Glycoconjugate Journal (1995) **12**: 99–107

Glycan structure of the S-layer glycoprotein of *Bacillus* sp. L420-91

PAUL KOSMA[1], CHRISTINA NEUNINGER[2], RUDOLF CHRISTIAN[3], GERHARD SCHULZ[4], and PAUL MESSNER[2*]

[1] *Institut für Chemie, Universität für Bodenkultur, A-1180 Wien, Austria*

[2] *Zentrum für Ultrastrukturforschung und Ludwig Boltzmann-Institut für Molekulare Nanotechnologie, Universität für Bodenkultur, A-1180 Wien. Austria*

[3] *Scientific Software Comp., A-1140 Wien, Austria*

[4] *Sandoz Forschungsinstitut Wien, A-1235 Wien, Austria*

Received 17 August, 1994, revised 13 October 1994

Preliminary taxonomic characterization of isolate L420-91 has revealed that this organism is closely related to the species *Bacillus aneurinolyticus*. The bacterium is covered by a squarely arranged crystalline surface layer composed of identical glycoprotein subunits with an apparent molecular mass in the range of 109 kDa. A total carbohydrate content of approximately 3.5% (wt/wt) was determined in the purified surface layer glycoprotein. Glycopeptides were obtained after exhaustive Pronase digestion and purification including gel filtration, ion exchange chromatography and HPLC. From the combined evidence of composition analysis, Smith degradation and nuclear magnetic resonance spectroscopy experiments we propose the following structure for the glycan chain of the surface layer glycoprotein:

$$\begin{array}{l}\rightarrow 3)\text{-}\alpha\text{-D-Rha}p\text{-}(1\rightarrow 3)\text{-}\alpha\text{-D-Rha}p\text{-}(1\rightarrow 2)\text{-}\alpha\text{-D-Rha}p\text{-}(1\rightarrow 2)\text{-}\alpha\text{-D-Rha}p\text{-}(1\rightarrow \\ \quad\quad 2 \quad\quad\quad\quad 2 \\ \quad\quad\uparrow \quad\quad\quad\quad\uparrow \\ \quad\quad 1 \quad\quad\quad\quad 1 \\ \quad\alpha\text{-D-Fuc}p3\text{NAc} \quad \alpha\text{-D-Fuc}p3\text{NAc}\end{array}$$

Keywords: Eubacteria, glycoprotein, surface layer (S-layer), glycan structure, ^{1}H and ^{13}C-NMR

Abbreviations: S-layer, crystalline surface layer; SDS-PAGE, sodium dodecyl sulphate polyacrylamide gel electrophoresis; RP-HPLC, reversed phase high performance liquid chromatography; HPAEC, high performance anion exchange chromatography; TFA, trifluoroacetic acid; NMR, nuclear magnetic resonance spectroscopy; COSY, correlated spectroscopy; TOCSY, total correlated spectroscopy; NOE, nuclear Overhauser enhancement; DRha, D-rhamnose; DFuc3NAc, 3-acetamido-3-deoxy-D-fucose; Man, mannose; Gro, glycerol; Thr, threonine, Gly, glycine; Tyr, tyrosine.

Introduction

The occurrence of glycosylated S-layer proteins [1] among eubacteria is now well established [2,3]. In our systematic survey of S-layer glycoproteins from isolates of thermophilic, aerobic bacteria from extraction towers of Austrian beet sugar factories we analysed approximately thirty cultures of the beet sugar campaign 1991. As a result four strains possessed glycosylated S-layer proteins. Among others isolate L420-91 was chosen for further characterization. Preliminary taxonomic analysis of isolate L420-91 [4] positioned this strain in close relationship to the recently reinvestigated species *Bacillus aneurinolyticus* [5]. A full taxonomic description of this organism will be published elsewhere (K. Meier-Stauffer *et al.*, in preparation).

In this communication we report the unusual composition of the repeating unit of the glycan chains of the S-layer glycoprotein of isolate L420-91. The hexasaccharide repeats consist

* To whom correspondence should be addressed at: Zentrum für Ultrastrukturforschung, Universität für Bodenkultur, Gregor-Mendel-Str. 33, A-1180 Wien, Austria.

of α-D-rhamnose and 3-acetamido-3-deoxy-α-D-fucose residues in a ratio of 2 : 1. Both sugars were never found before as constituents of prokaryotic glycoproteins [6] but are common components of lipopolysaccharides of Gram-negative eubacteria [7].

Materials and methods

Growth of bacteria Isolate L420-91 was obtained from Dr F. Hollaus (Österreichisches Zuckerforschungs-Institut, Tulln, Austria) and was grown in continuous culture with aeration (5 l. min^{-1}) at 57°C in S-VIII medium in a 10 l fermenter (Biostat E; Braun, Melsungen, Germany) as described [8].

Analytical methods and electron microscopy Carbohydrate and amino acid analyses, SDS-PAGE, electron microscopy, and sequence analysis were performed according to published methods [9]. The $[\alpha]_D^{20}$ value of the polysaccharide and the material after Smith-degradation was determined using a Perkin-Elmer model 243B polarimeter (Perkin-Elmer, Norwalk, CT).

