Microbial Biology

Structural characterization of the N-linked pentasaccharide decorating glycoproteins of the halophilic archaeon *Haloferax volcanii*

Lina Kandiba², Chia-Wei Lin³, Markus Aebi³, Jerry Eichler¹,², and Yann Guerardel¹,⁴

²Department of Life Sciences, Ben Gurion University of the Negev, PO Box 653, Beersheva 84105, Israel, ³Institute of Microbiology, ETH Zurich, CH-8093 Zurich, Switzerland, and ⁴Université de Lille, CNRS, UMR 8576-UGSF-Unité de Glycobiologie Structurale et Fonctionnelle, F 59000 Lille, France

To whom correspondence should be addressed: Tel: +972-8646-1343; Fax: +972-8647-9175; e-mail: jeichler@bgu.ac.il (JE); Tel: +33-3-20-43-69-41; Fax: +33-3-20-43-65-55; e-mail: yann.guerardel@univ-lille1.fr (YG)

Received 8 January 2016; Revised 2 February 2016; Accepted 3 February 2016

Abstract

N-Glycosylation is a post-translational modification performed in all three domains of life. In the halophilic archaea *Haloferax volcanii*, glycoproteins such as the S-layer glycoprotein are modified by an N-linked pentasaccharide assembled by a series of Agl (archaeal glycosylation) proteins. In the present study, mass spectrometry (MS) and nuclear magnetic resonance spectroscopy were used to define the structure of this glycan attached to at least four of the seven putative S-layer glycoprotein N-glycosylation sites, namely Asn-13, Asn-83, Asn-274 and Asn-279. Such approaches detected a trisaccharide corresponding to glucuronic acid (GlcA)-β1,4-GlcA-β1,4-glucose-β1-Asn, a tetrasaccharide corresponding to methyl-α1,4-GlcA-β1,4-galacturonic acid-α1,4-GlcA-β1,4-glucose-β1-Asn, and a pentasaccharide corresponding to hexose-1,2-[methyl-α1,4-]GlcA-β1,4-galacturonic acid-α1,4-GlcA-β1,4-glucose-β1-Asn, with previous MS and radiolabeling experiments showing the hexose at the non-reducing end of the pentasaccharide to be mannose. The present analysis thus corrects the earlier assignment of the penultimate sugar as a methyl ester of a hexuronic acid, instead revealing this sugar to be a methylated GlcA. The assignments made here are in good agreement with what was already known of the *Hfx. volcanii* N-glycosylation pathway from previous genetic and biochemical efforts while providing new insight into the process.

Key words: archaea, mass spectrometry, N-linked glycosylation, nuclear magnetic resonance, S-layer glycoprotein

Introduction

It is now clear that the protein content of a given cell exceeds what is encoded by the genome, with post-translational modifications being a major source of such proteomic expansion. Of the various processing events that a protein can undergo, N-glycosylation, namely the covalent attachment of a glycan to selected Asn residues of a target protein, is one of the most prevalent and likely the most complex. Indeed, it has been reported that over 50% of all eukaryotic proteins are N-glycosylated (Apweiler et al. 1999). Long thought to be restricted to eukaryotes, it has since become evident that archaea and bacteria also perform this post-translational modification. However, while N-glycosylation in bacteria is thought to be restricted to the delta/epsilon proteobacteria (Nothaft and Szymanski 2010), such post-translational protein-processing appears to be an almost universal trait in archaea (Kaminski, Lurie-Weinberger, et al. 2013). At the same time, the N-linked glycans decorating archaeal glycoproteins present a diversity not seen elsewhere (Schwarz and Aebi 2011; Eichler 2013), possibly related to the ability of these microorganisms to thrive in extremes of pH, temperature or salinity, as well as in other seemingly harsh environments. Still, our understanding of archaeal N-glycosylation pathways remains limited. In the
last decade, however, significant strides in redressing this situation have been taken in several archaea, including the methanogens Methanococcus voltae and Methanococcus maripaludis, the thermoacidophile Sulfolobus acidocaldarius and the halophile Haloferax volcanii (for review, see Jarrell et al. 2014).

