The first biantennary bacterial secondary cell wall polymer and its influence on S-layer glycoprotein assembly

Christian STEINDL*, Christina SCHÄFFER†¹, Thomas WUGEDITSCH†, Michael GRANINGER†, Irena MATECKO*, Norbert MÜLLER* and Paul MESSNER†

*Institut für Chemie, Johannes-Kepler-Universität Linz, A-4040 Linz, Austria, and †Zentrum für Ultrastrukturforschung und Ludwig Boltzmann-Institut für Molekulare Nanotechnologie, Universität für Bodenkultur Wien, A-1180 Wien, Austria

The cell surface of *Aneurinibacillus thermoaerophilus* DSM 10155 is covered with a square surface (S)-layer glycoprotein lattice. This S-layer glycoprotein, which was extracted with aqueous buffers after a freeze–thaw cycle of the bacterial cells, is the only completely water-soluble S-layer glycoprotein to be reported to date. The purified S-layer glycoprotein preparation had an overall carbohydrate content of 19%. Detailed chemical investigations indicated that the S-layer O-glycans of previously established structure accounted for 13% of total glycosylation. The remainder could be attributed to a peptidoglycan-associated secondary cell wall polymer. Structure analysis was performed using purified secondary cell wall polymer–peptidoglycan complexes. NMR spectroscopy revealed the first biantennary secondary cell wall polymer from the domain Bacteria, with the structure α -D-GlcpNAc-(1→3)- β -D-ManpNAc-(1→4)- β -D-GalpNAc-(1→3)- α -D-GlcpNAc-(1→3)- β -D-ManpNAc-(1→4)- β -D-GalpNAc-(1→3)- α -D-GlcpNAc-(1→4)-[α -D-GlcpNAc-(1→3)- β -D-ManpNAc-(1→4)- β -D-GalpNAc-(1→3)- α -D-GlcpNAc-(1→3)]- β -D-ManpNAc-(1→3)- α -D-GlcpNAc-(1→3)- β -D-ManpNAc-(1→3)- α -D-GlcpNAc-(1→3)- α -D-

GlcpNAc-(1→O)-PO₂[−]-O-PO₂[−]-(O→6)-MurNAc- (where MurNAc is *N*-acetylmuramic acid). The neutral polysaccharide is linked via a pyrophosphate bond to the C-6 atom of every fourth *N*-acetylmuramic acid residue, in average, of the A1 γ -type peptidoglycan. *In vivo*, the biantennary polymer anchored the S-layer glycoprotein very effectively to the cell wall, probably due to the doubling of motifs for a proposed lectin-like binding between the polymer and the N-terminus of the S-layer protein. When the cellular support was removed during S-layer glycoprotein isolation, the co-purified polymer mediated the solubility of the S-layer glycoprotein *in vitro*. Initial crystallization experiments performed with the soluble S-layer glycoprotein revealed that the assembly property could be restored upon dissociation of the polymer by the addition of poly(ethylene glycols). The formed two-dimensional crystalline S-layer self-assembly products exhibited the same lattice symmetry as observed on intact bacterial cells.

Key words: bacteria, biantennary peptidoglycan-associated polymer, NMR spectroscopy, structure determination.

INTRODUCTION

In addition to its cell-shape-determining/maintaining function, the cell wall of Gram-positive bacteria serves as an attachment site for various proteins, which may interact with the bacterial environment. In general, mechanisms for immobilization of proteins involve the covalent attachment to the peptidoglycan or the non-covalent binding to either the peptidoglycan or secondary cell wall polymers (SCWPs) [1]. The most common surface structures found on bacteria are monomolecular crystalline arrays formed of (glyco)proteinaceous subunits, which are termed surface (S)-layers (for reviews see [2,3]). The constituting subunits are held together and on to the supporting peptidoglycan layer by weak non-covalent forces. Extraction of S-layers usually requires the presence of dissociating agents (LiCl or EDTA) or chaotropic denaturants (guanidinium chloride or urea). Upon removal of these agents, the monomeric subunits usually spontaneously reassemble into sheet-like or cylindrical so-called S-layer self-assembly products (for review see [4]).

Preliminary insight has been gained only recently into the manner by which S-layers are attached to the Gram-positive cell surface. In the bacteria studied, there is strong evidence that

attachment is mediated via distinct types of peptidoglycan-associated SCWPs. They, in turn, recognize the N-terminal region of the respective S-layer protein [5,6]. Either S-layer homologous (SLH) domains or basic amino acids are involved presumably in a lectin-like binding between the S-layer protein and the SCWP. Basic amino acids are typically present at a high density in the S-layer N-terminal region, and they are known for their direct interaction with carbohydrates via electrostatic interactions and hydrogen bonds.

The structures of three such polymers from S-layer-carrying organisms have been elucidated so far. They constitute between 10 and 20% (w/w) of the peptidoglycan sacculus and include the acidic polysaccharide structures [$\rightarrow$ 4)- β -D-ManpA-2,3-(NAc)₂-(1→6)- α -D-Glcp-(1→4)- β -D-ManpA-2,3-(NAc)₂-(1→3)- α -D-GlcpNAc-(1→)_n ≈ 5 of *Geobacillus stearothermophilus* wild-type strains [7] and [$\rightarrow$ 3)(Pyr4,6)- β -D-ManpNAc-(1→4)- β -D-GlcpNAc-(1→)_n ≈ 12 (where Pyr is pyruvic acid) isolated from *Bacillus sphaericus* CCM 2177 [8] and *Paenibacillus alvei* CCM 2051 [9]. Furthermore, neutral SCWPs of identical composition and structure have been isolated from the peptidoglycan of *Thermoanaerobacterium thermosaccharolyticum* strains

Abbreviations used: DQF-COSY, double quantum-filtered correlated spectroscopy; ESI-TOF MS, electrospray-ionization time-of-flight MS; Hepp, heptopyranose; HMBC, heteronuclear multi-bond correlation spectroscopy; HSQC, heteronuclear single quantum correlation spectroscopy; *m*-A₂pm, diaminopimelic acid; MME 550, poly(ethylene glycol) 550 monomethylether; MurNAc, *N*-acetylmuramic acid; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; NS, number of scans; PEG 20000, poly(ethylene glycol) 20000; Rhap, rhamnopyranose; ROE, rotating-frame Overhauser effect; ROESY, rotating-frame nuclear Overhauser effect spectroscopy; S-layer, surface layer; SCWP, secondary cell wall polymer; SLH, S-layer homologous.

¹ To whom correspondence should be addressed (e-mail CRS@edv1.boku.ac.at).

E207-71 and D120-70 (C. Schäffer, H. Kählig, S. Zayni, A. Scheberl and P. Messner, unpublished work). Typically, these polysaccharides are composed of repeating units, resulting in variable chain lengths, and exhibit an overall molecular mass in the range of 4–6 kDa. As suggested from detailed analysis of the SCWPs of *G. stearothermophilus* NRS 2004/3a [7] and *P. alvei* CCM 2051 [9], they are bound directly via a pyrophosphate bridge to the C-6 atom of muramic acid residues of the peptidoglycan sacculus. Therefore, according to the scheme of Araki and Ito [10], these polymers should be classified in the group of 'neutral or acidic polysaccharides that cannot be affiliated to teichoic or teichuronic acids.'

Aneurinibacillus thermoaerophilus DSM 10155 is an exception among Gram-positive, S-layer-carrying organisms. It is the only known strain that possesses a glycosylated S-layer protein lattice, which is completely soluble in water and aqueous buffer systems during isolation. Furthermore, detailed analysis of the S-layer glycosylation revealed a heptose residue, which is a typical constituent of lipopolysaccharides of Gram-negative bacteria in the $[\rightarrow 4)\text{-}\alpha\text{-L-Rhap-(1}\rightarrow 3)\text{-}\beta\text{-D-glycero-D-manno-Hepp-(1}\rightarrow)_n\approx 15$ S-layer glycan repeats (where Rhap is rhamnopyranose and Hepp is heptopyranose) [11, 11a]. On the S-layer polypeptide, the O-glycans are bound via a variable $[\alpha\text{-L-Rhap-(1}\rightarrow 3)]_{n=0-2}\text{-}\beta\text{-D-GalpNAc-(1}\rightarrow$ core saccharide to distinct serine and threonine residues [12]. During S-layer glycoprotein isolation, minor amounts of a second glycan structure could be separated from the S-layer glycoprotein in the final purification step. As a corresponding glycosidic linkage could not be isolated, it was tempting to assume that this structure was a peptidoglycan-associated SCWP [13].