Isolation of S-layer glycoprotein and S-layer glycopeptide The S-layer glycoprotein was isolated as described previously [9]. After isolation of this material purity was checked both by SDS-PAGE and electron microscopy. After exhaustive Pronase digestion [10] a glycopeptide fraction was eluted in the void volume of the columns after gel filtration over Bio-Gel P-4 (65±20 μm, BioRad; 1.5 × 95 cm) and Bio-Gel P-100 (45–90 μm; 1 × 120 cm). Fractions giving a positive reaction for carbohydrates were pooled, lyophilized, and subjected to cation exchange chromatography on Dowex 50W-X8 resin as described earlier [9]. The material of interest which did not bind to the resin was pooled, lyophilized, and the purity was checked by RP-HPLC on an RP-18 column (Supersphere 100, 4 μm, 8 × 125 mm; Merck) [9]. Fractions with positive reaction for carbohydrates were combined, lyophilized and stored at –18°C.

Periodate oxidation and Smith-type hydrolysis A solution of the glycopeptide (46 mg) in 0.1 M sodium metaperiodate in 0.1 M sodium acetate buffer, pH 4.5 (4 ml) was treated in the dark for 6 days at 4°C [10]. To destroy excess periodate 0.4 ml ethylenglycol was added and the mixture was desalted on a Bio-Gel P-2 column (1 × 120 cm). The material which eluted in the void volume was reduced at pH 8.0 with sodium borohydride (40 mg). After 16 h at 22°C, the solution was neutralized with dilute acetic acid, desalted on the Bio-Gel P-2 column, and lyophilized. Smith-type hydrolysis of the periodateoxidized and reduced polymer was affected with 2.0 M trifluoroacetic acid at 22°C for 48 h and the degradation products were fractionated on a Bio-Gel P-4 column (65±20 μm; 1.5 × 95 cm). The purity of this material was determined by RP-HPLC on the RP-18 column as described before. The isolated purified product (12 mg) was analysed by NMR before and after sodium borohydride reduction and an aliquot (7 mg) was subjected to further Smith degradation. Prior to hydrolysis with 0.5 M trifluoroacetic acid the sample was desalted on a Bio-Gel P-2 column. Upon separation of the hydrolysis products on the Bio-Gel P-2 column (1 × 120 cm) and reduction final purification of the degradation products was performed by HPLC on a TSK HW40S column (1.6 × 100 cm; Merck) using water as eluent.

NMR experiments NMR experiments were performed on solutions of the poly- and oligosaccharides in 500 μl of ^2H_2O (99.8 atom % ^2H; Merck) on a Bruker AC 300 F instrument (^1H: 300.13 MHz, ^{13}C: 75.47 MHz). ^1H chemical shifts (δ) were determined at 297 K and 330 K, respectively, and referenced to an external standard sodium 3-trimethylsilyl-(2,2,3,3-$^2H_4$) propionate (δ = 0.00) in ^2H_2O. ^{13}C-NMR spectra were recorded at 297 K and 330 K over a spectral width of 25 kHz and reported relative to 1,4-dioxane as an external standard at 67.40 ppm.

The homonuclear ^1H-^1H chemical shift-correlated two-dimensional spectra were obtained at 297 K using the COSY-45° pulse sequence [11] and one and two step Relayed-COSY spectra [12] were obtained accordingly. The two-dimensional heteronuclear-correlated experiments were performed with simultaneous ^1H broadband decoupling using the Bruker pulse program XHCORR. Refocusing delays were adjusted to an average $J_{C\text{-}H}$ coupling constant of 145 Hz. Spectral widths of 10 000 Hz with 4096 data points, for ^{13}C, and 800 Hz with 512 data points, for ^1H were employed.

NOE measurements were performed in the difference mode with sequential irradiation of each line in a multiplet [13] at 300 and 500 MHz. Spin simulations were performed using a modified version of QCPE (Quantum Chemistry Program Exchange, Program 458) written by R. Christian.

Spectra of the reduced tetrasaccharide and all TOCSY spectra were recorded on a BRUKER AMX 500 instrument (^1H: 500.14 MHz) at 300 and 330 K. ^1H-^{13}C correlation spectra were recorded in the proton-detected mode at 300 K with a Bruker 5 mm inverse broad-band probe using Bruker reverse electronics.

Results

Characterization of S-layer glycoprotein and glycopeptide Micrographs of freeze-etched intact cells of isolate L420-91 showed that the bacteria are completely covered by a square S-layer lattice with a centre-to-centre spacing of the morphological units of approximately 10.4 nm (Fig. 1). By SDS-PAGE and periodic acid-Schiff staining reaction it was demonstrated that the S-layer protomer with an apparent molecular weight of 109 000 was the only carbohydrate-positive staining band on the gel. Chemical analysis of this material yielded a total of approximately 3.5% (wt/wt) carbohydrates with 1.4% (wt/wt) rhamnose and small amounts of mannose and *N*-acetylgalactosamine. Most probably these sugars are part of a core region between the repeating units

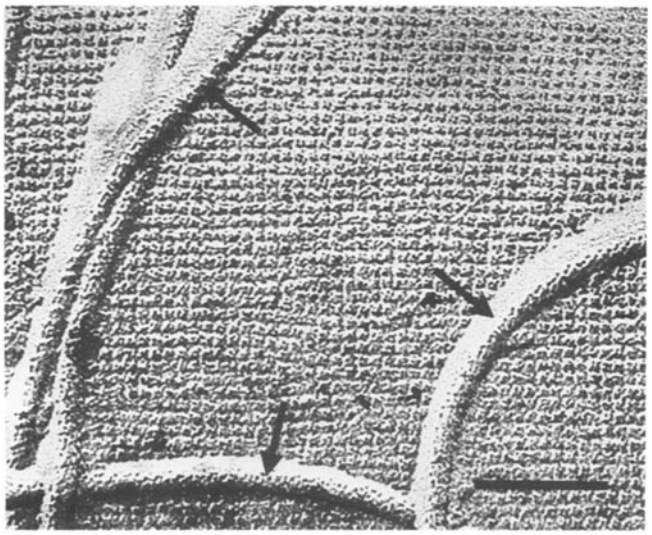


Figure 1. Electron micrograph of a freeze-etched preparation of an intact cell of *Bacillus* sp. L420-91 showing the square S-layer lattice; arrows, flagella. Bar = 100 nm.

and the polypeptide chain [9]. An unidentified peak in the HPAEC chromatogram was later identified by NMR as Fuc3NAc.