Hfx. volcanii cells, first isolated from the Dead Sea and hence requiring molar concentrations of salt for survival (Mullakhanbhai and Larsen 1973), are surrounded by a protein-based surface (S)-layer comprising a single glycoprotein species, the S-layer glycoprotein (Sumper et al. 1990). As the most abundant protein in Hfx. volcanii and containing seven putative N-glycosylation sites (Asn-13, Asn-83, Asn-274, Asn-279, Asn-370, Asn-498 and Asn-732), such post-translational modification of the S-layer glycoprotein has been the focus of numerous studies. The first such examination employed gas chromatography–mass spectrometry (GC–MS) to determine that Asn-13 and Asn-498 were modified by a linear chain of 10 β1–4-linked glycans, while Asn-274 and Asn-279 were modified by a glycan mostly comprising glycerol, with lesser amounts of galactose and idose (Mengele and Sumper 1992). In later studies, liquid chromatography–electrospray ionization mass spectrometry (LC–ESI-MS) revealed Asn-13 and Asn-83 to be modified by a pentasaccharide comprising a hexose (Hex), two hexuronic acids (HexA), a methyl ester of a HexA and a mannoside (Man) (Abu-Qarn et al. 2007; Guan et al. 2010; Magidovich et al. 2010). The pathway responsible for assembly of this pentasaccharide was delineated using bioinformatics, genetic and biochemical approaches (Eichler et al. 2013). It was subsequently reported that in Hfx. volcanii cells grown at lower salinity (i.e. in medium containing 1.75 rather than 3.4 M NaCl), Asn-498 was modified by a tetrasaccharide comprising a sulfated hexose, two hexoses and a rhamnose (Guan et al. 2012). The same glycan had been previously observed in cells grown in 1.25 M NaCl-containing medium attached to dolichol phosphate (DolP), the lipid carrier upon which N-linked glycans are assembled in Hfx. volcanii (Kuntz et al. 1997; Guan et al. 2010). More recently, studies employing matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) also detected the pentasaccharide bound to Asn-13 and Asn-83 and the tetrasaccharide bound to Asn-498 but also identified a novel glycan, N-acetylgalactosaminyl(1–3)hexose(-sulfoquinovose-hexose)_6, bound to Asn-732 (Parente et al. 2014).

Towards providing a precise description of the glycans N-linked to the Hfx. volcanii S-layer glycoprotein, the present study was undertaken. Using MS and nuclear magnetic resonance (NMR), the composition and structure of a pentasaccharide attached to at least four N-glycosylation sites were defined.

Results

Hfx. volcanii S-layer glycoprotein Asn-13, Asn-83, Asn-274 and Asn-279 are pentasaccharide modified

Analysis of the Hfx. volcanii S-layer glycoprotein amino acid sequence revealed seven sequons containing putative N-glycosylation sites (Asn-13, Asn-83, Asn-274, Asn-279, Asn-370, Asn-498 and Asn-732) (Sumper et al. 1990). Given the different compositions of glycans linked to several of the positions reported (Sumper et al. 1990; Abu-Qarn et al. 2007; Kaminski, Guan, et al. 2013; Parente et al. 2014), N-glycosylation was revisited. Accordingly, the S-layer glycoprotein was digested with Glu-C protease and analyzed by MS. Such digestion was expected to generate peptides comprising 152SFNKTIQGDRVFLGEE72 that includes Asn-13 ([M + H]^+=1926.955), 65NQPLGTYVDGSGATTPNVTLLAPRTIDSE35 that includes Asn-83 ([M + H]^+=3102.522), 270VGIANSATNTSGGSTGPTVE290 that includes Asn-274 and Asn-279 ([M + H]^+=1936.909), 347VINGTATSSDVAIVYRDDDGDWQLE393 that includes Asn-370 ([M + H]^+=2939.390), 492TGTVDI-NGTASSGANSVLHIVFVDE114 that includes Asn-498 ([M + H]^+=2279.140), and 721DTTSSSDNATDTDDTGGPTE253 that includes Asn-732 ([M + H]^+=2131.863) (putative sites of N-glycosylation are in bold). Subsequent nano-LC–MS/MS analysis and manual interpretation of the data revealed the presence of several of these peptides (i.e., those containing Asn-13, Asn-83 and Asn-274 and Asn-279), as well as glycosylated versions thereof.