The present work aims at the complete elucidation of the primary structure of the SCWP of *A. thermoaerophilus* DSM 10155 and the identification of its linkage to the peptidoglycan sacculus. Furthermore, due to its water solubility, initial crystallization experiments were performed with a purified S-layer glycoprotein preparation of that organism. A putative role of the SCWP in the assembly process of the S-layer glycoprotein is proposed. In this context, a possible scenario for the *in vivo* biological function of the SCWP is discussed.

EXPERIMENTAL

Analytical and other methods

The amino acid composition, including identification of muramic and diaminopimelic acids, and the carbohydrate composition were analysed as described previously [14]. The amount of cross-linking of peptidoglycan was determined by the nitrous acid procedure [15]. Carbohydrates in column effluents were detected using the thymol reagent [16]. Phosphate was determined colorimetrically [17]. Protein was determined by the method of Bradford [18]. SDS/PAGE of S-layer glycoprotein was carried out according to Laemmli using either 10 or 7.5% gels, and protein and carbohydrates were visualized with silver staining [19] and periodic acid–Schiff reagent [14], respectively. Electron microscopy of negatively stained and ultrathin-sectioned preparations of assembled S-layer glycoprotein was described elsewhere [20].

Growth of bacteria

A. thermoaerophilus DSM 10155 was grown under steady-state conditions in a 15 litre Biostat E fermenter (Braun, Melsungen, Germany) with an aeration rate of 6 litres/min and a dilution rate of $D = 0.3\text{ h}^{-1}$, at 55 °C and $\text{pH} = 7.0 \pm 0.15$ in modified SVIII medium [21]. Cells were harvested by continuous centrifugation

at 16000 *g* and 4 °C. The biomass was stored at –20 °C until use.

Isolation and purification of the SCWP–peptidoglycan complex

Thawed cell pellets from *A. thermoaerophilus* DSM 10155 ($\approx 100\text{ g}$, wet weight) were suspended in 250 ml of 50 mM Tris/HCl, pH 7.2, with gentle stirring for 30 min on ice, and centrifuged at 27200 *g* for 15 min. Cell pellets were washed six times with buffer, and the supernatants were collected for isolation of the S-layer glycoprotein as described below. For isolation of the SCWP–peptidoglycan complex, the pellets were resuspended in the same buffer and cells were broken by ultrasonication on ice, using three cycles of 3 min with intermittent cooling for 5 min (Branson 450 sonifier, macrotip, 90% output; Inula, Wien, Austria). Cell debris was separated from unbroken cells by centrifugation. Subsequently, broken cells were mixed with 250 ml of cold 1 M NaCl and the fraction sedimenting between 3200 and 48400 *g* for 25 min was collected. Further isolation of the SCWP–peptidoglycan complex from that fraction was carried out as described by Schäffer et al. [9], with slight modifications.

For purification, the crude preparation was applied to columns of Sephadex G-50 (2.5 cm $\times$ 100 cm; Amersham Biosciences, Uppsala, Sweden), eluted with 50 mM NH_4HCO_3 , and Bio-Gel P-30 (1.5 cm $\times$ 120 cm; Bio-Rad), eluted with 0.1 mM NaCl. The carbohydrate-containing pool was finally purified on a semi-preparative RP-18 column (Nucleosil 120-3C₁₈, 8 mm $\times$ 125 mm; Macherey-Nagel, Düren, Germany), equipped with an RP-18 guard column (Nucleosil 120-3C₁₈, 8 mm $\times$ 40 mm; Macherey-Nagel), using an acetonitrile/water gradient as described in [7]. Fractions were combined according to the UV elution profile at 280 nm and the carbohydrate content. The pool containing the SCWP–peptidoglycan complex was lyophilized and investigated by NMR spectroscopy.

NMR spectroscopy

A solution of the SCWP–peptidoglycan complex (15 mg/ml), containing 0.5 mg/ml NaN_3 , in 99.95% $^2\text{H}_2\text{O}$, pH 5.1, was used for most experiments. The NMR spectra were recorded at 313 K, non-spinning in 5 mm tubes at 500.13 MHz for ^1H and at 125.77 MHz for ^{13}C on a Bruker Avance DRX 500 spectrometer, using a 5 mm TXI probe with triple axis gradient coils and 100 W ^1H and 300 W heteronuclear linear amplifiers. ^1H -only experiments [rotating-frame nuclear Overhauser effect spectroscopy (ROESY), COSY, TOCSY] were run for increased resolution and dynamic range with a 1.5 mg/ml sample. Minor chemical shift changes were observed upon dilution. The chemical shifts reported here refer to the 15 mg/ml solution. Spectra were referenced to internal sodium trimethylsilyl propionate- d_4 at 0.0 p.p.m. for ^1H , and at –2.0 p.p.m. for ^{13}C . Typical 90° hard-pulse durations were 8 μs (^1H) and 18 μs (^{13}C); 90° pulses in decoupling sequences were set to 100 μs [^1H , wideband alternating-phase low-power technique for zero residue splitting ('WALTZ')] and 80 μs [^{13}C , globally optimized alternating-phase rectangular pulse (GARP)]. Repetition intervals and gradient durations were 1.5 ms each, with 0.15 ms recovery time. Standard, mostly gradient-enhanced ($\approx 60\text{ T/m}$ maximum gradient strength on a Bruker Acustar II unit), phase-sensitive pulse sequences as provided by the manufacturer were used [one-dimensional ^1H -NMR with pre-saturation of residual water signal; heteronuclear single quantum correlation spectroscopy (HSQC) and heteronuclear multi-bond correlation spectroscopy (HMBC), without ^1H decoupling for ^1H -coupled

spectra; NOESY mixing times were 300–500 ms, offset compensated ROESY used 200 ms spin lock at 3.2 kHz, TOCSY experiments used MLEV mixing sequences 22.55 μ s 90° pulses and binominal pulse water suppression). ^{31}P -NMR experiments were run as described previously [9]. Typical data sizes were 4 K $\times$ 512 ($f_2 \times f_1$) data points with single zero filling. For high-resolution HSQC spectra, 2 K data points were acquired in the f_1 dimension. Gaussian window functions were interactively optimized for the required resolution.

Electrospray-ionization time-of-flight MS (ESI-TOF MS) analysis

Of the NMR sample (15 mg/ml), 50 μ l were diluted with 0.5 ml of water and lyophilized twice to back-exchange ^1H for ^2H , and finally dissolved in 100 μ l of water. Of that sample, 10 μ l were mixed with 100 μ l of a solution of water/isopropanol/triethylamine (80:20:2; by vol.). Mass spectra were recorded on a Mariner time-of-flight electrospray ionization mass spectrometer (Perkin Elmer, Wellesley, MA, U.S.A.) in negative-ion mode using a nano-spray sample interface. Spray capillaries coated with Au/Pd (inner diameter, 0.69 mm), purchased from Protana (Odense, Denmark), were used with the medium length of the outermost tip. The sample was sprayed at 25 °C with a curtain gas flow of 1.2 litres/min. The spray tip potential was set at 1310 V, the nozzle potential was 80 V and the nozzle temperature was 90 °C. Software for data recording (Mariner Workstation 4.0) and data processing (Data Explorer 3.2) was supplied by Applied Biosystems (Foster City, CA, U.S.A.). We used 15 spectra for averaging after external calibration. Due to a low signal-to-noise ratio and the absence of an internal standard the mass accuracy was ± 2.6 .

Isolation and purification of S-layer glycoprotein

After washing of the biomass, the combined lyophilized supernatants were used as a source of S-layer glycoprotein. The material was solubilized at a concentration of 10 mg/ml in 50 mM Tris/HCl, 5 mM EDTA, pH 7.2, containing 40% $(\text{NH}_4)_2\text{SO}_4$ (buffer A). Precipitated protein was removed by centrifugation at 100 000 g for 45 min, and S-layer glycoprotein was extracted twice upon stirring for 90 min at 25 °C [11,11a]. The soluble S-layer glycoprotein was purified from the supernatant by column chromatography in a four-step procedure.