Purification of the glycopeptides after Pronase digestion was performed by gel filtration on Bio-Gel P-4 and P-100 columns. Even on the P-100 column the glycopeptide eluted in the void volume. To prove the purity of this material an aliquot was subjected to RP-HPLC on an RP-18 column. A broadened peak was obtained which, by HPAEC, showed Rha and Fuc3NAc in a ratio of 2 : 1 as the major constituents and small amounts of Man and amino acids. Fuc3NAc was identified by retention time in the HPAEC run. Amino acid and sequence analysis revealed that Thr and Gly were present in the peptide portion in a ratio of 1 : 1. However, there was no indication for the presence of Tyr which was frequently observed in other eubacterial S-layer glycoproteins [2,9]. Optical rotation measurement of this glycopeptide with the intact polysaccharide chain showed an $[\alpha]_D^{20}$ value of +130° (*c* 0.8, water).

Smith degradation The glycopeptide (compound **1**) was subjected to periodate oxidation and Smith-type hydrolysis with 2.0 M TFA. The major reaction product eluted from the Bio-Gel P-4 column with a K_{av} = 0.30. Composition analysis by HPAEC indicated the presence of equimolar amounts of Rha and Fuc3NAc. By NMR analysis this material was identified as a tetrasaccharide linked to a glyceraldehyde residue (compound **2**). An $[\alpha]_D^{20}$ value of +130° (*c* 0.9, water) was determined for this material. After reduction with $NaBH_4$ a stable product with a proximal glycerol residue was obtained (compound **3**; Fig. 2). Upon additional Smith degradation one of the hydrolysed and reduced degradation products eluted from

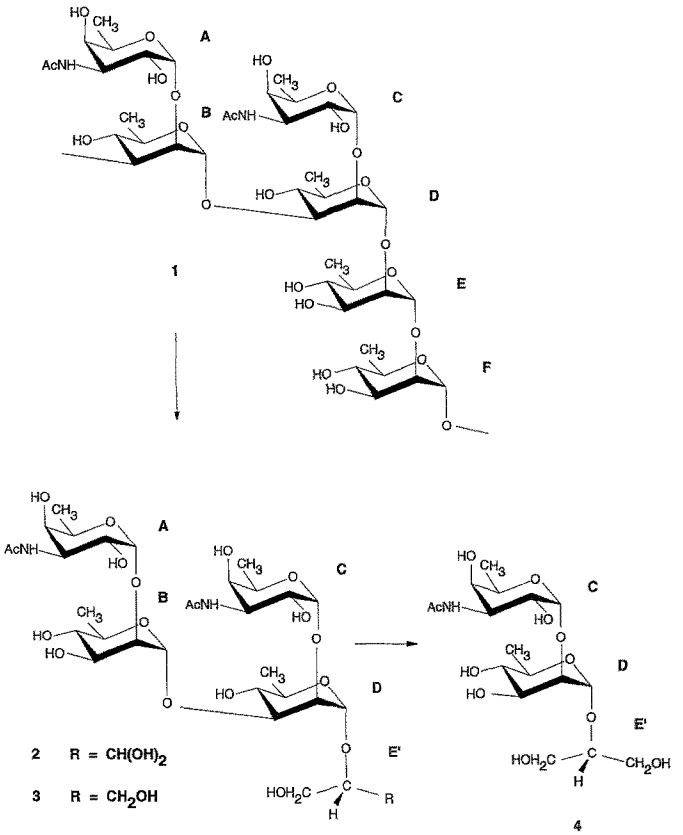


Figure 2. Repeating unit structure of the S-layer glycopeptide of *Bacillus* sp. L420-91 and its Smith degradation products.

the Bio-Gel P-2 with K_{av} = 0.40. Its structure was identified by NMR to be the disaccharide Fuc3NAc-Rha-Gro (compound **4**; Fig. 2; Tables 1–3).

NMR measurements The 300 MHz ^{1}H-NMR spectrum of glycopeptide **1**, recorded at 330 K, showed the presence of six protons in the range 5.10–5.24 ppm (four signals with couplings of 1.5 Hz, two signals with couplings of 3.9 Hz), twenty-two protons between 3.60–4.20 ppm, one signal of two protons at 3.49 ppm (dd, 2 × 10 Hz), two signals corresponding to six protons at 2.07 ppm and six methyl groups (coupling of 6.3–6.5 Hz) at 1.19–1.35 ppm. Comparison of the integral of the major signals to that of some signals with minor intensity indicated a chain length of at least fifteen repeats.

