**N-AcetylMuramnic Acid as Capping Element of α-D-Fucose-containing S-layer Glycoprotein Glycans from Geobacillus tepidamans GS5–97T**

Received for publication, February 15, 2005, and in revised form, March 16, 2005  
Published, JBC Papers in Press, March 21, 2005, DOI 10.1074/jbc.M501724200

Hanspeter Kählig, Daniel Kolarich, Sonja Zayni, Andrea Scheberl, Paul Kosma, Christina Schäffer, and Paul Messner

From the Center for NanoBiotechnology, University of Natural Resources and Applied Life Sciences, A-1180 Vienna, the Department of Organic Chemistry, University of Vienna, A-1090 Vienna, and the Department of Chemistry, University of Natural Resources and Applied Life Sciences, A-1190 Vienna, Austria

Geobacillus tepidamans GS5–97T is a novel Gram-positive, moderately thermophilic bacterial species that is covered by a glycosylated surface layer (S-layer) protein. The isolated and purified S-layer glycoprotein SgtA was ultrastructurally and chemically investigated and showed several novel properties. By SDS-PAGE, SgtA was separated into four distinct bands in an apparent molecular mass range of 106–166 kDa. The three high molecular mass bands gave a positive periodic acid-Schiff staining reaction, whereas the 106-kDa band was nonglycosylated. Glycosylation of SgtA was investigated by means of chemical analyses, 600-MHz nuclear magnetic resonance spectroscopy, and electrospray ionization quadrupole time-of-flight mass spectrometry. Glycopeptides obtained after Pronase digestion revealed the glycan structure \[-2)α-L-Rhap-(1→3)-α-D-Fucp-(1\rightarrow \text{mannose} \text{ residue with } D\text{-fucopyranose having never been identified before as a constituent of S-layer glycans. The rhamnose residue at the nonreducing end of the terminal repeating unit of the glycan chain was di-substituted. For the first time, (R)-N-acetylMuramic acid, the key component of prokaryotic peptidoglycan, was found in an α-linkage to carbon 3 of the terminal rhamnose residue, serving as capping motif of an S-layer glycan. In addition, that rhamnose was substituted at position 2 with a β-N-acetylgalcosamine residue. The S-layer glycan chains were bound via the trisaccharide core \[-2)α-L-Rhap-(1→3)α-L-Rhap-(1→3)-α-L-Rhap-(1→ to carbon 3 of β-D-galactose, which was attached in O-glycosidic linkage to serine and threonine residues of SgtA of *G. tepidamans* GS5–97T.  

Glycosylated surface layer (S-layer) proteins represent an interesting cell surface feature of prokaryotic organisms (1, 2).

In the domain Bacteria, the S-layer glycoproteins of several bacilli and clostridia are best investigated. Recently, the detailed chemical structure and the S-layer glycan biosynthesis (slg) gene cluster of *Geobacillus stearothermophilus* strain NRS 2004/3a have been published (3, 4). These analyses support the general concept of a three-partite architecture of S-layer glycoproteins from organisms of the domain Bacteria. Usually, they are O-glycosidically linked to the S-layer polypeptide, and they possess relatively conserved core structures but extensive variations of the S-layer glycans, even within a given species (1, 2). The size distribution of S-layer glycans can vary from short heterosaccharides (5) up to chains with ~50 repeat units (6).

At the nonreducing end of the S-layer glycan chains of several strains, modifications such as 3-O-methyl or 2-O-methyl groups have been found (3, 7, 8). According to genetic analyses of several slg gene clusters (9), S-layer glycan biosynthesis and LPS O-antigen biosynthesis may follow similar routes (10, 11). These are in principle the Way (polymerase)-dependent and the ATP-binding cassette (ABC) transporter-dependent biosynthesis routes. Recently, nonreducing terminal modifications have been shown to determine the chain length of O-antigens of *Escherichia coli* (12). Currently, the slg gene cluster of *Geobacillus* GS5–97T is under investigation. Preliminary results indicate that the formation of its S-layer glycan may involve enzyme activities comparable with those proposed for the biosynthesis pathway of the *G. stearothermophilus* NRS 2004/3a S-layer glycan (4). A major difference, however, is the absence of O-methyl groups as capping elements at the nonreducing end of the S-layer glycan chains of *G. tepidamans* strain GS5–97T (this paper). Instead, (R)-N-acetylMuramic acid (MurNAc) and GlcNAc, two major constituents of bacterial peptidoglycan (13), have been found. In this report, we describe the complete S-layer glycoprotein glycan structure of *G. tepidamans* strain GS5–97T and suggest that the MurNAc and GlcNAc peptidoglycan components are involved in chain length determination of S-layer glycans.

**Materials and Methods**

Bacterial Strain and Growth Conditions— *G. tepidamans* GS5–97T (ATCC BAA-942\(^\text{a}\), DSM 16325\(^\text{b}\)) was enriched from a sample of sugar beet extraction juice (obtained from Dr. F. Hollaus; Zuckerforschung Tulln GmbH, Tulln, Austria) (14) upon cultivation in modified SVIII medium (SVIII/Glc, containing 0.5% glucose instead of sucrose) under increased aeration at 55 °C. For analysis of the S-layer glycoprotein, the organism was grown under steady-state conditions as described elsewhere (3). The cell pellet was stored at ~20 °C.