The nano-LC–MS profile revealed [M + 2H]^2+ peaks consistent with the calculated mass of the Asn-13-containing peptide modified by a monosaccharide (hexose (Hex)), a disaccharide (Hex, HexA), a trisaccharide (Hex, HexA2), a tetrasaccharide (Hex, HexA2, methylated HexA (methylHexA)) and a pentasaccharide (Hex2, HexA2, methyl-HexA) (Table I). In terms of peak intensity, the [M + 2H]^2+ peak at m/z 1316.07, thought to correspond to the tetrasaccharide-modified peptide, predominated (Figure 1A). To confirm the assignment of the peptide and the composition of the bound glycan, two different tandem MS (MS/MS) fragmentation approaches were adopted. Higher-energy collision dissociation (HCD) fragmentation led to the appearance of fragment ions derived from cleavage of the peptide backbone (Figure 1B, top panel). The HCD MS/MS fragmentation pattern revealed consecutive peptide fragment ions from y5 to y14, confirming the peptide to be the Asn-13-containing SFNKTIQGDRVFLGEE proteolytic fragment. However, because the high-energy conditions used in such fragmentation could lead to a loss of information regarding the bound glycan, low-energy collision-induced dissociation (CID) was performed. Based on the sequential neutral loss of monosaccharides observed in the CID MS/MS profile, the Asn-13-bound glycan was confirmed as -Hex-HexA-A-methylHexA (Figure 1B, lower panel). These findings were in agreement with earlier MALDI-TOF MS-based reports defining the glycan bound to Asn-13 of the Hfx. volcanii S-layer glycoprotein (Abu-Qarn et al. 2007).

The nano-LC–MS profile also revealed [M + 3H]^3+ peaks consistent with the calculated mass of the Asn-83-containing peptide (m/z 1063.86), as well as the same peptide modified by the pentasaccharide (Hex-HexA-A-methylHexA-Hex; m/z 1352.60) linked to Asn-13 and its tetrasaccharide precursor (m/z 1298.59) (Table I; Supplementary data, Figure S1A). To confirm the assignment of the peptide and the composition of the bound glycan, MS/MS was performed following HCD fragmentation of the peptide. Such analysis of the non-glycosylated peptide confirmed it to indeed correspond to the Asn-83-containing NQPLGTYVDGSGATTPNVTLLAPRTIDSE proteolytic fragment (Supplementary data, Figure S1B). The same peptide backbone was detected when the tetrasaccharide-modified peptide was examined by HCD MS/MS, which also confirmed that Asn-83 was modified by -Hex-HexA-A-methylHexA (Supplementary data, Figure S1C). As such, these findings were also in agreement with those previously reported (Abu-Qarn et al. 2007).

Digestion of the Hfx. volcanii S-layer glycoprotein Glu-C endopeptidase was predicted to generate an [M + 2H]^2+ peak at m/z 968.958, corresponding to a peptide containing both Asn-274 and Asn-279. Accordingly, the nano-LC MS profile revealed a peak at m/z 968.963, as well as additional peaks consistent with the peptide in question being modified twice by the pentasaccharide Hex-HexA2-methylHexA-Hex and its various precursors (Table I). Examples of the various glycans linked to the VGIANNSATNTSGGSTGPTVE peptide are presented in Figure 2. The [M + 3H]^3+ peak that eluted at 30–35 min corresponded to the peptide modified by two pentasaccharides (m/z 1223.79), by a
pentsaccharide and the trisaccharide precursor (m/z 1169.77), by two
tetrosaccharide precursors (m/z 1115.75), by the tetrascaccharide and
trisaccharide precursors and by two trisaccharide precursors (m/z 989.05) (Figure 2A). The [M + 2H]^{2+} peaks that eluted at 35–40 min
corresponded to the peptide modified by the trisaccharide precursor (m/z 1226.92), by the trisaccharide and monosaccharide precursors (m/z 1307.05), by the trisaccharide and disaccharide precursors (m/z 1395.06) and by the tetrasaccharide and disaccharide precursors (m/z 1490.09) of the pentsaccharide (Figure 2B). Finally, the [M + 2H]^{2+} peaks that eluted at 40–48 min corresponded to the peptide modified by the disaccharide precursor (m/z 1130.00) and by the monosaccharide and disaccharide precursors (m/z 1219.3). These results thus revealed that both Asn-274 and Asn-279 are modified by a pentsaccharide comprising Hex-HexA_{2}-methylHexA-Hex, as well as its precursors. As such analysis did not distinguish between modification of Asn-274 and Asn-279, the order of glycosylation of the two residues is not known, assuming such a preference exists.