The proton-decoupled ^{13}C-NMR spectrum, recorded at 330 K, contained two signals attributable to carbonyl carbons (175.12 and 175.24 ppm), four signals for six anomeric carbons of pyranoid sugars (101.09–101.72 ppm), and five signals of six carbons corresponding to linkage sites (77.39–78.80 ppm). There were twelve signals of sixteen carbons in the range 67.2–73.3 ppm and two signals indicating carbons connected to nitrogen at 52.22 and 52.31 ppm. The two signals of methyl groups at 22.90 and 22.96 ppm together

Table 1. Carbon shifts[a].

Unit /		*Temperature*	*Carbon*							
	Compound	*(K)*	*1*	*2*	*3*	*4*	*5*	*6*	*7*[b]	*8*[b]
A	**1**	330	101.09	67.18	52.31	71.19	68.04	16.11	175.24	22.96
A	**3**	330	101.42	67.37	52.09	71.21	68.07	16.05	175.18	22.89
A	**3**	300	101.51	67.29	52.00	71.07	68.08	16.12	175.26	22.81
A	**2**	297	101.49	67.30	52.00	71.09	68.08	16.13	175.26	22.81
B	**1**	330	101.72	78.71	77.39	72.96	70.27	17.44		
B	**3**	330	101.60	80.22	70.97	72.99	70.22	17.50		
B	**3**	300	101.60	80.22	70.97	72.99	70.22	17.50		
B	**2**	297	101.60	80.19	70.99	73.00	70.23	17.51		
C	**1**	330	101.22	67.28	52.22	71.19	68.07	16.13	175.12	22.90
C	**3**	330	101.15	67.27	52.24	71.21	68.06	16.07	175.18	22.89
C	**3**	300	101.36	67.14	52.11	71.07	68.08	16.12	175.31	22.81
C	**2**	297	101.30	67.16	52.12	71.09	68.08	16.13	175.30	22.81
C	**4**	300	101.68	67.39	52.09	71.18	68.22	16.20	175.33	22.90
D	**1**	330	101.57	78.71	77.59	73.30	70.46	17.36		
D	**3**	330	99.16	79.26	77.94	73.05	69.98	17.34		
D	**3**	300	99.19	79.34	77.91	72.99	69.90	17.27		
D	**2**	297	99.36	79.25	77.94	73.00	69.99	17.25		
D	**4**	300	99.08	80.91	71.18	73.20	69.86	17.50		
E	**1**	330	101.58	78.80	70.93	73.16	70.16	17.63		
E′	**3**	330	62.33	79.40	61.61					
E′	**3**	300	62.23	79.43	61.46					
E′	**2**	297	89.91	81.03	60.76					
E′	**4**	300	62.29	79.35	61.40					
F	**1**	330	101.54	78.54	70.98	73.16	70.09	17.57		

[a] Shifts are relative to TMS (dioxane at 67.4) at the temperature indicated.
[b] Assignment of acetyl group signals are based on intensity.

with the carbonyls indicate these two carbons to be acetylated. Five signals of six carbons at 16.11–17.63 ppm suggest the presence of 6-deoxy sugars. The relative number of carbons per signals was determined by integrating the spectra. The absence of CH_2-groups was confirmed by means of a DEPT135 experiment. Thus, the polysaccharide was assumed to be composed of hexasaccharide repeating units which contain only 6-deoxy sugars, two of them having an additional acetamido group. Further assignments were made with the aid of GATED-decoupled ^{13}C-NMR spectra, which confirmed the presence of two *N*-acetyl groups ($^2J_{C,H}$~6 Hz). The anomeric carbon signals had $^1J_{C,H}$ in the range of 172–176 Hz which, in conjunction with the small $^3J_{H,H}$ coupling constants, are indicative of equatorially oriented anomeric protons [14]. The cluster of the signals centred around 78.7 ppm (three signals of four carbons in the decoupled spectrum) displayed $^1J_{C,H}$ values of 152 Hz, which are consistent with axially oriented oxygens at C-2 of rhamnopyranosyl residues [15], whereas the smaller values of 146 Hz at 77.39 and 77.59 are indicative of an axial arrangement of hydrogen. Similarly, the spacings of 140 to 146 Hz for the carbon signals at 52.22/52.31, 67.18/67.28 and 68.07/68.04 are compatible with axial hydrogens.

Due to line-broading and signal overlap at 300 MHz, first order $^3J_{H,H}$ values could only be extracted for a limited number of signals. Some of the proton-proton and proton-carbon connectivities were established using two-dimensional techniques and thus the two protons at 3.49 ppm (triplet, 10 Hz) attributable to H-4 of α-rhamnose systems correlated to 73.16 ppm in the carbon spectrum. Furthermore, the nitrogen bearing carbons had connected protons assigned as H-3 and are therefore 3-acetamido sugars.

For further structural analysis the glycopeptide was subjected to Smith degradation giving tetrasaccharide **2**, amenable to analysis by 300 MHz ^{1}H-NMR spectroscopy. The ^{1}H-NMR spectrum of **2**, recorded at 297 K, displayed five protons in the anomeric region, nineteen protons between 3.50–4.20 ppm,

Table 2. Proton shifts[a]