Analytical and General Methods—Monosaccharide and amino acid analyses were performed according to published procedures (3, 15).

---

\(\text{a}\) This work was supported by the Austrian Science Fund Projects P15612 and P15840 (to P. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

To whom correspondence should be addressed: Center for NanoBiotechnology, University of Natural Resources and Applied Life Sciences, A-1180 Vienna, Austria. Tel.: 43-1-47654-2203; Fax: 43-1-4789112; E-mail: christina.schaeffef@boku.ac.at.

The abbreviations used are: S-layer, bacterial cell surface layer; ESI Q-Tof MS, electrospray ionization quadrupole time-of-flight mass spectrometry; Fuc, fucose; HMBC, heteronuclear multiple bond correlation spectroscopy; HPAEC/PAD, high performance anion-exchange chromatography/pulsed amperometric detection; HPLC, high performance liquid chromatography; HSQC, heteronuclear single quantum coherence; LPS, lipopolysaccharide; MurNAc, N-acetylmuramic acid; NOESY, nuclear Overhauser enhancement spectroscopy; MS, mass spectrometry; nt, nucleotide; Rha, rhamnose; TOCSY, total correlation spectroscopy; ABC, ATP-binding cassette; GPS, glycopeptide species.

---

\(\text{b}\) S. Zayni, R. Novotny, P. Messner, and C. Schäffer, unpublished data.

---
Optical rotation measurement was performed as described previously (8). Chemical deglycosylation of S-layer glycoprotein was performed by the method of Edge et al. (16). SDS-PAGE with Coomassie Blue staining for protein and periodic acid-Schiff staining for carbohydrate was carried out as described previously (17). N-terminal sequencing of glycoproteins and glycopeptides according to standard protocols (7) was done after semidry blotting (3). Electron microscopy of thin-sectioned and freeze-etched preparations of bacterial cells as well as of negatively stained assembled S-layer glycoprotein was performed according to published procedures (18).

Isolation and Purification of S-layer Glycoprotein Species—The isolation of S-layer glycoprotein and the preparation of S-layer self-assembly products followed published methods (17). For separation of individual glycopeptide species, the S-layer glycoprotein (self-assembly products) was suspended at a concentration of 10 mg/ml in 60 mM Tris-HCl buffer, pH 6.8, containing 2.5% SDS, 0.1 M NaCl, and 0.025% 3-mercaptoethanol, incubated for 2 min at 100 °C, and pelleted. Separation of batchings of 30 mg of S-layer glycoprotein by preparative SDS-PAGE with 25 mM Tris-HCl buffer, pH 8.3, containing 192 mM glycine and 0.1% SDS (running buffer), was performed over 21 h at a constant power of 50 watts at room temperature, using the Bio-Rad model 491 Prep Cell electrophoresis apparatus. The supernatant was loaded onto a gel tube (diameter = 38 mm) consisting of 10% acrylamide gel in 125 mM Tris-HCl buffer, pH 8.8, supplemented with 0.1 M NaCl. Elution was done with running buffer without SDS at a flow rate of 0.8 ml/min; fractions of 8 ml were collected, and the total eluent volume was 1200 ml. Fractions were pooled according to the protein banding pattern on 7.5% SDS-polyacrylamide gels upon silver staining (19). Individual pools were dialyzed at 4 °C against distilled water until salt-free and lyophilized.

Preparation of S-layer Glycopeptides—S-layer glycopeptides were obtained after proteolytic degradation (20 h) of S-layer self-assembly products with V8 protease (Sigma). Purification of glycopeptides is described in detail elsewhere (20), except that chromatofocusing was performed using a pH gradient between 8.5 and 5.5. The glycopeptide pools (designated CF I through CF VI) were desalted using Sephadex G-10 and finally purified by semi-preparative RP(C18)-HPLC (8). Individual CF fractions were lyophilized. Yields of glycopeptides were 1.6 mg for CF I (pH range of 8.5–7.9), 7.2 mg for CF II (pH range of 7.9–7.2), 0.8 mg for CF III (pH range of 7.2–7.0), 8.9 mg for CF IV (pH range of 6.9–6.8), and 1.7 mg for CF V (pH range of 6.8–6.7). After dilution in 50% acetonitrile containing 0.5% formic acid. Spectra were recorded on an Applied Biosystems 470A MALDI TOF mass spectrometer. 

Characterization of S-layer Glycoprotein Species—Starting with ~350 g (wet pellet) of biomass of G. tepidamans GS5–97T, revealed that the bacterium is completely covered with an oblique S-layer glycoprotein lattice. The recently determined lattice parameters (a = 11.2 nm, b = 7.9 nm, γ = 80°) were confirmed in that cultivation (14). All assigned two-dimensional cross-peaks with standard deviations of 0.003 ppm for 1H and 0.015 ppm for 13C. The proton-proton coupling constants were extracted from experimental or calculated (program XSIM 971120, provided by K. Marat, University of Manitoba, Winnipeg, Manitoba, Canada) 1H NMR spectra. 1H-13C coupling constants were taken from appropriate slices of coupled two-dimensional HSQC spectra.

