysodeikticus* concentrations were adjusted to an optical density (OD) of 1.0 at

435 550 nm. The concentration of S-layer protein was 2.5 mg/mL, and agglutination was

436 indicated by decreased absorbance. Agglutination of *M. lysodeikticus* by S-layer protein

437 is indicated by squares, and the densities of *M. lysodeikticus* controls without S-layer

438 protein are indicated by circles. Agglutination was inhibited using 10 mM *N*-

439 acetylglucosamine (GlcNAc) (triangles).

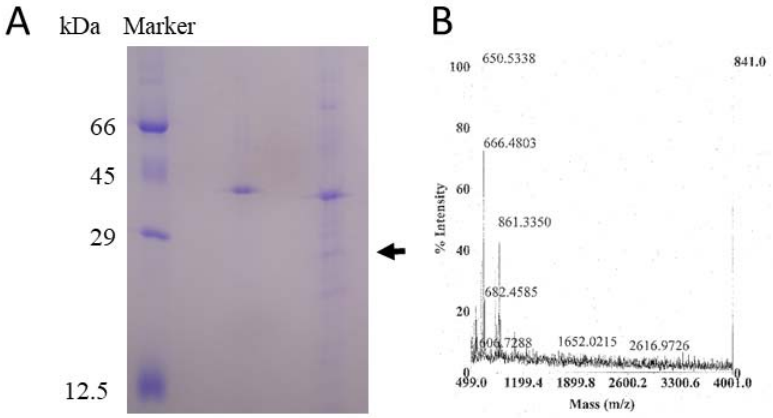
Fig. 7. Small-angle X-ray scattering of the surface (S)-layer protein.

The size and molecular mass of S-layer protein were measured using small-angle X-ray scattering. Guinier (A) and Kratky (B) plots for S-layer protein were generated in the absence (solid line) and presence of *N*-acetylglucosamine (GlcNAc) (broken line).

Fig. 8. Effect of temperature on the size of the surface (S)-layer protein.

Z-average values of the S-layer protein were measured using dynamic light scattering between 40°C and 90°C.

Fig. 1



C

1 MEVFDMKVKK IAALAVGAAV AGATLGLASA QPEVPSVPKD FFVKDGEPNV

51 KIVVGSEGAA MDVVSAADIA AAIGSLLYTE KDVEVSDVSV VVKKDITDDP

101 EDIPVFNNFY PDSDQVFKEG TDELDKIPAW WNGGYDEDGN PVFNEDLSHS

151 AWADGTYSGP YELELDNVAS YDDYELTGIL TLKNITLKEI AEKEVDWKDI

201 NDFKDLTVVI DSVVLNTTLE VYNWTKTETE KDPVTGEETT VTKWEISNVQ

251 PGSGYSIKET VSEGATAGET LKVLDKEFPI VFVGKAIKSG QIVDQEDSSS

301 MVPIMEQNSM TRDRK

Fig. 2

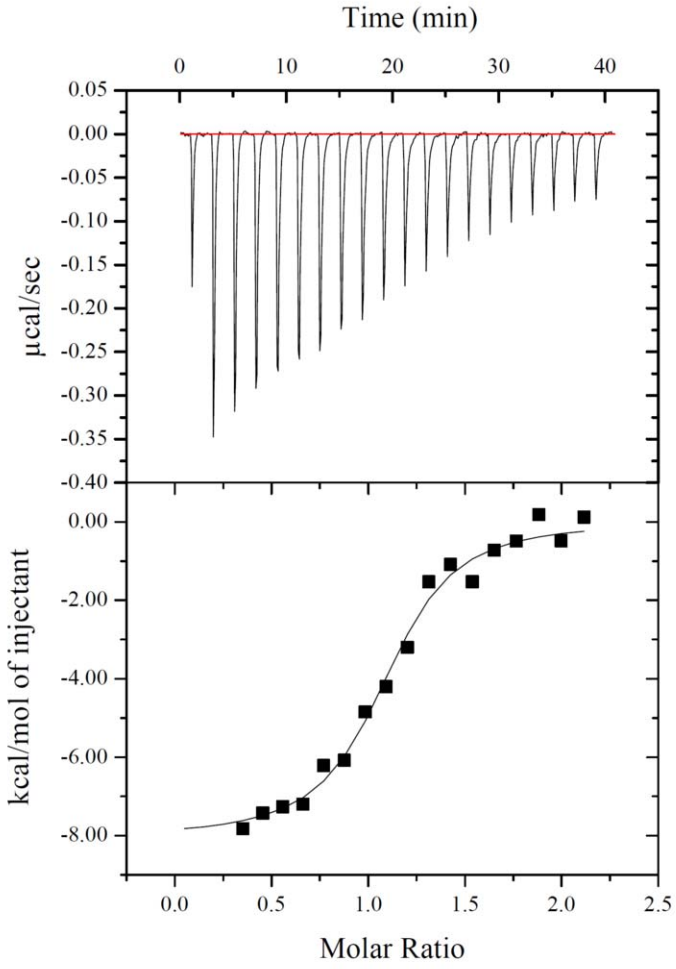

Time (min)
0
10
20
30
40
0.05
0.00
-0.05
-0.10
-0.15
-0.20
-0.25
-0.30
-0.35
-0.40
µcal/sec
0.00
-2.00
-4.00
-6.00
-8.00
kcal/mol of injectant
0.0
0.5
1.0
1.5
2.0
2.5
Molar Ratio