Unit /		*Temperature*	*Proton*[b]						
	Compound	*(K)*	*1(1a)*	*2(1b)*	*3(2)*	*4(3a)*	*5(3b)*	*6*	*7*[c]
A	**1**	330	5.155	3.826	4.195	3.774	4.110	1.185	2.066
A	**3**	330	5.076	3.826	4.195	3.753	4.153	1.165	2.053
A	**3**	300	5.076	3.824	4.208	3.752	4.161	1.164	2.055
A	**2**	297	5.076	3.823	4.208	3.749	4.159	1.164	2.055
B	**1**	330	5.243	4.109	3.955	3.763	3.855	1.354	
B	**3**	330	5.298	4.060	3.858	3.593	3.801	1.323	
B	**3**	300	5.338	4.067	3.872	3.615	3.818	1.332	
B	**2**	297	5.344	4.066	3.877	3.620	3.826	1.335	
C	**1**	330	5.113	3.826	4.189	3.775	4.135	1.195	2.070
C	**3**	330	5.118	3.821	4.185	3.765	4.132	1.182	2.060
C	**3**	300	5.127	3.824	4.198	3.765	4.139	1.182	2.062
C	**2**	297	5.134	3.823	4.198	3.763	4.131	1.180	2.062
C	**4**	300	5.097	3.841	4.221	3.764	4.201	1.181	2.063
D	**1**	330	5.170	4.172	4.002	3.786	3.779	1.312	
D	**3**	330	5.122	4.051	4.008	3.739	3.901	1.310	
D	**3**	300	5.138	4.075	4.009	3.756	3.926	1.316	
D	**2**	297	5.160	4.104	4.012	3.760	4.006	1.310	
D	**4**	300	5.196	4.024	3.927	3.613	3.882	1.314	
E	**1**	330	5.095	4.114	3.921	3.492	3.733	1.293	
E′	**3**	330	3.721	3.705	3.816	3.768	3.678		
E′	**3**	300	3.726	3.706	3.822	3.776	3.680		
E′	**2**	297	5.105	–	3.661	3.881	3.743		
E′	**4**	300	3.721	3.701	3.809	3.775	3.675		
F	**1**	330	5.209	4.060	3.867	3.495	3.775	1.313	

[a] Proton shifts are iterated values and are accurate to 0.001 ppm.
[b] Numbering of glycerol is given in parentheses.
[c] Assignment of acetyl groups are based on intensity.

two CH_3CO signals at 2.05 and 2.06 ppm, two CH_3-groups at 1.31 and 1.32 ppm and two CH_3-signals at 1.17 and 1.18 ppm, indicating loss of two 6-deoxy sugar residues during Smith degradation. The presence of a 2-linked glyceraldehyde residue was inferred from 75.47 MHz ^{13}C-NMR spectra which showed signals at 89.91, 81.03, and 60.76 ppm, indicative of this Smith degradation fragment [16]. Furthermore, since the spectrum contained thirtyone carbon signals, and taking into account the presence of two acetamido groups, the structure was assigned to a tetrasaccharide unit. Due to the inherent instability of the aldehyde group, the tetrasaccharide was subsequently reduced to give compound **3,** which was completely analyed by 500 MHz ^{1}H-NMR spectroscopy using TOCSY, inverse C,H-correlation, and NOE measurements.

H,H and C,H-connectivities as well as $J_{H,H}$ values established the structure of the two 6-deoxy sugar residues B and D as α-rhamnopyranosyl residues, which was evident from the *trans* diaxial $J_{3,4}$ and $J_{4,5}$ values (9.8 Hz) as well as the smaller coupling constants $J_{2,3}$ (3.4 and 3.2 Hz) and $J_{1,2}$ (1.7 Hz). Similarly, the nitrogen connected carbon signals of units A and C at 52.00 and 52.11 ppm were correlated to H-3 signals with one *trans* diaxial coupling to H-2 (11.2 and 11.4 Hz) and a smaller value for $J_{3,4}$ (3.1 Hz), whereas H-4 revealed a coupling constant $J_{4,5}$ of 1.0 Hz. Thus, residues A and C were identified as 3-acetamido-3,6-dideoxy-α-galactopyranosyl (αFuc3NAc) residues. Since the ^{13}C-NMR data, which are also in close agreement with published values [17] for αFuc3NAc-methyl glycoside, contain no downfield shifted signals, residues A and C correspond to unsubstituted side chain residues. Further analysis of the ^{13}C-NMR chemical shift values revealed the presence of a 2-monosubstituted and a 2,3-disubstituted α-rhamnopyranosyl residue. Therefore, at least one Fuc3NAc residue has to be linked to the O-2 of rhamnopyranosyl unit B. Considering the upfield shift of the

Table 3. Proton-proton coupling constants[a]

Unit / Compound		*Temperature (K)*	*Ja,b* 1,2 [1a,1b/1a,2	1b,2	2,3 2,3a/2,3b	3a,3b][b]	3,4	4,5	5,6
A	**1**	330	3.8		11.4		3.1	1.0	6.6
A	**3**	330	3.9		11.2		3.1	1.2	6.6
A	**3**	300	3.8		11.4		3.1	1.0	6.6
A	**2**	297	3.8		11.4		3.1	1.0	6.6
B	**1**	330	1.6		3.1		9.8	9.8	6.3
B	**3**	330	1.7		3.4		9.8	9.8	6.3
B	**3**	300	1.8		3.4		9.9	9.7	6.2
B	**2**	297	1.8		3.4		9.9	9.7	6.2
C	**1**	330	3.8		11.4		3.1	1.0	6.6
C	**3**	330	3.8		11.4		3.1	1.1	6.6
C	**3**	300	3.8		11.4		3.1	1.0	6.6
C	**2**	297	3.8		11.4		3.1	1.0	6.6
C	**4**	300	3.8		11.4		3.1	1.0	6.6
D	**1**	330	1.6		3.3		9.8	9.8	6.3
D	**3**	330	1.7		3.2		9.7	9.6	6.2
D	**3**	300	1.8		3.2		9.8	9.8	6.3
D	**2**	297	1.8		3.2		9.8	9.8	6.3
D	**4**	300	1.6		3.4		9.9	9.7	6.3
E	**1**	330	1.6		3.2		9.8	9.8	6.3
E′	**3**	330	−12.0/4.0	5.9	3.8/6.0	−12.3			
E′	**3**	300	−12.1/4.1	5.7	3.9/6.1	−12.3			
E′	**2**	297	4.7		3.2/6.4	−12.8			
E′	**4**	300	−12.1/4.1	5.7	3.9/6.1	−12.4			
F	**1**	330	1.6		3.2		9.8	9.8	6.3