RESULTS

General Description of the Organism—Electron microscopy of freeze-etched cells of G. tepidamans GS5–97T revealed that the bacterium is completely covered with an oblique S-layer glycoprotein lattice. The recently determined lattice parameters (a = 11.2 nm, b = 7.9 nm, γ = 80°) were confirmed in that cultivation (14). The overall degree of glycosylation of the S-layer protein is ~8.5% (w/w). On SDS-PAGE gels, the S-layer glycoprotein showed a four-banded protein staining pattern, with the individual bands corresponding to apparent molecular masses of 106, 124, 130, and 166 kDa. The three high molecular mass bands, but not the 106-kDa band, gave an intensive periodic acid-Schiff staining reaction on the gel. Upon chemical deglycosylation of the S-layer glycoprotein, a single band corresponding to the 106-kDa band remained visible on the gel, indicating that this band represents the nonglycosylated protomeric unit of the S-layer glycoprotein lattice of G. tepidamans GS5–97T.

Characterization of S-layer Glycoprotein Species—Starting with ~350 g (wet pellet) of biomass of G. tepidamans GS5–97T, ~500 mg (dry weight) of S-layer glycoprotein was obtained. Negative staining of the self-assembly products formed upon dialysis of the guanidinium hydrochloride-extracted S-layer glycoprotein revealed predominantly open-ended cylinders of glycoprotein monolayers, with an average diameter of 200 nm and an average length of 1.6 μm. 60 mg of that material was used for purification of the nonglycosylated protomeric unit of the S-layer glycoprotein of G. tepidamans GS5–97T.
Muramic Acid and Fucose as S-layer Glycoprotein Constituents

N-acetylglucosamine:Ser:MurNAc:Ala = 17:20:8:0.8:0.8:1:0.8:1 for CF II and Fuc:Rha:Gal:GlcNAc:Thr:MurNAc:Glx = 16:5:20:0.9:0.9:1:0.9:0.9 for CF IV. (The putative linkage amino acids are underlined.) Setting the putative linkage amino acid Ser (CF II) and Thr (CF IV), respectively, to the value of 1 indicated identical carbohydrate composition of the two glycopeptides. The finding of MurNAc and GlcNAc in molar ratio of 1:1 in both the preparations was initially interpreted as a contamination of these S-layer glycopeptide preparations with peptidoglycan. Repeated RP-HPLC purification, however, did not result in removal of these components. Optical rotation measurement gave a value of $[\alpha]_D^{20} +65^\circ$ (c 0.5, H$_2$O) for CF II. Analysis of possible isomers of N-acetylmuramic acid in glycopeptides CF II and CF IV indicated that, even after spiking with an authentic standard, only one peak appeared on the HPAEC (41.6 min), confirming the exclusive presence of (R)-N-acetylmuramic acid in the S-layer glycopeptides.

**NMR Studies**—Structural investigations by NMR spectroscopy were performed on a sample of 5.3 mg of purified CF II. In the $^1$H NMR spectrum, two intense signals were observed in the anomeric region (Fig. 1, δ = 5.143 and 4.959) as well as two intense aliphatic CH$_3$ groups (δ = 1.312 and 1.195) suggesting two deoxy sugars for the repeating unit. Detailed analysis of these spin systems revealed a rhamnose (sugar E, for the indication of the residues see Fig. 2) and a fucose (sugar F) residue, both as pyranosyl units, and according to the measured coupling constants for the anomeric signals (Table I, $3J_{HH} = 1.0$ and 3.7 Hz and $J_{H,C} = 170.0$ and 169.7 Hz) both in a-configuration. In two-dimensional HMBC spectra, cross-peaks via the glycosidic bonds (Fig. 2 and Table II) proved a value of $[\alpha]_D^{20} +65^\circ$ (c 0.5, H$_2$O) for CF II. Analysis of possible isomers of N-acetylmuramic acid in glycopeptides CF II and CF IV indicated that, even after spiking with an authentic standard, only one peak appeared on the HPAEC (41.6 min), confirming the exclusive presence of (R)-N-acetylmuramic acid in the S-layer glycopeptides.

**FIG. 1.** $^1$H NMR spectrum of glycopeptide CF II showing the anomeric region (the residues are indicated as in Table I and Fig. 5).

**FIG. 2.** HMBC spectrum of glycopeptide CF II showing the glycosidic linkages between the sugar residues (A-F, see Table I and Fig. 5) and from GalpJ to the serine K.

The linkage region to the peptide served as starting point for the more elaborate analysis of the distal and proximal parts of CF II. Typical amino acid signals of a serine residue, O-glycosidically linked to a hexose, could be identified. The corresponding two-dimensional HMBC and two-dimensional NOESY cross-peaks between the anomeric center of the linkage sugar and the serine CH$_2$ group are listed in Table II. Compared with the about 20 times more abundant signal intensity of the disaccharide repeating unit, these signals as well as those for all other terminal residues were very weak in all NMR spectra. A detailed NMR analysis was only performed with CF II. In CF IV, threonine signals were identified; however, the homogeneity of the sample was not sufficient to get reliable linkage information. In CF II, the monosaccharide attached to the serine residue turned out to be a galactopyranose (sugar J) in $\beta$-configuration (Table I, $3J_{HH} = 7.2$ Hz and $J_{H,C} = 161.1$ Hz), connected to the polymer at position 3 according to the $^{13}$C chemical shift of 81.01 ppm. A two-dimensional HMBC cross-peak from carbon 3 of that residue led to an anomeric proton at 5.026 ppm. The trace at this chemical shift in a two-dimensional TOCSY spectrum (mixing time = 120 ms) revealed a signal in the aliphatic region indicative of a 6-deoxy sugar, which was identified as $\alpha$-Rha (sugar I, $J_{H,C} = 171.4$ Hz). In addition, three further rhamnose residues (sugars C, G, and H) and one additional $\alpha$-Fuc residue (sugar D) could be assigned besides the repeating unit constituents. The very small variability in the carbohydrate building blocks of the S-layer glycopeptide obscured the NMR analysis because of severe signal overlap in the spectra.