Step 1 was interfacial salting out. The supernatant from $(\text{NH}_4)_2\text{SO}_4$ precipitation was applied in 50 ml aliquots to a Sepharose CL-6B column (1.6 cm $\times$ 20 cm; Amersham Biosciences) equilibrated with buffer A at a flow rate of 0.6 ml/min, and the elution profile was recorded at 280 nm. Non-precipitating S-layer glycoprotein was eluted with 50 ml of buffer A. The column was regenerated with 60 ml of 50 mM Tris/HCl and 5 mM EDTA, pH 7.2 (buffer B), and re-equilibrated. Carbohydrate-positive fractions (2 ml) were combined and dialysed against repeated changes of 10 mM Tris/HCl, pH 7.7 (buffer C). In step 2 the glycoprotein pool from step 1 was applied in 10 ml aliquots to a DEAE HiPrep column (Amersham Biosciences) equilibrated with buffer C. After washing the column with 100 ml of buffer C, the S-layer glycoprotein and residual proteins were eluted with a 200 ml gradient from 0 to 1 M KCl in buffer C at a flow rate of 5 ml/min, and 4 ml fractions were collected. Individual fractions were analysed by SDS/PAGE (10% gels). In step 3, those fractions containing S-layer glycoprotein were pooled and precipitated with $(\text{NH}_4)_2\text{SO}_4$ as described in step 1, and interfacial salting out and dialysis were carried out as described above. For final purification (step 4), the S-layer glycoprotein from step 3 was applied to a MonoQ HR 5/5

column (Amersham Biosciences) equilibrated with 20 mM Tris/HCl, pH 7.7 (buffer D). The column was washed at a flow rate of 1 ml/min with 10 ml of buffer D and the (glyco)protein was eluted by a 60 ml gradient from 0 to 3 M KCl in buffer D. Fractions (2 ml) containing the S-layer glycoprotein according to the SDS/PAGE profile were combined, sterile filtered (0.14 μ m) and stored in 500 μ l aliquots at -20 °C until used for microcrystallization experiments. For the determination of amino acid and carbohydrate composition and SDS/PAGE analysis, the material was thoroughly dialysed until salt-free and lyophilized.

Microcrystallization

Microcrystallization experiments were carried out with the MonoQ-purified S-layer glycoprotein of *A. thermoaerophilus* DSM 10155 at a final protein concentration of 11 mg/ml of 0.3 M KCl in buffer D. For initial screening, the crystallization salts and buffers provided with the Crystal Screen 2™ Kit (Hampton Research, Laguna Niguel, CA, U.S.A.) were used in the sitting-drop method. Solvent reservoirs contained 100 μ l of crystallization solution and the drop was mixed from 1 μ l of S-layer glycoprotein and 1 μ l of crystallization solution. Crystal Screen 2™ was performed in duplicates for incubation at room temperature and at 4 °C. Drops were examined daily under stereo microscope. Aliquots of 5 μ l of those solutions showing precipitation were investigated in detail by electron microscopy after 14 days of incubation at 4 °C. Accordingly, crystallization conditions were optimized by varying the concentration and the pH value of the crystallization solutions. For detailed investigations, 30% poly(ethylene glycol) 20000 (PEG 20000; Hampton Research) and 100% poly(ethylene glycol) 550 monomethylether (MME 550; Hampton Research) were used. The pH was varied between 4 and 8, and the final concentrations of the crystallization polymers were 5–15% for PEG 20000 and 15–35% for MME 550.

Assembly of S-layer glycoprotein

To investigate the putative role of the SCWP in the assembly process of the S-layer glycoprotein of *A. thermoaerophilus* DSM 10155, the microcrystallization approach giving the most promising result was scaled-up. For this purpose, to 1 ml of S-layer glycoprotein (corresponding to 11 mg of protein) the same volume of either 20% PEG 20000 in 0.2 M Hepes, pH 7.0 (PEG 20000 solution), or 40% MME 550 in 0.2 M Hepes, pH 7.0 (MME 550 solution), was added. Reactions were carried out for 14 days at 4 °C. Samples were centrifuged at 65 000 g for 20 min at 4 °C, followed by repeated washes of the pellet with either PEG 20000 solution or MME 550 solution, both diluted 1:2 with water. Supernatants were collected and concentrated to a final volume of ≈ 1 ml on a rotary evaporator. The crystallization polymer and buffer substances were removed by gel-permeation chromatography, using columns of Bio-Gel P-30 and P-60 (Bio-Rad) in the case of the PEG 20000 approach, and a Sephadex G-10 column (Amersham Biosciences) in the other case. Elution was done with 0.1 mM NaCl, column dimensions were 1.0 cm $\times$ 120 cm, and fractions of 1 ml were collected and pooled according to the refraction index and the carbohydrate content. The pelleted material from either crystallization approach was resuspended into 7 ml of the respective wash buffer and thoroughly dialysed against water. Salt-free supernatants and pellets were lyophilized and analysed by SDS/PAGE, and for amino acid and carbohydrate composition.

Table 1 ^1H and ^{13}C chemical shift data (δ , p.p.m.) of the SCWP of *A. thermoaerophilus* DSM 10155 ($^2\text{H}_2\text{O}$, 313 K, pH 5.1) and specific nuclear Overhauser effects (NOEs) and rotating-frame Overhauser effects (ROEs) as well as ($^{13}\text{C} \rightarrow ^1\text{H}$)-HMBC correlations (HMBC corr.) relevant for identification of glycosidic linkages

Glycosidic linkages are indicated by residue designation and position number.