[a] Proton coupling constants are iterated values and are accurate to 0.2 Hz (digital resolution).
[b] Numbering of glycerol part.

C-1 signal (99.19 ppm) in unit D, this 2,3-disubstituted residue must be linked to the glycerol entity, which leaves two possibilities for O-2 or O-3 as attachment sites for the rhamnose and the second Fuc3NAc residue. Final verification was obtained by NOE difference measurements at 300 K. At 300 MHz, upon irradiation of H-1 in unit B, inter-residue NOEs were observed for H-3 of residue D (+4.2%) and H-5 of Fuc3NAc residue A (+3.1%) which confirms the 1→3 linkage between the rhamnosyl units. In addition, as frequently observed for 3-substituted manno systems, when substituted by an α-manno system of the same stereochemical series, an NOE was observed at the overlapping positions of H-2 of units B and D (intra- and inter-residue NOE of 4.2%). Discrimination between the Fuc3NAc residues was not straightforward. Irradiation of the H-1 in unit A at 500 MHz, resulted in a small positive NOE (+1.6%) for H-2 of unit B, a negative NOE (−2.6%) was seen for H-2 of residue A. However, since both H-2 signals of B and D were very similar in chemical shift, it was of crucial importance to check for a valid assignment. Most notably, although identical signal patterns occurred at 300 and 330 K, at elevated temperature both H-2 signals B and D changed place. However, selective decoupling of H-3 of unit B verified the assignment of H-2 at 300 K. Final proof was achieved by spin-simulation of the ^{1}H-NMR spectra, which showed close agreement between calculated and experimental spectra (Fig. 3). Additional confirmation of the structural assignments was obtained by a second Smith degradation/reduction performed with the tetrasaccharide derivative **2**, which furnished the disaccharide derivate **4**. Analysis of the ^{1}H and ^{13}C-NMR data (Tables 1–3) revealed the presence of an α-rhamnosyl unit (D) linked to the O-2 of glycerol, which is substituted in position 2 by an

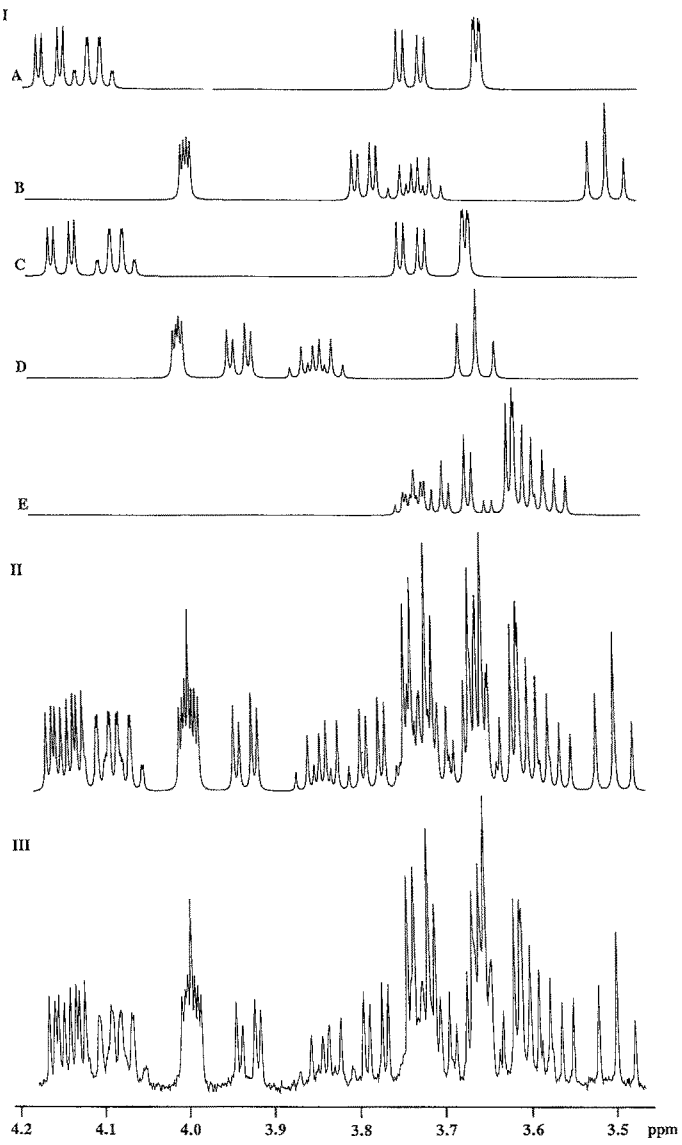


Figure 3. Partial 500 MHz ^{1}H-NMR spectrum (3.4–4.2 ppm) of the reduced tetrasaccharide (compound **3**). I, simulated subspectra of units A–E′; II, simulated spectrum (units A–E′); III, experimental spectrum (for details see Results).