The comparison of the NMR data from CF II with those of the recently published S-layer glycopeptide from *G. stearothermophilus* NRS 2004/3a (3) enabled further signal assignments. In that organism, the polyrhhamnan S-layer glycan chain is connected via a $\alpha$-3-$\alpha$-Rhap(1$\rightarrow$3)-$\alpha$-Rhap(1$\rightarrow$) to a $\beta$-D-Galp-Ser or a $\beta$-D-Galp-Thr linkage unit. In analogy to these data, the same rhamnose disaccharide could be found 1$\rightarrow$3-linked to the terminal galactose of CF II (sugars H and I). From rhamnose H heteronuclear long range as well as nuclear Overhauser effect correlations established an additional rhamnose G in an $\alpha$-1$\rightarrow$3 linkage (Fig. 2). This latter Rhap residue provides the attachment site to the polysaccharide via an...
| Carbohydrate signals | 2-N-Acetyl  | 3-O-Lactyl |
|----------------------|------------|------------|
|                      | CO | CH₃ | NH | CH | CH₃ |
| A, β-D-GlcNAc-(1→ | | |
| I  4.544 (7.9) | 4.627 | 3.723 | 3.722 | 3.507 | 3.379 | 3.895/3.777 | 2.941 |
| 103.34 (162.2) | 56.85 | 73.50 | 70.38 | 76.88 | 61.35 | 174.91 | 23.42 |
| II  4.504 | 3.769 | 3.771 | 3.514 | 3.365 | 3.293 | 7.949 (8.4) |
| III  4.492 | 3.778 | 3.772 | 3.519 | 3.366 | 3.366 | 3.896/3.781 | 2.937 |
| 103.85 | 56.61 | 73.49 | 70.25 | 76.79 | 61.30 | 23.33 |
| B, α-MurNAc-(1→ | | |
| I  5.271 (3.1) | 3.657 (3.1/10.6) | 3.764 | 3.744 | 3.943 | 3.835/3.719 | 2.129 |
| 103.31 (170.5) | 55.21 | 77.56 | 73.29 | 60.34 | 174.80 | 23.42 |
| II  5.280 | 3.678 | 3.760 | 3.946 | 10.046 (3.3) | 4.620 | 1.436 |
| III  5.292 | 3.598 | 3.767 | 3.743 | 3.942 | 3.839/3.702 | 2.135 |
| 103.85 | 56.61 | 73.49 | 70.25 | 76.79 | 61.30 | 23.33 |
| C, α-L-Rhap-(1→ | | |
| I  5.119 (1.4) | 4.090 | 3.833 | 3.342 | 3.754 | 1.272 |
| 102.09 (173.8) | 77.54 | 71.96 | 71.11 | 71.44 |
| III  5.117 | 4.045 | 3.813 | 3.316 | 3.745 | 1.267 |
| 102.04 | 77.54 | 80.46 | 71.69 | 71.09 |
| D, α-L-Fucp-(1→ | | |
| I  4.977 | 3.877 | 3.988 | 3.924 | 1.184 |
| 98.66 | 68.07 | 77.90 |
| III  4.981 | 3.867 | 3.989 | 1.186 |
| E, α-L-Rhap-(1→ | | |
| I  5.143 (1.0) | 4.065 (1.0/3.1) | 3.975 (3.1/9.7) | 3.554 (9.7/9.6) | 3.862 (9.6/6.1) | 1.312 (6.1) |
| 100.55 (170.0) | 77.02 (150.5) | 72.63 (145.3) | 70.24 (143.0) | 17.33 (128.4) |
| F, α-L-Fucp-(1→ | | |
| I  4.995 (3.7) | 3.981 (3.7/10.7) | 4.018 (10.7/2.0) | 4.019 (10.1) | 4.417 (10.7/2.0) | 1.195 (6.5) |
| 98.67 (169.7) | 67.92 (147.8) | 78.58 (142.5) | 72.60 (147.3) | 15.94 (128.3) |
| G, α-L-Rhap-(1→ | | |
| I  5.165 | 4.067 | 3.973 | 1.322 |
| 100.61 (170.1) | 77.03 | 70.24 |
| H, α-L-Rhap-(1→ | | |
| I  5.026 | 4.171 | 3.908 | 3.576 | 1.307 |
| 100.08 (171.4) | 70.73 | 70.13 |
| II  5.026 | 4.171 | 3.908 | 3.576 | 1.307 |
| 100.08 (171.4) | 70.73 | 70.13 |
| J, α-L-Galp-(1→ | | |
| I  4.533 (7.2) | 3.724 | 3.733 | 4.024 | 3.744 | 3.795/3.758 |
| 103.23 (161.1) | 70.73 | 70.19 |
| III  4.546 | 3.719 | 3.746 | 1.294 |
| 103.23 (161.1) | 70.73 | 61.78 |
|  |  |  |  |  |  |  |  |
| **Amino acid signals** | α | β | γ | NH² | CO |
| K, α-L-Ser- | | | | |
| I  4.343 (6.8/4.2) | 4.279 (11.8/6.8) | 4.172 (11.4/2) |
| 53.84 | 68.09 | 167.48 |
| II  4.337 | 4.337 | 4.192 | 174.93 |
| III  4.335 (7.2/4.6) | 4.251 (11.3/7.2) | 4.193 (11.3/4.6) |
| 53.84 | 67.91 |
| L, α-L-Ala- | | | | |
| I  4.403 (7.3) | 1.428 (7.3) | 174.47³ |
| 50.84 | 17.27 |
| II  4.411 | 1.428 | 174.47³ |
| III  4.412 (7.1) | 1.427 (7.1) |
| 50.69 | 17.22 |
| M, α-L-Aex | | | | |
| I  4.707 | 2.951 | 175.5³ |
| 50.54 | 36.69 | 174.47³ |
| II  4.627 | 2.998 | 8.210 (7.9) |
| III  4.508 (4.9/7.2) | 2.840 (16.2/4.9) | 2.777 (16.2/7.2) |
| 52.27 | 38.35 |