Finally, no peptides including Asn-370, Asn-498 or Asn-732 were detected.

**NMR characterization of an S-layer glycoprotein N-linked pentasaccharide**

To more precisely define the composition and structure of the N-linked glycan moiety attached to the *Hfx. volcanii* S-layer glycoprotein, NMR was performed. These experiments relied on the production and purification of S-layer glycoprotein-derived glycopeptides, rather than released free glycans, since no endoglycosidase that can release archaeal N-linked glycans is currently known. Thus, glycopeptides were generated from the S-layer glycoprotein by exhaustive proteolytic digestion with *Streptomyces griseus* pronase. As described in Materials and Methods, the glycopeptides were then purified by multiple chromatographic steps, including anion exchange and gel filtration. Following the purification process, three fractions were analyzed, namely a trisaccharide, a tetrascaccharide and a pentsaccharide.

**Trisaccharide**

{H NMR analysis of the glycopeptides revealed three anomeric signals at δ 4.543 (C), 4.568 (B) and 4.99–5.02 (A) in 1:1:1 ratio, thus establishing the presence of a major oligosaccharide composed of three monosaccharides (Figure 3A). Upon close examination, the A-H1 signal was shown to comprise several sub-signals ranging from 4.989 and 5.023, for which identical spin systems and very similar chemical shifts were identified. This strongly suggested that residue A is a single residue linked to a heterogeneous family of peptides, as would be expected upon non-specific proteolytic digestion by pronase.

The proton spin systems of residues A, B and C were established through COSY, TOCSY (Figure 3B) and HSQC (Figure 3C) and HSQC-ToCSY 1H,13C experiments (Figure 3D and E). All three residues showed similar 1J_{H1-H1}, 1J_{H2-H2}, 1J_{H1-H3} and 1J_{H4-H5} coupling constant patterns (L:L:L:L), which established that residues A, B and C present β-glucos configurations. Chemical shifts of H1 to H5, as well the observation of H6,6′ at 3.922/3.816 ppm, established residue A as a β-glucos residue (Table II). The observation of highly deshielded H5 signals at 3.982 and 3.889 ppm, as well as 2J_{H1-C} correlations between B-H5 and a carbon at 175.3 and between C-H5 and a carbon at 176.0 in a 1H,13C HMBC experiment pointed to residues B and C as being β-glucos residues (GlcA) residues (Table II).

The 1H,13C HSQC spectrum showed A-C1 to be strongly shielded at 80.3 ppm (Figure 3C). This unusual value strongly suggested that A is linked via a C1-N linkage to the nitrogen atom of an Asn residue (Kelly et al. 2009). This was confirmed by the observation of a 1J_{H1-H5} NOE between A-H1 at 5.023 ppm and the H2 of an Asn residue at 2.949 ppm (data not shown). The spin system of the substituted Asn residue was established by 1H,13C COSY and fits with that of previously described N-glycosylated Asn residues (Kelly et al. 2009). Likewise, the strong deshielding of A-C4 to δ 79.5, when compared with 71 ppm for a non-substituted β-glucose (Glc) residue, established that A is substituted at position C4. Residues B and C showed very