Fig. 3

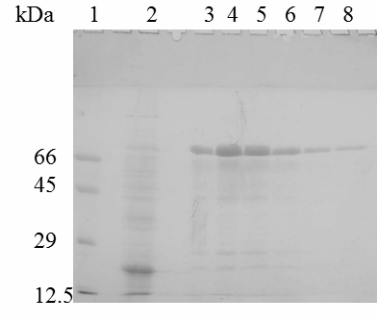


Fig. 4

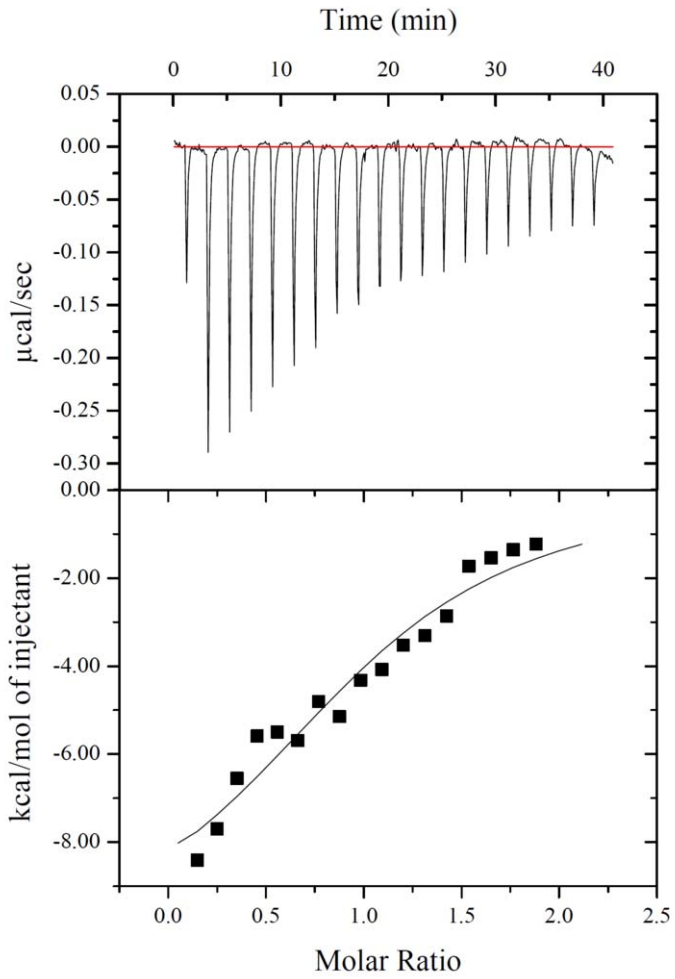

Time (min)
0
10
20
30
40
0.05
0.00
-0.05
-0.10
-0.15
-0.20
-0.25
-0.30
µcal/sec
0.00
-2.00
-4.00
-6.00
-8.00
kcal/mol of injectant
0.0
0.5
1.0
1.5
2.0
2.5
Molar Ratio

Fig. 5

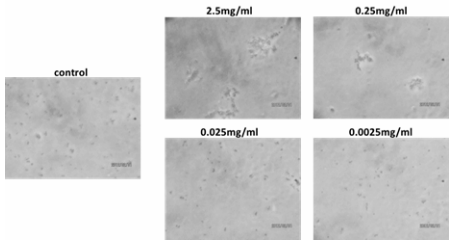


Fig. 6

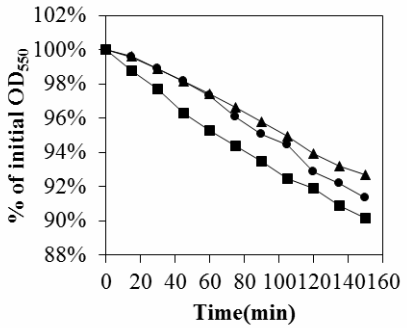

102%
100%
98%
96%
94%
92%
90%
88%
% of initial OD550
0 20 40 60 80 100120140160
Time(min)

Fig. 7

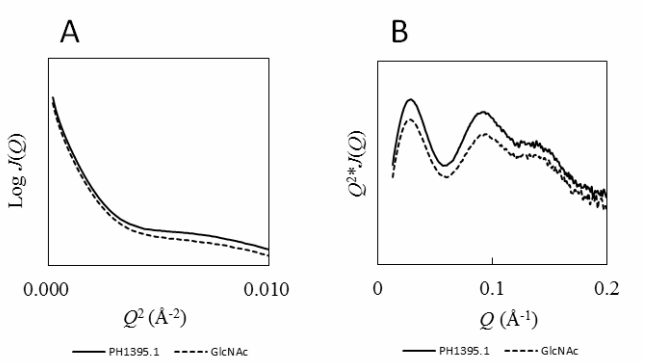

A
B
Log I(Q)
0.000
0.010
Q² (Å⁻²)
PH1395.1
GlcNAc
Q²*I(Q)
0
0.1
0.2
Q (Å⁻¹)
PH1395.1
GlcNAc

Fig. 8