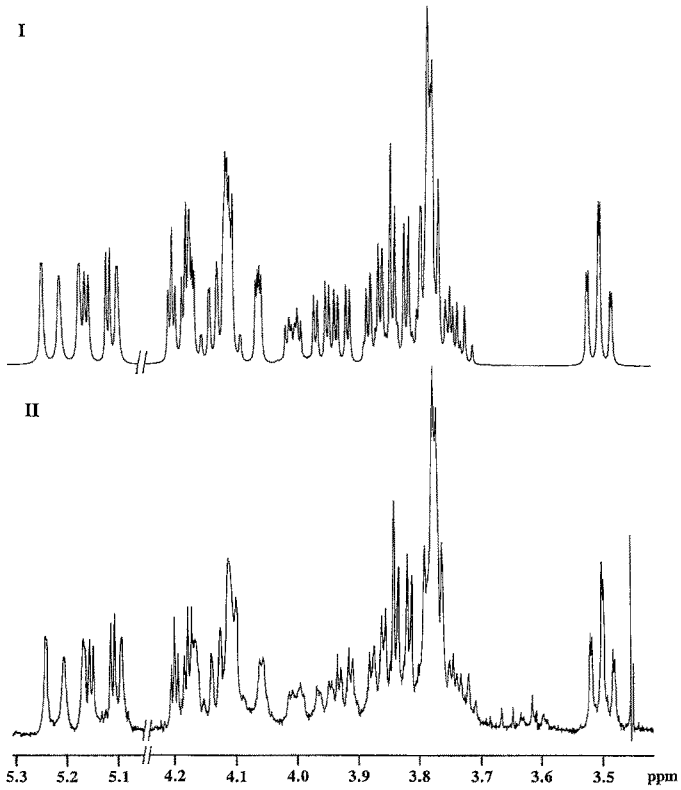


Figure 4. Partial 500 MHz ^{1}H-NMR spectrum (3.4–5.3 ppm) of the polysaccharide (compound **1**). I, simulated spectrum (units A–F); II, experimental spectrum (for details see Results).

αFuc3NAc residue (C). Thus, the occurrence of two α-(1→2)-connected Fuc3NAc monosaccharide side chain units was unambiguously established.

The optical rotation determined for **3**, $[\alpha]_D^{20}$ +149° (*c* 0.4, water), is in between the calculated value [18–20] of +165° (all sugars D) and +136° (one rhamnose L). Any other combination would be significantly less dextro rotating (e.g. +103°, two rhamnoses L; −165°, all sugars L). Since Klyne's rule was not developed for sterically crowded arrangements as in the 2,3-substituted rhamnose, it can be safely assumed that the rhamnose involved in the branch does not contribute to the optical rotation. In this case the calculated value of +149° (two DFuc3NAc and one DRha) perfectly fits the measured value.

The former assumption that the rhamnose involved in the branch does not contribute to optical rotation and both rhamnoses belong to the same stereochemical series is verified by the similar ^{13}C chemical shifts observed for both Fuc3NAcs and the substitution shifts measured for both rhamnoses. If DFuc3NAc is linked to LRha a significant upfield shift should have been observed for carbon-1 of Fuc3NAc (97–98 ppm) and carbon-2 of Rha [7].

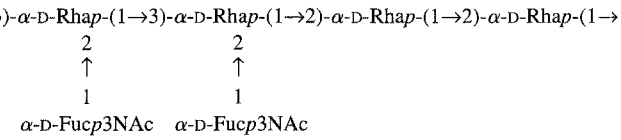
Assignment of the polysaccharide signals Having established the shifts and coupling constants of the individual sugar subunits in the tetrasaccharide compound **3** (Fig. 3), it was possible to refine iteratively the 500 MHz 1-D proton NMR spectrum of polysaccharide **1** (Fig. 4, partial spectrum) using the proton shift positions derived from a TOCSY experiment and the coupling constants from the tetrasaccharide subunits as a starting guess. Using the results of the iteration in combination with that from CH-correlation, the presence of three pairs of closely related sugars was obvious: two unsubstituted αFuc3NAc residues (linked to two neighbour rhamnoses as evident from Smith degradation), two 2,3-disubstituted α-rhamnoses, and two 2-monosubstituted α-rhamnoses.

Since the carbon signals of the aforementioned three pairs of closely related sugar subunits were too similar for a safe differentiation, and since the proton spectrum was too

crowded to make use of NOE spectroscopy, the final assignment of the pairs was made using the following argument. The anomeric protons of both rhamnoses B (5.24 ppm) and F (5.21 ppm) linked to position 3 of a rhamnose, which is substituted in 2 by Fuc3NAc, should be shifted downfield similar to the shift at 5.3 ppm observed in the tetrasaccharide for unit B. By comparison, the anomeric protons of rhamnoses D and E, which are linked to position 2 of the monosubstituted rhamnoses, appeared upfield (5.17 ppm for D and 5.10 ppm for E). Finally, visual inspection of CPK models of the poly- and tetrasaccharide revealed that Fuc3NAc units designated C should be the one most similar in proton and carbon shifts due to their similar chemical environment (Figs. 3, 4).