**Table I. Summary of glycopeptides from the Hfx. volcanii S-layer glycoprotein treated with Glu-C endopeptidasea**

<table>
<thead>
<tr>
<th>Asn of interestb</th>
<th>Peptide sequence</th>
<th>Observed m/z</th>
<th>Glycopeptide mass [M + H]^+</th>
<th>Glycan structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn-13</td>
<td>SFNKTIQSGDRVFLGE</td>
<td>1045.011²⁺</td>
<td>2089.029</td>
<td>Hex</td>
</tr>
<tr>
<td></td>
<td>1133.022²⁺</td>
<td>2265.061</td>
<td>HexHexA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1221.042²⁺</td>
<td>2441.093</td>
<td>HexHexA₂</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1316.066²⁺</td>
<td>2631.141</td>
<td>HexHexA₁(methylHexA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1397.082²⁺</td>
<td>2793.195</td>
<td>HexHexA₂(methylHexA)Hex</td>
<td></td>
</tr>
<tr>
<td>Asn-83</td>
<td>NQPLGTVDGSGSAT</td>
<td>1063.862²⁺</td>
<td>3189.592</td>
<td>HexHexA₁(methylHexA)</td>
</tr>
<tr>
<td></td>
<td>TPNVTLLAPRTDSE</td>
<td>1298.585²⁺</td>
<td>3893.761</td>
<td>HexHexA₂(methylHexA)Hex</td>
</tr>
<tr>
<td></td>
<td>1352.601²⁺</td>
<td>4055.810</td>
<td>HexHexA₁(methylHexA)Hex</td>
<td></td>
</tr>
<tr>
<td>Asn-247, 249</td>
<td>VGIANSATNTSGSTG</td>
<td>968.963²⁺</td>
<td>1936.934</td>
<td>Hex</td>
</tr>
<tr>
<td></td>
<td>PTVE</td>
<td>1049.987²⁺</td>
<td>2098.974</td>
<td>Hex</td>
</tr>
<tr>
<td></td>
<td>1131.014²⁺</td>
<td>2261.028</td>
<td>Hex₂</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1138.003²⁺</td>
<td>2275.006</td>
<td>HexHexA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1219.030²⁺</td>
<td>2437.060</td>
<td>HexHexA₂</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1226.919²⁺</td>
<td>2452.838</td>
<td>HexHexA₂</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1307.046²⁺</td>
<td>2613.092</td>
<td>HexHexA₂</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1395.063²⁺</td>
<td>2789.126</td>
<td>HexHexA₃</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1402.070²⁺</td>
<td>2803.140</td>
<td>HexHexA₃</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1490.087²⁺</td>
<td>2979.174</td>
<td>HexHexA₄</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1052.405²⁺</td>
<td>3153.215</td>
<td>HexHexA₃(methylHexA)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1115.754²⁺</td>
<td>3345.262</td>
<td>HexHexA₄(methylHexA)A₂</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1169.772²⁺</td>
<td>3507.316</td>
<td>HexHexA₄(methylHexA)A₂</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1223.790²⁺</td>
<td>3669.370</td>
<td>HexHexA₄(methylHexA)A₂</td>
<td></td>
</tr>
</tbody>
</table>

aAll data are supported by mass matching and confirmed by at least one MS/MS spectrum.
bThe Asn residues of interest are indicated in bold.
similar spin systems, except for H4/C4 that was strongly deshielded to δ 3.758/81.8 in residue B, when compared with δ 3.548/72.7 in residue C. This established that residue B is substituted at position C4 whereas C is non-substituted (Figure 3C). It is noteworthy that the B and C spin systems were very similar to those from a GlcA-β1,4-GlcA-β1,4-GlcA-β1,4-GlcA-α1-Asn-peptide. Together, these experiments established the structure of the major trisaccharide bound to the glycopeptides as GlcA-β1,4-GlcA-β1,4-GlcA-β1,4-GlcA-α1-Asn-peptide.

Tetrasaccharide

As with the trisaccharide, NMR analysis of the tetrasaccharide revealed three sets of signals A, B and C, assigned according to their chemical shifts and spin systems as β-Glc, β-GlcA and β-GlcA, respectively (Figure 4A–D). In addition to these three signals, two signals were identified in the anomeric region at 1H/13C 5.023/69.0 (Figure 4C) and 4.891/102.4 (Figure 4D). Determination of the spin systems associated with these two signals established that both originated from a single monosaccharide, residue D. Thus, the 1H signal at 4.891 ppm was identified as D-H1 and the 1H signal at 5.023 was identified as D-H5 (Table III). In addition, an intense O–CH₃ signal was identified at δ 3.485/61.1.