Residue	H/C atom							
	1	2	3	4	5	6a/b	CO	CH_3
A	5.41	3.97	3.78	3.58	3.93			
α -GlcNAc	101.50	50.30	74.90	69.76				
NOE, ROE to ^1H			B1					
HMBC corr. to ^1H			B1					
B	5.23	3.98	3.79	3.69	3.82	3.71/3.85		2.00
α -GlcNAc	99.69	53.77	82.05	74.14	72.60	61.40	175.11	23.13
NOE, ROE to ^1H	A3		C1, C2	C1, C2	C1			
C	4.86	4.33	3.98	3.82	3.84	3.78/3.83		2.00
β -ManNAc	100.88	55.18	74.40	74.25	74.72	61.92	176.54	23.45
NOE, ROE to ^1H	B3, B4, B5	B3, B4	D1	D1				
HMBC corr. to ^1H	B3		D1					
D	5.23	3.98	3.79	3.69	3.82	3.71/3.85		2.00
α -GlcNAc	99.69	53.77	82.05	74.14	72.60	61.40	175.11	23.13
NOE, ROE to ^1H	C3, C4		E1, E2	E1, E2	E1			
HMBC corr. to ^1H			E1					
E	4.86	4.33	3.98	3.99	3.84	3.88/3.90		2.00
β -ManNAc	100.88	55.18	74.73	74.50	74.72	62.07	176.54	23.45
NOE, ROE to ^1H	D3, D4, D5	D3, D4	F1	F1	F1			
HMBC corr. to ^1H			F1	F1				
F	5.27	4.21	3.90	4.18	3.87	3.73/3.91		2.01
α -GlcNAc	98.67	49.65	77.82	69.73	71.78	62.68	175.42	23.50
NOE, ROE to ^1H	E3		G1	G1, G2	G1			
G	4.57	3.69	3.45	3.68	3.81	3.53/3.60		2.02
β -GalNAc	103.69	56.78	76.09	80.22	77.86	60.80	175.77	23.70
NOE, ROE to ^1H	F3, F4, F5	F4, H1	H1, H2	H1				
HMBC corr. to ^1H	F3							
H	4.87	4.49	4.07	3.69	3.86	3.67/3.81		2.04
β -ManNAc	100.75	55.14	74.79	74.86	74.65	62.00	176.93	23.49
NOE, ROE to ^1H	G2, G3, G4	G3	I5, I6a			I1		
HMBC corr. to ^1H	G4							
I	5.27	4.21	3.90	4.18	3.87	3.73/3.91		2.01
α -GlcNAc	98.67	49.65	77.82	69.73	71.78	62.68	175.42	23.50
NOE, ROE to ^1H	H6b		K1	K1, K2	H3, K1	H3		
HMBC corr. to ^1H			K1					
K	4.57	3.69	3.45	3.68	3.81	3.53/3.60		2.02
β -GalNAc	103.69	56.78	76.09	80.22	77.86	60.80	175.77	23.70
NOE, ROE to ^1H	I3, I4, I5	I4, L1	L1, L2	L1				
HMBC corr. to ^1H				L1				
L	4.87	4.49	4.07	3.69	3.86	3.67/3.81		2.04
β -ManNAc	100.75	55.14	74.79	74.86	74.65	62.00	176.93	23.49
NOE, ROE to ^1H	K2, K3, K4	K3	M1	M1				
HMBC corr. to ^1H				M1				
M	5.23	3.97	3.59	3.69	3.82	3.68/3.81		2.00
α -GlcNAc	99.69	53.77	69.83	74.38	72.60	61.41	175.11	23.13
NOE, ROE to ^1H	L3, L4							
F'	5.28	4.33	3.93	4.25	3.90	3.75/3.94		2.01
α -GlcNAc	98.67	50.24	76.65	75.18	72.86	62.68	175.55	23.50
NOE, ROE to ^1H	E4, E5		G'1	G'1, G'2	G'1			
G'	4.51	3.63	3.41	3.68	3.86	3.54/3.62		2.02
β -GalNAc	104.33	56.69	75.69	80.80	76.66	60.80	175.77	23.70
NOE, ROE to ^1H	F'3, F'4, F'5	F'4, H'1	H'1, H'2	H'1				
HMBC corr. to ^1H	F'3							
H'	4.87	4.49	4.05	3.67	3.86	3.67/3.81		2.04
β -ManNAc	100.75	55.14	74.17	74.86	74.65	62.00	176.93	23.49
NOE, ROE to ^1H	G'2, G'3, G'4	G'3	I'5, I'6a			I'1		
HMBC corr. to ^1H	G'4							
I'	5.28	4.33	3.93	4.25	3.90	3.75/3.94		2.01
α -GlcNAc	98.67	50.24	76.65	75.18	72.86	62.68	175.55	23.50
NOE, ROE to ^1H	H'6b		K'1	K'1, K'2	H'3, K'1	H'3		
HMBC corr. to ^1H			K'1					

Table 1 (contd.)

Residue	H/C atom							
	1	2	3	4	5	6a/b	CO	CH ₃
K'	4.51	3.63	3.41	3.68	3.86	3.54/3.62		2.02
β -GalNAc	104.33	56.69	75.69	80.80	76.66	60.80	175.77	23.70
NOE, ROE to ¹ H	I'3, I'4, I'5	I'4, L'1	L'1, L'2	L'1				
HMBC corr. to ¹ H				L'1				
L'	4.87	4.49	4.05	3.67	3.86	3.67/3.81		2.04
β -ManNAc	100.75	55.14	74.17	74.86	74.65	62.00	176.93	23.49
NOE, ROE to ¹ H	K'2, K'3, K'4	K'3, M'1	M'1	M'1				
HMBC corr. to ¹ H				M'1				
M'	5.23	3.97	3.59	3.69	3.82	3.68/3.81		2.00
α -GlcNAc	99.69	53.77	69.83	74.38	72.60	61.41	175.11	23.13
NOE, ROE to ¹ H	L'2, L'3, L'4							
P*	4.97	3.87	3.83	3.89	3.52	3.96/4.28		2.02
α -MurNAc	98.28	55.51	77.65	71.80	71.35	72.90	176.13	23.35
P†	4.53	3.80	3.70					
β -MurNAc	101.86	53.85	77.88					

* Additional ¹H/¹³C shifts for 3-*O*-lactyl: 1.38/19.71; 4.52/78.99; 177.02.

† Additional ¹H/¹³C shifts for 3-*O*-lactyl: 1.37/18.00; 4.40/79.60; 176.33.

RESULTS

Preparation of the SCWP–peptidoglycan complex

From ≈ 100 g (wet weight) of *A. thermoaerophilus* DSM 10155 biomass, 410 mg of crude peptidoglycan preparation was obtained. The peptidoglycan was confirmed to be of the A1 γ -type, as was indicated by a molar ratio of *N*-acetylmuramic acid (MurNAc)/Glu/Ala/diaminopimelic acid (*m*-A₂pm) (1.0:1.0:1.6:1.0; *m*-A₂pm was arbitrarily set to the value of 1), and the degree of cross-linking of the peptidoglycan was 53%, based on the distinction between cross-linked and underivatized *m*-A₂pm.

The SCWP–peptidoglycan complex was purified from the crude peptidoglycan preparation following a protocol preserving the linkage between the SCWP and the peptidoglycan. After separation of the SCWP–peptidoglycan complex from the majority of glycan-free, low-molecular-mass peptidoglycan breakdown products by Sephadex G-50 chromatography, 47% of that material was recovered within the fractionation range of a Bio-Gel P-30 column. Final purification was accomplished by semi-preparative reversed-phase (C₁₈) HPLC. The material of interest was eluted as a single but broad peak between 32 and 40 min, corresponding to an acetonitrile concentration of 10.6–13.3%; residual, peptidoglycan-derived peptides were eluted at 100% acetonitrile. The total amount of purified SCWP–peptidoglycan complex was ≈ 50 mg, which corresponds to an overall content of SCWP in the peptidoglycan of $\approx 13\%$ (w/w). Carbohydrate analysis revealed the constituents GalNAc, ManNAc and GlcNAc in the molar ratio 1:1.4:2.2 (GalNAc was arbitrarily set to the value of 1). Combining carbohydrate and amino acid data, a MurNAc/GlcNAc molar ratio of 1:3.5 was found. This finding indicated that the excess of GlcNAc was due to the associated SCWP and that the preparation indeed contained an intact SCWP–peptidoglycan complex, with the SCWP, on average, constituting two-thirds (by weight) of this complex. The molar ratio of phosphate/MurNAc was determined to be 8:1. Assuming a pyrophosphate linkage between the SCWP and the peptidoglycan moieties, this was indicative of an average substitution of every fourth MurNAc residue with a SCWP. This material was subjected to structure analysis by NMR spectroscopy.

NMR analysis of the SCWP–peptidoglycan complex

NMR structure analysis was based on identification of the ¹H and ¹³C resonance frequencies of the sugar units and on the classification of the monosaccharide type, based on well-known chemical shift ranges [9,22,23] and intra-unit through-the-bond correlations from (H,H)-double quantum-filtered correlated spectroscopy (DQF-COSY), (H,H)-TOCSY and (H,C)-HSQC spectra, acquired at natural isotope abundance. In the following interpretation of the NMR data, capital bold letters denote individual monosaccharide units, as defined in Table 1, following their sequence in the proposed oligosaccharide structure.

In the anomeric region of a coupled (H,C)-HSQC spectrum (results not shown), different monosaccharide groups were distinguished by well-resolved signals of most monomeric units. The ¹³C signals at about 98.7 p.p.m. (**F**, **F'**, **I**, **I'**), 99.7 p.p.m. (**B**, **D**, **M**, **M'**) and 101.50 p.p.m. (**A**) were assigned to C1 resonances with a ¹J_{CH}-coupling constant of 172 Hz, indicating α -configuration, whereas the resonances around 100.8 p.p.m. (**C**, **E**, **H**, **H'**, **L**, **L'**), 103.7 p.p.m. (**G**, **K**) and 104.3 p.p.m. (**G'**, **K'**) were in the β -configuration, according to ¹J_{CH}-coupling constants of about 162 Hz. The signals at ≈ 98.2 p.p.m. and ≈ 101.7 p.p.m. belonged to the two anomeric forms of reducing MurNAc, in an α (**P**)/ β (**P'**) ratio of about 7:3.