Because of a similar sterical arrangement of the branched parts in compounds **3** and **1** which is in agreement with the NMR data (Tables 1–3) the $[\alpha]_D^{20}$ value of +130° can be explained. It is close to the calculated value of +132° (two DFuc3NAc and two DRha), whereas +157° would be expected for two DFuc3NAcs and four DRhas. Again, the verification, that the lower $[\alpha]_D$ value is not caused by one LRha, is based on ^{13}C NMR data where the similarity of substitution shifts indicates that all sugars belong to the D-series. Upfield shifts for carbon-1s would have to be expected if one of the sugars belongs to the L-series [7].

Therefore, from the combined evidence the following structure for the polysaccharide is proposed:



Discussion

The ^1H and ^{13}C-chemical shift data, in conjunction with proton NOE data from the S-layer glycoprotein of *Bacillus* sp. L420-91 established its repeating unit structure (Fig. 2). The configuration and $^4C_1$-conformation of the monosaccharide constituents was confirmed by comparison of spin simulated and experimental spectra. Optical rotation measurements of the glycopeptide with the intact polysaccharide chain and the Smith degradation products confirmed these observations. From the absence of tyrosine signals in the proton NMR spectrum of the polysaccharide [2, 9] it was obvious that the glycan chain was not linked via the recently discovered O-glycosidic tyrosine linkages in this glycopeptide. The type of linkage is not yet known.

Like *Bacillus stearothermophilus* [6], *Thermoanaerobacter (Clostridium) thermohydrosulfuricus* [2, 6, 9], *Clostridium thermosaccharolyticum* [2, 6] and most other eubacteria which possess glycosylated proteins [2], *Bacillus* sp. L420–91 belongs to the Bacillaceae family. The structural analysis of its glycan chain revealed two notable differences in comparison with previously analysed S-layer glycoproteins. First, the complex hexasaccharide repeating unit is composed only of Rha and Fuc3NAc residues in the D-configuration which is uncommon in S-layers but frequent in lipopolysaccharides of Gram-negative bacteria [7]. Second, as already mentioned, there is yet no indication for the presence of an O-glycosidic linkage via tyrosine which was observed in several eubacterial S-layer glycoproteins [2, 9]. A detailed analysis of this linkage region is currently in progress.

Although to date nothing is known about the biological functions of glycosylated S-layers in eubacteria the close similarity of the glycan structure to O-antigens of lipopolysaccharides [7] is intriguing. Whether this reflects a functional relationship of the glycans of both types of structures remains to be established. Further analyses are required to unravel the principles of biosynthesis and function of the glycans of this S-layer glycoprotein.

Acknowledgements

We thank Manuela Glozik and Sonja Zayni for excellent technical assistance, Dr Karola Vorauer for sequencing, and Professor Uwe B. Sleytr for continuous interest in this project. The work was supported in part by grants from the Austrian Science Foundation, project S7201-MOB and the Austrian Ministry of Science and Research.

References

1. Messner P, Sleytr UB (1992) In *Advances in Microbial Physiology*, vol. 33 (Rose AH, ed). pp. 213–75. London: Academic Press.
2. Messner P, Schuster-Kolbe J, Schäffer C, Sleytr UB, Christian R (1993) In *Advances in Bacterial Paracrystalline Surface Layers*, (Beveridge TJ, Koval SF, eds). pp. 95–107. New York: Plenum Press.
3. Beveridge TJ (1994) *Curr Opinion Struct Biol* **4**:204–12.
4. Meier-Stauffer K, Kuen B, Busse H-J, Hollaus F, Lubitz W, Sleytr UB, Messner P (1994) *Proc 7th Int Congr Bacteriol Appl Microbiol*. Prague, Abstr. BC-1/30.
5. Shida O, Takagi H, Kadowaki K, Yano H, Abe M, Udaka S, Komagata K (1994) *Int J Syst Bacteriol* **44**:143–50.
6. Messner P, Sleytr UB (1991) *Glycobiology* **1**:545–51.
7. Lipkind GM, Shashkov AS, Nifant'ev NE, Kochetkov NK (1992) *Carbohydr Res* **237**:11–22.
8. Sleytr UB, Sára M, Küpcü Z, Messner P (1986) *Arch Microbiol* **146**:19–24.
9. Bock K, Schuster-Kolbe J, Altman E, Allmaier G, Stahl B, Christian R, Sleytr UB, Messner P (1994) *J Biol Chem* **269**:7137–44.
10. Altman E, Brisson J-R, Messner P, Sleytr UB (1991) *Biochem Cell Biol* **69**:72–8.
11. Bax A, Freeman R, Morris G (1981) *J Magn Reson* **42**:164–70.
12. Bax A, Drobuy G (1985) *J Magn Reson* **61**:306–20.
13. Neuhaus D (1983) *J Magn Reson* **53**:109–14.
14. Bock K, Lundt I, Pedersen C (1973) *Tetrahedron Lett* **13**:1037–40.
15. Christian R, Schulz G, Unger FM, Messner P, Küpcü Z, Sleytr UB (1986) *Carbohydr Res* **150**:265–72.

16. Lipkind GM, Nifant'ev NE, Shashkov AS, Kochetkov NK (1990) *Can J Chem* **68**:1238–50.
17. L'vov V, Tochtamysheva N, Shashkov AS, Dmitriev L, Capek K (1983) *Carbohydr Res* **112**:233–7.
18. Klyne W (1950) *Biochem J* **47**:xli–xlii.
19. Čapek K, Šteffková J, Jarý J (1966) *Collect Czech Chem Commun* **31**:1854–63.
20. Fischer E, Bergmann M, Rabe A (1920) *Ber* **53**:2362–88.