a The chemical shift data set II was measured in a solution of 90% H₂O, 10% D₂O.
b The chemical shift data set III derive from the sample after re-exchange the solvent from H₂O back to D₂O.
c The assignments are interchangeable.
Shift region of the NH protons indicating an amino sugar (Fig. 3). Additionally, there was a second sugar spin system in that area including an anemic signal at 4.5 ppm, partly overlapping with galactose J-H1 (Fig. 1). Both amino sugars could be assigned to be of the 2-acetamido-2-deoxyglucose type, the first one in $\alpha$-(H$_{\text{JH1}}$ = 3.1 Hz and $\text{J}_{\text{HC}}$ = 170.5 Hz) and the latter in $\beta$-configuration ($\text{J}_{\text{H1}}$ = 7.9 Hz and $\text{J}_{\text{HC}}$ = 162.2 Hz). Additionally, the first glucosamine is substituted at position 3 with an $\alpha$-1-carboxydetyl residue, thus corresponding to muramic acid. Diffusion difference experiments (20, 24) clearly proved these amino sugars as being constituents of CF II, as no relative intensity changes were seen between these signals and the rest of the glycan chain upon alteration of the pulsed field gradient amplitude in the stimulated echo experiments. Detailed analysis was finally possible following the aforementioned observation of significant shift differences for this terminal part after exchanging the solvent from D$_2$O to H$_2$O. Repetition of the NMR experiments after re-exchanging the solvent back to D$_2$O moved these signals even further, resulting in a third data set III (Table I). These normally undesirable effects on the chemical shifts, possibly due to changes in the pH or salt concentration, allowed for the elucidation of the complete spin system of the two amino sugars and consequently for determination of the connectivity (Table II). The $\beta$-GlcNAc residue A is 1-2-linked to the $\alpha$-1-Rhap C, as inferred from a long range cross-peak between its anomic proton and carbon 2 of that rhamnose residue. The second set of signals correlating to the $\alpha$-configured GlcNAc residue revealed the presence of a lactyl ether group at carbon 3. This $\alpha$-D-MurNAc B is 1-3-linked, as shown by the appropriate HMBC cross-peak from its H-1 to C-3 of sugar C (Fig. 2).

\section*{N-terminal Sequencing of CF IV—As mentioned previously, full NMR analysis of CF IV was impossible because of insufficient homogeneity of the sample. Only the presence of a Thr residue could be identified unambiguously. Therefore, this sample was split, and one part was subjected to N-terminal sequencing and the other to MS analysis. The major portion of this fraction possessed the sequence TA, but there was also a minor portion present with the sequence TQ.}

\section*{Mass Spectrometry—Mass spectra were recorded for glycopeptides CF II and CF IV and interpreted on the basis of the known S-layer glycan structure elucidated for CF II. Based on the NMR and on the compositional data, the theoretical average masses of the glycopeptide molecules, including the expected numbers of disaccharide repeating units, were calculated and compared with the masses calculated from the experimentally acquired data under consideration of the different charge-specifying adducts (Table IV). Charge states from +3 to +5 were observed with the +4 state giving the best signals, which were therefore used to calculate glycopeptide masses (Fig. 4 and Table IV). Multiple peaks of the same charge state were observed for each glycopeptide due to different combinations of the charge-specifying adducts with hydrogen, sodium, and potassium (Fig. 4 and Table V). In CF II, the overall degree of polymerization of disaccharides varies between $n = 18$ and $n = 22$ with $n = 20$ being the most abundant form. The spectrum of CF IV, which was insufficient for NMR investigations, was evaluated in analogy to the CF II spectrum. The masses exactly matched with the S-layer glycans composed of 18 to 22 disaccharide repeats bound to a threonine residue, with the second amino acid being a glutamine. Additionally, a set of signals corresponding to a minor form of the glycopeptide containing an alanine instead of a glutamine was detected as well. This concurred with the compositional data. Thus, it is conceivable to assume that on SgTA two types of threonine-linked S-layer $\alpha$-glycans are present with the sequence signa-
IV, besides the major form of a Thr/Gln (CF IVa), minor amounts of a Q-TOF MS measurements, is shown in Fig. 5.