With the notable exception of β -D-GalpNAc, for which the ¹H chemical shifts also deviated significantly from the published values [23], it was possible to identify the spin systems by straightforward interpretation of the chemical shifts and connectivities found in (H,H)-DQF-COSY (Figure 1), (H,H)-TOCSY and coupled (H,C)-HMBC (Figure 2). The assignments are summarized in Table 1. Further discrimination between the sugar configurations was possible from H2 and H3 coupling constant magnitudes. We found glucose-type sugars (**A**, **B**, **D**, **F**, **F'**, **I**, **I'**, **M**, **M'**), with a ³J_{H2H3} of about 10 Hz, and mannose-type sugars (**C**, **E**, **H**, **H'**, **L**, **L'**), with a ³J_{H2H3} of about 5 Hz, as estimated from DQF-COSY. In the spin systems assigned to β -D-GalpNAc (**G**, **G'**, **K**, **K'**), the chemical shifts and coupling constants did not allow straightforward identification of the sugar conformation type. In the usually prevailing conformation (⁴C₁), nuclear Overhauser effect (NOE) correlations would

Figure 1 Part of a $(^1\text{H}, ^1\text{H})$ -COSY spectrum ($f_1 \times f_2 = 1024 \times 4096$, number of scans (NS) = 8) of the SCWP

Lines highlight the spin system for GlcNAc (F, F') and ManNAc (H, H', L, L'). The expansion shows the coupling pattern $^3J_{23}$ and $^3J_{34}$ of GalNAc (G, G') coinciding with K, K').

occur between the *N*-acetyl-methyl resonances and the axial protons H1, H2 and H3 of β -D-GalpNAc. Instead, there were only NOE correlations to H4 and H5 (Figure 3). This behaviour was explained by a ring flip of the 1,4-glycosylated sugar into the $^1C_4$ conformation. Vicinal ^1H - ^1H coupling magnitudes (see detail in Figure 1) further corroborated this conformation; both $J_{\text{H}2\text{H}3}$ and $J_{\text{H}3\text{H}4}$ were < 5 Hz.

For the C2 resonances all monomeric units exhibited chemical shifts (≈ 50 p.p.m.) typical of acetylated *N*-glycosidic carbons. Several of the ^1H singlets from acetyl groups (at ≈ 2 p.p.m.) were found to correlate to carbonyl-C resonances at ≈ 175 p.p.m. in (H,C)-HMBC spectra. Since well-defined NOE cross-peaks were found from these acetyl methyl groups (≈ 2 p.p.m.) to protons of the attached monosaccharide units (e.g. H1 and H2, see Figure 3), we concluded, in accordance with the chemical analysis, that all monomers were 2-acetamido-2-deoxy-pyranoses. In the cases of the β -D-ManpNAc units (C, E, H, H', L, L'), the HMBC

spectra also showed correlation peaks from the carbonyl carbon atoms to H2 of β -D-ManpNAc. Combining this information with the integral ratio of a one-dimensional ^1H spectrum, we concluded that the SCWP from *A. thermoaerophilus* DSM 10155 consisted of α -D-GlcpNAc, β -D-ManpNAc and β -D-GalpNAc in a ratio of 9:6:4. This is in accordance with the high-performance anion-exchange chromatography with pulsed electrochemical detection ('HPAEC/PED') data.

It was also possible to identify the resonances of a fragment of the peptidoglycan portion, which contained four amino acids (2 Ala, Glu, *m*-A₂pm) and MurNAc (P, P') at the reducing end, indicating a cleavage of the glycosidic linkages between the β -D-GlcpNAc units of the peptidoglycan backbone and MurNAc. Due to severe overlapping of individual signals at this field strength, it was not possible to assign unambiguously all the signals arising from the MurNAc residue, especially in the case of the β -anomer. Nevertheless, the identification of monomeric

Figure 2 Part of an (H,C)-HMBC spectrum ($f_1 \times f_2 = 4096 \times 4096$, NS = 32)

Lines mark the chemical shifts of named ^{13}C atoms. Named signals identify the correlated protons. Of special interest are those corresponding to C3 and C4 atoms of the branching ManNAc (E).

Figure 3 Detail of an (H,H)-NOESY spectrum ($f_1 \times f_2 = 512 \times 4096$, NS = 16)

Figure shows NOE cross-peaks from *N*-acetyl-methyl groups. For GalNAc (G, G', K, K') there are only cross-peaks to H4 and H5.

building blocks laid the foundation for the analysis of the inter-monomer connectivities and the ^{13}C chemical shifts provided evidence for the positions of glycosidic substitutions.

Through-the-space ^1H - ^1H NOEs or rotating-frame Overhauser effects (ROEs) and through-the-bonds long-range ^1H - ^{13}C J -coupling across the glycosidic linkages in HSQC-type HMBC spectra were used to derive the connectivities between individual monosaccharide units. Glycosidic linkages were identified from

NOESY and ROESY correlation peaks of the anomeric protons (see Figure 4, left-hand panel and Table 1). The almost complete assignment of the ^1H -NMR signals of the oligosaccharide allowed identification of inter-monomer correlations also from other protons of the different hexose units as shown in Figure 4 (right-hand panel).

Some ambiguities from accidental degeneracy of ^1H chemical shifts could be resolved by the use of complementary information

Figure 4 Details of an (H,H)-ROESY spectrum ($f_1 \times f_2 = 512 \times 4096$, NS = 16)

ROE cross-peaks for branching are shown in the left-hand panel. Identification of the linkage GlcNAc (I, I') (α 1-3) ManNAc (H, H') is shown in both panels by very distinct ROE cross-peaks.

from long-range H-C couplings ($H \rightarrow C \rightarrow O \rightarrow C$ or $C \rightarrow O \rightarrow C \rightarrow H$) to well-resolved ^{13}C chemical shifts of glycosidic carbons (74–82 p.p.m.). Almost all glycosidic linkages could be identified in HMBC spectra (optimized for 10 Hz coupling), as summarized in Table 1. For example, C4 of β -D-GalpNAc (G', K') was found at a chemical shift of 80.80 p.p.m. In Figure 2, one can see the intra-monomer scalar coupling to H5 and the inter-monomer coupling to the anomeric proton of β -D-ManpNAc (H', L') at 4.87 p.p.m.

The branching of the SCWP glycan at β -D-ManpNAc (E) is of unique structural importance. Due to the difficulties arising from near degenerate chemical shifts of H3 (3.98 p.p.m.) and H4 (3.99 p.p.m.) of this sugar unit, the branching was confirmed by two approaches: (i) there were ROE cross-peaks from H1 of α -D-GlcpNAc (F') to H4 of E and from H1 of α -D-GlcpNAc (F) to H3 (E) (Figure 4, left-hand panel), and (ii) as shown in Figure 2, we saw $^3J_{\text{CH}}$ couplings from C3 of β -D-ManpNAc (E) to H1 of α -D-GlcpNAc (F') and from C4 of β -D-ManpNAc (E) to H1 of α -D-GlcpNAc (F) in the HMBC spectra.

The connections between the repeating trisaccharide units in each branch were also of particular interest in this context. For these few glycosidic linkages no correlation signals were found in the HMBC spectra. However, the very specific and distinct inter-residue ROE cross-peaks, in particular (I, I')-H1/(H, H')-H6b (Figure 4, left-hand panel) and (H, H')-H3/(I, I')-H5,6 (Figure 4, right-hand panel), unequivocally identified these links.

Regarding the pyrophosphate bridge which is linking the SCWP to C6 of MurNAc (P, P'), the ^{31}P -NMR spectra of this

SCWP-peptidoglycan complex were very similar to those described for the SCWP-peptidoglycan complex of *P. alvei* CCM 2051 [9], which corroborates the pyrophosphate structure.

The complete structure of the SCWP from *A. thermoaerophilus* DSM 10155 as derived from the NMR data and chemical analysis is shown in Figure 5.

ESI-TOF MS

Based on preliminary matrix-assisted laser desorption-ionization time-of-flight ('MALDI-TOF') measurements [13] and chemical analyses (this study), the expected m/z range of the SCWP-peptidoglycan complex was ≈ 4000 . Positive-ion mode ESI-TOF mass spectra of the sample under acidic conditions provided no detectable signals in the corresponding m/z ranges for singly or doubly charged cations. According to the amino acid composition, no positive charge was expected for the intact complex, but a single positive charge would originate from the SCWP-free peptidoglycan fragment. Despite a low signal-to-noise ratio, in strongly alkaline solution, only a single signal was found between masses 3000 and 6000, the range expected for triply charged ions (results not shown). Its averaged m/z of 1347.1 ± 2.6 was in good agreement with 19 N-acetylated hexose units with an attached pyrophosphate group ($\text{C}_{152}\text{H}_{248}\text{N}_{19}\text{O}_{102}\text{P}_2^{3-}$, m/z 1345.22). Since fragmentation during ionization could be excluded due to the very mild ionization method used [24], the heterogeneous sample contained a significant amount of SCWP attached to pyrophosphate, from which the peptidoglycan part had been cleaved.

Figure 5 Schematic drawing of the biantennary SCWP-peptidoglycan complex isolated from *A. thermoaerophilus* DSM 10155

Bold letters designate sugar monomer units as referred to in the NMR data. Dotted lines indicate the peptidoglycan portion, which was not subject of the present study. GN, D-GlcNAc; MN, D-ManNAc; GaN, D-GalNAc; MurN, D-MurNAc.

Assembly of the S-layer glycoprotein

Even though during fermenter cultivation of *A. thermoaerophilus* DSM 10155 only low amounts of S-layer glycoprotein (80–100 mg/l) could be recovered from culture supernatants [13], the glycoprotein was extractable as a water-soluble protein after a freeze-thaw cycle of the bacterial cells [11,11a]. The glycoprotein, which was purified to apparent homogeneity by several chromatographic steps, represented 4% (w/w) of the biomass used for its extraction. A silver-stained SDS/polyacrylamide gel (7.5%) showed a single, sharp carbohydrate-staining protein band with an apparent molecular mass of 153 kDa (results not shown), in accordance with previously published data [11,11a]. Detailed amino acid and carbohydrate analyses revealed that 13% of glycosylation (by weight) of that material originated from the S-layer O-glycans, which are composed of, on average, 15 [→4)- α -L-Rhap-(1→3)- β -D-glycero-D-manno-Hepp-(1→] disaccharide repeating units [11,11a,12] and an [α -L-Rhap-(1→3)]_{n-0-2}- β -D-GalpNAc-(1→ core saccharide; an associated biantennary SCWP (this study) accounted for an additional 6% of glycosylation.

Due to its complete water solubility, this S-layer glycoprotein was suitable for microcrystallization experiments. Precipitates were obtained especially well when using crystallization solutions containing poly(ethylene glycols). Other alcohols and different uni- or bivalent ions from the Crystal Screen 2™ Kit had no effect on the crystallization. Using MME 550 at a final concentration of 20% at pH 7.0 as crystallization agent, the formation of flat, mono-layered self-assembly products with square lattice symmetry was induced (Figure 6A). Computer-enhanced image analysis of negatively stained preparations of that material showed that the S-layer glycoprotein lattice was identical to that observed on intact *A. thermoaerophilus* DSM 10155 cells [25]. Interestingly, with PEG 20000 at a final concentration of 10% using the same buffer system, again no three-dimensional crystals, just stacks of bi-layered self-assembly products, were obtained. This was clearly demonstrated by thin-sectioned preparations of assembled S-layer glycoprotein (Figure 6B). These products are the first self-assembly products ever obtained with the S-layer glycoprotein from *A. thermoaerophilus* DSM 10155.

To unravel the putative role of the SCWP in the self-assembly process of the S-layer glycoprotein, pelleted self-assembly products and corresponding supernatants derived from the scaled-up crystallization experiments were investigated. It became evident that the content of SCWP in the pellet had dropped dramatically in comparison with the material initially used for crystallization. Extensive washing with buffer sequentially removed the SCWP. The SCWP recovered from the corresponding

supernatant of the MME 550 approach was 85% after three washing steps and $\approx$ 98% after seven washing steps, respectively. In the case of the PEG 20000 approach, the self-assemblies retained 30% of SCWP after three washes. Thus assembled S-layer glycoprotein from *A. thermoaerophilus* DSM 10155 was essentially free of SCWP.

DISCUSSION

In this study, we report on the detailed analysis of the SCWP from the thermophilic Gram-positive organism *A. thermoaerophilus* DSM 10155. As deduced from NMR spectroscopy data, the three most salient features of that SCWP are (i) the biantennary oligosaccharide structure, (ii) the structural homogeneity of the oligosaccharide moiety and (iii) the pyrophosphate linkage to the peptidoglycan sacculus. In contrast, previously characterized SCWPs of S-layer-carrying bacilli had a variable number of constituent repeating units [7,9], and biantennary SCWPs have, thus far, not been reported for a member of the domain Bacteria. It is presumed that, due to its unique structure, the SCWP of *A. thermoaerophilus* DSM 10155 serves several important *in vivo* and *in vitro* biological functions.

In general, SCWPs are, together with S-layer glycoproteins and the peptidoglycan, the major glycosylated compounds of the Gram-positive cell wall [1]. SCWPs include mainly anionic structures [26] but, less frequently, also neutral structures, as has been demonstrated in this and other recent studies (Schäffer, C., Kählig, H., Zayni, S., Scheberl, A. and Messner, P., unpublished work). They are either covalently linked to the peptidoglycan network (teichoic and teichuronic acids) or tethered to a lipid anchor (lipoteichoic acids or lipoglycans) [1]. Interestingly, all SCWPs of S-layer-carrying organisms completely analysed so far are anchored to the peptidoglycan via a pyrophosphate bridge [7,9], which makes them distinctly different from teichoic and teichuronic acid-like structures. The pyrophosphate bond might represent a pre-determined cleavage site with possible relevance in the S-layer glycoprotein biosynthesis [27], because pyrophosphates are commonly found in biosynthetic intermediates of different bacterial cell wall polymers. Even though the exact biological function of SCWPs is not fully understood yet, several general functions have been attributed to them. These include the binding of bivalent cations, which keeps the peptidoglycan sacculus in an expanded state by charge repulsion, the binding of proteins to create an acidic cell wall, or the formation of a barrier to prevent diffusion of nutrients and metabolites [1]. However, these functions mainly apply for charged polymers.

As a newly defined function, SCWPs serve as anchors for S-layer (glyco)proteins to the bacterial cell wall. Surface display,

Figure 6 Negatively stained preparation of self-assembled, sheet-like S-layer glycoprotein obtained after treatment with MME 550 (A) and thin-sectioned preparation of stacks of S-layer glycoprotein obtained after treatment with PEG 20000 (B)

Scale bar, 100 nm.

based on a presumably lectin-like interaction with the N-terminus of the respective S-layer (glyco)protein, utilizes at least two different mechanisms. On the N-terminus of the S-layer protein SbsB of *G. stearothermophilus* PV72/p2, the co-existence of a SCWP-binding domain and a peptidoglycan-binding domain, which contains a typical SLH motif, was identified [28,29]. N-terminal SLH motifs interacting with SCWPs were demonstrated for the S-layer proteins SlpA of *Clostridium thermocellum* [30,31], the S-layer glycoprotein of *Thermoanaerobacterium thermosulphurigenes* EM1 [32], the S-layer protein SbpA of *Bacillus sphaericus* CCM 2177 [8], as well as EA1 and Sap of *Bacillus anthracis* [5]. For the last two organisms, pyruvylated monosaccharides are the ligands of the SLH domain [5]. However, for the well-investigated S-layer proteins of *G. stearothermophilus* wild-type strains, alternative cell-wall-anchoring mechanisms must have evolved, because these S-layer proteins are devoid of SLH motifs [21,33]. Interestingly, the *G. stearothermophilus* S-layer proteins SbsA [34], SbsC [35] and SgsE [21] possess highly conserved N-terminal regions, rich in arginine, tyrosine and lysine residues, which recognize a distinct 2,3-diacetamido-2,3-dideoxymannuronic acid-containing SCWP [7].