Whereas for CF II solely a Ser/Ala/Asp peptide was confirmed, in CF III, besides the major form of a Thr/Gln (CF IVa), minor amounts of a Ser/Ala/Asp peptide were confirmed, in CF IV, besides the major form of a Thr/Gln (CF IVa), minor amounts of a Thr/Ala peptide (CF IVb) were found.

**Table III**

|          | Calculated | Experimental |
|----------|------------|--------------|
|          | D-D (L-L)  | D-L (D-L)    |
|          | C1 C2 C3   | C1' C2 C3    |
| C-D      | 97.6 74.5 103.5 78.5 102.9 77.90 |
| D-E      | 102.7 81.8 98.9 78.1 98.66 78.02 |
| E-F      | 97.6 74.5 103.5 78.5 100.55 78.58 |
| F-G      | 99.9 78.1 98.67 77.03 |
| G-H      | 103.0 78.8 97.6 75.5 100.61 79.43 |
| H-I      | 103.0 78.8 97.6 75.5 103.08 79.07 |
| I-J      | 97.6 78.2 103.5 82.2 103.05 81.01 |

P and P' are the glycosylated and the glycosylating pyranose residue, respectively.

Relative absolute configuration of the residues P and P'.

For the indication of the residues see Table I and Fig. 5.

**DISCUSSION**

S-layer glycoproteins are one of the best investigated systems of prokaryotic protein glycosylation. If present, they are among the most abundant cellular proteins, indicating their pivotal role for bacterial organisms. Although no precise functions of S-layers in general, and S-layer glycoproteins in particular, have been determined thus far, it can be assumed that they provide a selection advantage for bacteria under the competitive conditions of the natural habitat (e.g., provision of a hydrophilic coat comparable with LPS to bacteria; involvement in general cell surface phenomena). This is supported by the observation that the S-layer glycans present in fresh isolates may be lost after prolonged cultivation of the bacteria in rich laboratory media. Usually bacterial S-layer glycans consist strain-specifically of long, highly variable polysaccharide chains, made of identical repeating units with up to 100–120 monosaccharide residues, O-glycosidically linked to the S-layer polypeptide.

In this contribution, we report on the unusual S-layer glycan structure of *G. tepidamans* GS5–97T. The carbohydrate chains consist of disaccharide repeating units with the structure \(-\text{(D}-\text{L-Rha})_n\text{-L-Rha}\) (18–22). L-Rhamnose is a frequent constituent of bacterial polysaccharides, including S-layer glycans (2, 25). D-Fucose, on the other hand, is a rare constituent of bacterial polysaccharides, which is almost exclusively present in LPS O-antigens (26, 27). Because in S-layer glycans of Gram-positive bacteria typical sugar constituents of LPS are frequently found, it is reasonable to assume that in Gram-positive bacteria the ability to glycosylate proteins was acquired during evolution by lateral gene transfer from Gram-negative organisms. This is supported by the presence of transposases within recently sequenced chromosomal S-layer glycan biosynthesis (slg) gene clusters of different bacteria (9), allowing the division of the cluster into distinct groups of genes with similar G + C contents. Depending on the structure of the encoded S-layer glycan, polycistronic slg clusters are approximately 16–25 kb in size, and based on data base comparison, they contain typical glycan biosynthesis components, including nucleotide sugar pathway genes that are clustered in an operon, sugar transferase genes, glycan-processing genes, and transporter genes (9). slg gene clusters from Bacillaceae seem to be much less organized than the clusters encoding the biosynthesis of other bacterial polysaccharides, such as LPS O-antigens of Gram-negative bacteria (28) or exopolysaccharides of lactic acid bacteria (29). Housekeeping genes that map outside the slg gene clusters are additionally required for S-layer protein glycosylation.

**Investigation of the slg gene cluster of *G. stearothermophilus* NRS 2004/3a,** which possesses a polyrhamnan S-layer glycan with about 15 repeating units, that is terminated by 2-O-methylrhamnose (3), revealed, among other proteins, an ABC transporter and a protein with a predicted methyltransferase...
domain (4). Because G. tepidamans GS5–97T is taxonomically related to G. stearothermophilus NRS 2004/3a (3) and preliminary sequencing data indicate the presence of a similarly organized slg gene cluster, it is conceivable to assume that a comparable S-layer glycan biosynthesis pathway occurs in that strain. This would involve the successive addition of sugar residues from nucleotide-activated precursors to the nonreducing end of the nascent saccharide chain by specific glycosyltransferases, utilization of an ATP-binding cassette dependent export system, and transfer of the chain to the O-glycosylation site target sequences on the S-layer polypeptide by ligase. In analogy to the ABC transporter-dependent LPS O-antigen biosynthesis route of Gram-negative bacteria (10, 11), nonreducing terminal modifications seem to play a pivotal role in S-layer glycan chain termination (12). In different LPS O-antigens of Gram-negative bacteria besides the common nonreducing terminal O-methyl groups (30–32), unusual terminal residues such as 3-deoxy-β-manno-2-ulosonic acid (33), or 2,3,4-tri-2,3,4-trideoxy-α-galacturonamide (34) have been described. Hence, the finding that in G. tepidamans GS5–97T MurNAc and GlcNAc residues represented so far unknown capping elements of the S-layer glycan chain was quite unexpected.