Initial sequencing data from the S-layer protein SatA of *A. thermoaerophilus* DSM 10155 (Pföstl, A., Novotny, R., Schäffer, C. and Messner, P., unpublished work), showing the absence of an SLH domain, in combination with the knowledge of the composition of the SCWP (this study), indicate that neither SLH motifs nor polymer-inherent charge phenomena are likely to be involved in the cell-surface display strategy of the S-layer glycoprotein from this organism. Apparently, the highly regular and symmetric arrangement of the biantennary SCWP enables very efficient binding to the S-layer polypeptide, thereby linking the S-layer glycoprotein to the peptidoglycan. The concept that a specific binding takes place is supported by the fact that residual amounts of SCWP were always co-purified with the S-layer glycoprotein. Furthermore, due to its defined length, the SCWP might keep the S-layer protomers at a certain distance

from the peptidoglycan layer and promote correct folding and surface assembly of the newly synthesized S-layer subunits. A comparable effect is attributed to chaperones in the eukaryotic calnexin system [36]. Interestingly, in *Aeromonas salmonicida*, a similar interplay between the O-polysaccharide and the A-layer, as presumed for *A. thermoaerophilus* DSM 10155, is necessary for correct surface protein assembly. In this case, the O-polysaccharide is associated with and extended through the A-layer, thereby anchoring the A-layer to the cell surface. However, the mode of association between the O-polysaccharide and the protein has not been characterized in detail [37].

Initial crystallization attempts on the S-layer glycoprotein of *A. thermoaerophilus* DSM 10155 indicate that the SCWP does not only play an important role for cell wall anchoring *in vivo*, but also mediates the water-solubility of the S-layer *in vitro*. Whereas during fermenter cultivation of the bacteria only a low amount of S-layer glycoprotein, presumably due to lysed cells, was found in culture supernatants, it was extractable as a water-soluble glycoprotein after a freeze-thaw cycle of the cells. Comparable solubility of S-layer proteins has been described for *Campylobacter fetus* [38], *Azotobacter vinelandii* [39] and *Brevibacillus brevis* 47 [40]. However, in these cases the water solubility relied on removal of bivalent cations, which are required for S-layer protein assembly. Consequently, whereas *in vivo* the S-layer protomers of *A. thermoaerophilus* DSM 10155 are assembled into a square S-layer lattice on the bacterial cell surface, isolated subunits no longer displayed the ability to self-assemble into ordered arrays (this study). Interestingly, after addition of modified or non-modified poly(ethylene glycols) as crystallizing agents to an S-layer glycoprotein preparation, the self-assembly property could be restored *in vitro*. The self-assembly process was preceded by dissociation of the SCWP from the S-layer glycoprotein. This suggests that, on the intact bacterial cell, the SCWP holds the S-layer protomers in close proximity, which enables their inter- and intra-molecular interactions and thus S-layer glycoprotein lattice formation to occur. The high density

of SCWP in the peptidoglycan layer ($\approx 25\%$ of the MurNAc residues are substituted with the polymer) might additionally ensure the correct assembly of the S-layer protomers on the cell surface during all stages of cell growth and division. When the supporting peptidoglycan layer is removed during S-layer glycoprotein preparation, the covalent bonds between the SCWP and the peptidoglycan backbone are cleaved, releasing the S-layer glycoprotein with attached SCWP. The high degree of overall glycosylation (19%) of purified S-layer glycoprotein is likely to be responsible for its water solubility, when detached from the cellular support. Besides the increased hydrophilicity of the S-layer glycoprotein, the 20–30 nm-long S-layer glycan chains might provide an umbrella-like coating of the protomers, resulting in the steric hindrance of assembly [41].

It is conceivable that crystallization polymers substitute the SCWP at its lectin-like binding sites on the S-layer protein, thereby functioning as macromolecular crowding agents [42]. The different S-layer self-assembly products obtained after the addition of either MME 550 or PEG 20000 might reflect different reaction mechanisms being responsible for the dissociation of the SCWP–peptidoglycan complex and, further on, for the *in vitro* re-assembly of the S-layer protomers. It is conceivable that with the small molecule MME 550, which, at least to some extent, resembles the shape and properties of a methyl glycoside, the specific binding between the non-reducing-end sugars of the SCWP and the binding pocket(s) in the S-layer polypeptide could be perturbed. With PEG 20000 obviously a more profound molecular crowding mechanism [42] is responsible for the fast formation of stacks of double-layered S-layer assembly products by dehydration acting on the S-layer polypeptide (Figure 6B). The resulting situation might generally be seen as mimicry of the excluded volume effects thought to prevail on the bacterial cell surface *in vivo*. Thus only under conditions of molecular crowding can the glycosylated S-layer protomers of *A. thermoaerophilus* DSM 10155 interact with each other through hydrophobic interactions. This interpretation supports the general assumption of a lectin-type like binding between S-layer (glyco)proteins and SCWP, because lectin-type binding involves mostly hydrophobic forces and hydrogen bonds, and only rarely electrostatic forces, since many carbohydrates are uncharged [43]. Due to the generally quite low binding affinities of lectins for carbohydrate ligands [44], in the case of *A. thermoaerophilus* DSM 10155, the biantennary structure of the SCWP, with two antennae of equal length, may result in a doubling of the number of the binding sites of the SCWP for the S-layer glycoprotein and so could trigger a form of glycoside cluster effect, allowing a biologically relevant binding strength to be attained.

The mechanism underlying the binding process between the S-layer (glyco)proteins and the SCWP requires further investigations by techniques such as surface plasmon resonance spectroscopy and NMR spectroscopy to unravel fully the significance of this interesting strategy of bacterial cell surface display.

We thank Rudolf Christian for help with the interpretation of the NMR data and Andrea Scheberl and Sonja Zayni for excellent technical assistance. Critical reading of the manuscript by Dr Iain Wilson and Dr Friedrich Sreenc is gratefully acknowledged. This work was supported by the Austrian Science Fund, projects P14209-MOB (to P.M.), P15380 (to N.M.) and P13906 (to W. Buchberger and N.M.).

REFERENCES

- Navarre, W. W. and Schneewind, O. (1999) Surface proteins of Gram-positive bacteria and mechanisms of their targeting to the cell envelope. *Microbiol. Mol. Biol. Rev.* **63**, 174–229
- Sleytr, U. B., Messner, P., Pum, D. and Sára, M. (1999) Crystalline bacterial cell surface layers (S-layers): from supramolecular cell structure to biomimetics and nanotechnology. *Angew. Chem. Int. Ed.* **38**, 1034–1054
- Sára, M. and Sleytr, U. B. (2000) S-layer proteins. *J. Bacteriol.* **182**, 859–868
- Sleytr, U. B., Sára, M., Pum, D., Schuster, B., Messner, P. and Schäffer, C. Self-assembly protein systems: microbial S-layers. In *Biopolymers*, vol. 7 (Steinbüchel, A., ed.), Wiley-VCH, Weinheim, in the press
- Mesnage, S., Fontaine, T., Mignot, T., Delepiere, M., Mock, M. and Fouet, A. (2000) Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *EMBO J.* **19**, 4473–4484
- Sára, M. (2001) Conserved anchoring mechanisms between crystalline cell surface S-layer proteins and secondary cell wall polymers in Gram-positive bacteria? *Trends Microbiol.* **9**, 47–50
- Schäffer, C., Kählig, H., Christian, R., Schulz, G., Zayni, S. and Messner, P. (1999) The diacetamidodideoxyuronic-acid-containing glycan chain of *Bacillus stearothermophilus* NRS 2004/3a represents the secondary cell wall polymer of wild-type *B. stearothermophilus* strains. *Microbiology* **145**, 1575–1583
- Ilk, N., Kosma, P., Egelseer, E. M., Mayer, H. F., Sleytr, U. B. and Sára, M. (1999) Structural and functional analyses of the secondary cell wall polymer of *Bacillus sphaericus* CCM 2177 that serves as an S-layer-specific anchor. *J. Bacteriol.* **181**, 7643–7646
- Schäffer, C., Müller, N., Mandal, P. K., Christian, R., Zayni, S. and Messner, P. (2000) A pyrophosphate bridge links the pyruvate containing secondary cell wall polymer of *Paenibacillus alvei* CCM 2051 to muramic acid. *Glycoconj. J.* **17**, 681–690
- Araki, Y. and Ito, E. (1989) Linkage units in the cell wall of Gram-positive bacteria. *CRC Crit. Rev. Microbiol.* **17**, 121–135
- Kosma, P., Wügeditsch, T., Christian, R., Zayni, S. and Messner, P. (1995) Glycan structure of a heptose-containing S-layer glycoprotein of *Bacillus thermoaerophilus*. *Glycobiology* **5**, 791–796
- Erratum (1996) *Glycobiology* **6**, 5
- Wügeditsch, T., Zachara, N. E., Puchberger, M., Kosma, P., Gooley, A. A. and Messner, P. (1999) Structural heterogeneity in the core oligosaccharide of the S-layer glycoprotein from *Aneurinibacillus thermoaerophilus* DSM 10155. *Glycobiology* **9**, 787–795
- Wügeditsch, T. (1998) Strukturanalyse des S-Schicht Glykoproteinglykans und Zellwand-Aminozuckerpolymers von *Aneurinibacillus thermoaerophilus* DSM 10155. Doctoral thesis, Universität für Bodenkultur Wien, Wien, Austria
- Altman, E., Schäffer, C., Brisson, J.-R. and Messner, P. (1995) Characterization of the glycan structure of a major glycopeptide from the surface layer glycoprotein of *Clostridium thermosaccharolyticum* E207–71. *Eur. J. Biochem.* **229**, 308–315
- Fordham, W. D. and Gilvarg, C. (1974) Kinetics of cross-linking of peptidoglycan in *Bacillus megaterium*. *J. Biol. Chem.* **249**, 2478–2482
- Winzler, R. J. (1955) Determination of serum glycoproteins. *Methods Biochem. Anal.* **11**, 279–311
- Ames, B. N. (1966) Assay of inorganic phosphate, total phosphate and phosphatases. *Methods Enzymol.* **VIII**, 115–118
- Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254
- Heukeshoven, J. and Dernick, R. (1985) Simplified method for silver staining of proteins in polyacrylamide gels and the mechanism of silver staining. *Electrophoresis* **6**, 103–112
- Messner, P., Hollaus, F. and Sleytr, U. B. (1984) Paracrystalline cell wall surface layers of different *Bacillus stearothermophilus* strains. *Int. J. Syst. Bacteriol.* **34**, 202–210
- Schäffer, C., Wügeditsch, T., Kählig, H., Scheberl, A., Zayni, S. and Messner, P. (2002) The surface layer (S-layer) glycoprotein of *Geobacillus stearothermophilus* NRS 2004/3a. Analysis of its glycosylation. *J. Biol. Chem.* **277**, 6230–6239
- Katzenellenbogen, E., Romanowska, E., Witkowska, D. and Shaskov, A. S. (1992) The structure of the O-specific polysaccharide from *Hafnia alvei* strain 38 lipopolysaccharide. *Carbohydr. Res.* **231**, 51–54
- Gronberg, G., Nilsson, U., Bock, K. and Magnusson, G. (1994) Nuclear magnetic resonance and conformational investigations of the pentasaccharide of the Forssman antigen and overlapping di-, tri-, and tetra-saccharide sequences. *Carbohydr. Res.* **257**, 35–54
- Kearle, P. (2000) A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry. *J. Mass Spectrom.* **35**, 804–817
- Meier-Stauffer, K., Busse, H.-J., Rainey, F. A., Burghardt, J., Scheberl, A., Hollaus, F., Kuen, B., Makristathis, A., Sleytr, U. B. and Messner, P. (1996) Description of *Bacillus thermoaerophilus* sp. nov., to include sugar beet isolates and *Bacillus brevis* ATCC 12990. *Int. J. Syst. Bacteriol.* **46**, 532–541
- Naumova, I. B. and Shaskov, A. S. (1997) Anionic polymers in the cell walls of Gram-positive bacteria. *Biochemistry (Moscow)* **62**, 947–982

- 27 Schäffer, C. and Messner, P. (2001) Glycobiology of surface layer proteins. *Biochimie* **83**, 591–599
- 28 Ries, W., Hotzy, C., Schocher, I., Sleytr, U. B. and Sára, M. (1997) Evidence that a secondary cell wall polymer recognizes the N-terminal part of the S-layer protein from *Bacillus stearothermophilus* PV72/p2. *J. Bacteriol.* **179**, 3892–3898
- 29 Sára, M., Egelseer, E. M., Dekitsch, C. and Sleytr, U. B. (1998) Identification of two binding domains, one for peptidoglycan and another for a secondary cell wall polymer, on the N-terminal part of the S-layer protein SbsB from *Bacillus stearothermophilus* PV72/p2. *J. Bacteriol.* **180**, 6780–6783
- 30 Lemaire, M., Miras, I., Gounon, P. and Béguin, P. (1998) Identification of a region responsible for binding to the cell wall within the S-layer protein of *Clostridium thermocellum*. *Microbiology* **144**, 211–217
- 31 Chauvaux, S., Matuschek, M. and Béguin, P. (1999) Distinct affinity of binding sites for S-layer homologous domains in *Clostridium thermocellum* and *Bacillus anthracis* cell envelopes. *J. Bacteriol.* **181**, 2455–2458
- 32 Brechtel, E. and Bahl, H. (1999) *Thermoanaerobacterium thermosulfurigenes* EM1 S-layer homology domains do not attach to peptidoglycan. *J. Bacteriol.* **181**, 5017–5023
- 33 Egelseer, E. M., Leitner, K., Jarosch, M., Hotzy, C., Zayni, S., Sleytr, U. B. and Sára, M. (1998) The S-layer proteins of two *Bacillus stearothermophilus* wild-type strains are bound via their N-terminal region to a secondary cell wall polymer of identical chemical composition. *J. Bacteriol.* **180**, 1488–1495
- 34 Kuen, B., Sára, M. and Lubitz, W. (1995) Heterologous expression and self-assembly of the S-layer protein SbsA of *Bacillus stearothermophilus* in *Escherichia coli*. *Mol. Microbiol.* **19**, 495–503
- 35 Jarosch, M., Egelseer, E. M., Mattanovich, D., Sleytr, U. B. and Sára, M. (2000) S-layer gene *sbsC* of *Bacillus stearothermophilus* ATCC 12980: molecular characterization and heterologous expression in *Escherichia coli*. *Microbiology* **146**, 273–281
- 36 Molinari, M. and Helenius, A. (2000) Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. *Science* **288**, 331–333
- 37 Griffiths, S. G. and Lynch, W. H. (1990) Characterization of *Aeromonas salmonicida* variant with altered surfaces and their use in studying surface protein assembly. *Arch. Microbiol.* **154**, 308–312
- 38 Pei, Z., Ellison, R. T., Lewis, R. V. and Blaser, M. J. (1988) Purification and characterization of a family of high molecular weight surface-array proteins from *Campylobacter fetus*. *J. Biol. Chem.* **263**, 6416–6420
- 39 Schenk, S. P., Earhardt, C. F. and Wyss, O. (1977) A unique envelope protein for *Azotobacter vinelandii*. *Biochem. Biophys. Res. Commun.* **77**, 1452–1458
- 40 Yamada, H., Tsukagoshi, N. and Uda, S. (1981) Morphological alterations of cell wall concomitant with protein release in a protein-producing bacterium, *Bacillus brevis* 47. *J. Bacteriol.* **148**, 322–332
- 41 Montreuil, J. (1984) Spatial conformation of glycans and glycoproteins. *Biol. Cell.* **51**, 115–131
- 42 Martin, J. and Hartl, F.-U. (1997) The effect of macromolecular crowding on chaperonin-mediated protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1107–1112
- 43 Mirelman, D. and Ofek, I. (1986) Introduction to microbial lectins and agglutinins. In *Microbial Lectins and Agglutinins: Properties and Biological Activity* (Mirelman, D., ed.), pp. 1–19, John Wiley & Sons, New York
- 44 Lee, R. T. and Lee, Y. C. (2000) Affinity enhancement by multivalent lectin-carbohydrate interaction. *Glycoconj. J.* **17**, 543–551

Received 26 June 2002/28 August 2002; accepted 29 August 2002

Published as BJ Immediate Publication 29 August 2002, DOI 10.1042/BJ20020988