MurNAc and GlcNAc are the major glycan constituents of all bacterial peptidoglycans and form [α-1→4]-β-D-glucuronic acid, α-1→3)-β-D-mannosamine, or (1→6)-D-glucuronic acid chains of variable length (13). Most interestingly, MurNAc is known to occur not only in peptidoglycan but has also been found as a rare constituent of repeating units in different LPS O-antigens and capsular polysaccharides (21, 35–37). In the S-layer glycan of G. tepidamans GS5–97T, however, MurNAc and GlcNAc are exclusively present at the nonreducing end of the glycan chain and therefore seem to be involved in chain length determination. This is supported by the absence of a putative methyltransferase gene in the slg gene cluster of that organism. Different from peptidoglycan, the MurNAc residue in the S-layer glycan is α-linked. This change of anomeric configuration may be explained by the action of a different glycosyltransferase, which facilitates the transfer of this predicted capping MurNAc residue to the terminal repeating unit. Similar observations have been made in LPS O-antigens (38, 39). Although not yet experimentally demonstrated, the cellular pool of the nucleotide-activated precursors UDP-MurNAc and UDP-GlcNAc can obviously be accessed by the organism for consecutive addition of MurNAc and GlcNAc to the nonreducing end of the S-layer glycan chains. The mechanism of the S-layer glycosylation pathway is not known yet, but the parallels of this route and ABC transporter-dependent O-antigen biosynthesis are striking, which suggests

Table V
Detailed analysis of ESI Q-TOF MS results for the n = 20 glycoform

![Figure 4](http://www.jbc.org)
that this biosynthesis system is more widespread in bacteria (12).

Acknowledgments—We thank Dr. Michael Puchberger (Department of Chemistry, University of Natural Resources and Applied Life Sciences, Vienna, Austria) for initial 300 MHz NMR measurements and Dr. Herbert Lindner (Institute of Medical Chemistry and Biochemistry, University Innsbruck, Austria) for protein sequence analysis. The Q-TOF Ultima Global was funded by the Austrian Council for Research and Technology Development.

REFERENCES

1. Schafer, C., and Messner, P. (2001) Biochimie (Paris) 83, 591–599
2. Messner, P., and Schafer, C. (2003) in Progress in the Chemistry of Organic Natural Products (Henz, W., Falk, H., and Tamm, C., eds) pp. 51–124, Springer-Verlag, Wien, Germany
3. Schaffer, C., Wugeditsch, T., Kahlig, H., Scheberl, A., Zayni, S., and Messner, P. (2002) J. Biol. Chem. 277, 6230–6239
4. Novotny, R., Schaffer, C., Straus, J., and Messner, P. (2004) Microbiology 150, 953–965
5. Christian, R., Schulz, G., Schuster-Kolbe, J., Allmaier, G., Schmid, E. R., Sleytr, U. B., and Messner, P. (1999) Glycobiology 9, 407–414
6. Altman, E., Cruissen, J.-R., Gagne, S. M., Kolbe, J., Messner, P., and Sleytr, U. B. (1994) Biochim. Biophys. Acta 1217, 71–77
7. Bock, K., Schuster-Kolbe, J., Altman, E., Allmaier, G., Stahl, B., Christian, R., Sleytr, U. B., and Messner, P. (1994) J. Biol. Chem. 269, 7137–7144
8. Schaffer, C., Muller, N., Christian, R., Graninger, M., Wugeditsch, T., Scheberl, A., and Messner, P. (1999) Glycobiology 9, 407–414
9. Novotny, R., Pfoestl, A., Messner, P., and Schaffer, C. (2004) Glycoconj. J. 20, 435–447
10. Reetz, C. R. H., and Whitfield, C. (2002) Acc. Chem. Res. 35, 570–575
11. Valvano, M. A. (2003) Front. Biosci. 8, S452–S471
12. Clarke, B. R., Cuthbertson, L., and Whitfield, C. (2004) J. Biol. Chem. 279, 35709–35718
13. van Heijenoort, J. (2001) Glycobiology 11, 25R–36R
14. Schaffer, C., Franck, W. L., Scheberl, A., Kosma, P., McDermott, T. R., and Messner, P. (2004) Int. J. Syst. Evol. Microbiol. 54, 2301–2306
15. Schaffer, C., Wugeditsch, T., Neuninger, C., and Messner, P. (1996) Microb. Drug Resist. 2, 17–23
16. Edge, A. S., Faltyniek, C. R., Hof, L., Reichert, L. E., and Weber, P. (1981) Anal. Biochem. 118, 131–137
17. Altman, E., Schaffer, C., Brisson, J.-R., and Messner, P. (1995) Eur. J. Biochem. 229, 308–315
18. Messner, P., Hollaus, F., and Sleytr, U. B. (1984) Int. J. Syst. Bacteriol. 34, 202–210
19. Heukeshoven, J., and Dernick, R. (1985) Electrophoresis 6, 103–112
20. Schaffer, C., Dietrich, K., Unger, B., Scheberl, A., Rainey, F. A., Kahlig, H., and Messner, P. (2000) Eur. J. Biochem. 267, 5482–5492
21. Gunawardena, S., Reddy, G. P., Wang, Y., Kolli, V. S., Orlando, R., Morris, J. G., and Bush, C. A. (1998) Carbohydr. Res. 309, 65–76
22. Allman, F., Schweisser, S., and Weber, C. (1995) Glycocon. J. 12, 84–93
23. Schaffer, C., Kahlig, H., Christian, R., Schulz, G., Zayni, S., and Messner, P. (1999) Microbiology 145, 1575–1583
24. Kahlig, H., Dietrich, K., and Dermer, S. (2002) Monatsh. Chem. 133, 589–598
25. Samul, G., and Reeves, P. (2003) Carbohydr. Res. 338, 2503–2519
26. Yoshida, Y., Nakano, Y., Naka, T., Yamashita, Y., and Koga, T. (1999) J. Biol. Chem. 274, 16933–16939
27. Katenellenbogen, E., Kochatchova, N. A., Zakonovys, G. V., Witkowski, D., Bogoljubskaya, M., Shashkov, A. S., Gamian, A., and Knila, Y. A. (2003) Eur. J. Biochem. 270, 2732–2738
28. Reeves, P. R., and Wang, L. (2002) Curr. Top. Microbiol. Immunol. 264, 109–135
29. Jolly, L., and Stingele, F. (2001) Int. Dairy J. 11, 733–745
30. Ito, T., Higuchi, T., Hirobe, M., Hiramatsu, K., and Yokota, T. (1994) Carbohydr. Res. 256, 113–128
31. Forsberg, L. S., Bhat, U. R., and Carlson, R. W. (2000) J. Biol. Chem. 275, 18851–18863
32. Vinogradov, E. (2002) Carbohydr. Res. 337, 961–963
33. Vinogradov, E., Friedl, E., MacLean, L. L., Perry, M. B., Petersens, B. O. Duus, J. O., and Whitfield, C. (2002) J. Biol. Chem. 277, 25070–25081
34. Vinogradov, E., Peppler, M. S., and Perry, M. B. (2000) Eur. J. Biochem. 267, 7230–7237
35. Senchenkova, S. N., Shashkov, A. S., Knila, Y. A., Kochetkov, N. K., Zych, K., and Sidorkczyk, Z. (1996) Carbohydr. Res. 293, 71–75
36. Kochatchova, N. A., Zakonovys, G. V., Bystrova, O. V., Zolkowska, A., Wykrota, M., Shashkov, A. S., Knila, Y. A., and Rozalski, A. (2002) Carbohydr. Res. 337, 1667–1671
37. Bereplov, A. V., Tavrovskaya, A., Shashkov, A. S., Zolkowska, A., Senchenkova, S. N., Rozalski, A., and Knila, Y. A. (2002) Carbohydr. Res. 337, 2463–2468
38. Olslohn, M. M., Petersen, B. O., Duus, J., Haverkamp, J., Thomas-Oates, J. E., Bock, K., and Holst, O. (2000) Eur. J. Biochem. 267, 2014–2027
39. Sedowska, I., Brisson, J.-R., Thibault, P., Richards, J. C., Lam, J. S., and Altman, E. (2000) Eur. J. Biochem. 267, 1640–1650
40. Lipkind, G. M., Shashkov, A. S., Knila, Y. A., Vinogradov, E. V., and Kochetkov, N. K. (1998) Carbohydr. Res. 275, 59–75
41. Lipkind, G. M., Shashkov, A. S., Nifant'ev, N. E., and Kochetkov, N. K. (1992) Carbohydr. Res. 237, 11–22

Muramic Acid and Fucose as S-layer Glycoprotein Constituents 20299

Fig. 5. Schematic drawing of the complete S-layer glycan structure of G. tepidamons GS5–97 derived from different S-layer glycopeptide preparations (CF II and CF IV a, b).

B (R) D-MurNAc α1.3
D-GloNAc α1.3 L-Rha C
A D-Fuc α1.2
E L-Rha α1.3
D-Fuc α1.2 L-Rha α1.3
G L-Rha α1.3
H L-Rha α1.3
I L-Rha α1.3
J D-Gal β1.6 R
K

R = Ser-Ala-Asp in CF II
Thr-Gln in CF IV a
Thr-Ala in CF IV b
N-Acetylmuramic Acid as Capping Element of α-D-Fucose-containing S-layer Glycoprotein Glycans from *Geobacillus tepidamans GS5–97*

Hanspeter Kählig, Daniel Kolarich, Sonja Zayni, Andrea Scheberl, Paul Kosma, Christina Schäffer and Paul Messner

*J. Biol. Chem.* 2005, 280:20292-20299.
doi: 10.1074/jbc.M501724200 originally published online March 21, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M501724200

Alerts:

• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 40 references, 7 of which can be accessed free at http://www.jbc.org/content/280/21/20292.full.html#ref-list-1