Analysis of the spin systems of residues A, B and C showed that they differed from those observed in the trisaccharide fraction. This was reflected in the C4 chemical shift that changed from 72.7 to 82.9 ppm, indicative of residue C being substituted at the C4 position in the tetrasaccharide, whereas it was non-substituted in the trisaccharide (compare Figures 3C and 4C). The 3JH1-H2, 3JH2-H3 and 3JH3-H4 coupling constant pattern of D was established as S:L:S, confirming that D presents an α-galacto configuration. Furthermore, the HMBC spectrum showed correlation between D-H5 and a carboxylic

![Figure 1. Nano-LC-MS/MS analysis shows that Hfx. volcanii S-layer glycoprotein Asn-13 is modified by di- to pentasaccharides. (A) The MS spectrum shown includes the m/z 1316.07 (z = 2) peak corresponding to the Asn-13-containing SFNKTIQSDRVFLGE peptide (represented by the gray bar, with Asn-13 highlighted) modified by the tetrasaccharide Hex-HexA-HexA-methylHexA. (B) MS/MS spectra of the m/z 1316.07 (z = 2) peak following HCD (upper panel), as obtained upon mass-dependent acquisition by the LTQ-Orbitrap Velos mass spectrometer, or CID (lower panel), as obtained by ion trap fragmentation methods, show successive y ions released from the peptide backbone and the sequential neutral loss of tetrasaccharide-derived monosaccharides, according to the schemes presented in the upper left corner of each panel. In (A) and each panel of (B), the identity of the sugar responsible for the difference in mass between two peaks is indicated.](https://academic.oup.com/glycob/article-abstract/26/7/745/2388913/bib-1)
carbon at 174.4 ppm (data not shown). Thus, the $^1$H and $^{13}$C chemical shifts were consistent with C4-substituted α-D-galacturonic acid (GalA), except for D-H5, which was downfield-shifted >0.3 ppm, in comparison with what appears in the literature (Yang et al. 2013). Such an unusual chemical shift has thus far been observed either in residues of D-galacturonamide (GalAN), where the carboxyl group is replaced by a carbamoyl group (CO–NH$_2$) (Perry and MacLean 1999; Contreras Sánchez-Matamoros et al. 2013) or in residues of D-galacturonic acid in which the carboxylic acid is methyl-esterified (CO–OCH$_3$) (Yang et al. 2013). As no evidence from MS or MS/MS experiments conducted in the past or as part of the present reports supports the presence of a hexuronamide residue in the tetrasaccharide, the possibility that D is GalAN could be discounted. At the same time, the $^1$H/$^{13}$C chemical shift of an O-methyl group was noted at 3.485/61.1 ppm, an observation that would support the identification of a C6-methyl-esterified GalA. However, the –O–CH$_3$ group chemical shift did not fit with that of CO–OCH$_3$ groups, as reported in the literature (~3.6/55) (Yang et al. 2013). To confirm the values reported in the literature, GalA was methyl-esterified upon incubation with MeOH/HCl, as described in Materials and Methods. As reflected in Supplementary data,
The NMR results determined the fourth sugar of the tetrasaccharide to be Met-O-4-GlcA rather than a methyl ester of a HexA, as previously reported (Magidovich et al. 2010). As such, additional MS experimentation was carried out. Here, alkaline hydrolysis was performed to determine which of these two methyl-modified sugars was part of the N-linked tetrasaccharide. Were the fourth sugar a methyl-esterified GlcA, then mild alkaline hydrolysis would result in a 14 mass unit reduction. If, however, the fourth sugar were O-methylated, then such treatment would not affect the mass of the glycan. Accordingly, an S-layer glycoprotein-derived tryptic peptide containing Asn-13 was examined by liquid chromatography–electron spray ionization MS as described previously (Kaminski et al. 2010), in some cases following mild alkaline hydrolysis. Such analysis revealed the presence of [M + 2H]2+ ion peaks at m/z 1143.44, corresponding to the Asn-13-containing peptide modified by the tetrasaccharide, in both the control and mild base hydrolysis-treated samples (not shown). This observation provided further support for reassigning the fourth sugar as Me-O-4-GlcA rather than as a methyl ester of a HexA, as previously claimed.

Finally, NOESY revealed a full set of inter-residue correlations (B-H1/A-H3, B-H1/A-H4, D-H1/B-H3, D-H1/B-H4, C-H1/D-H3 and C-H1/D-H4) that strongly suggested that the tetrasaccharide had the sequence C-D-B-A. In agreement with the strong deshielding of A-C4 (δ 79.4), B-C4 (δ 80.1) and D-C4 (δ 79.3) carbons and the nOe connectivities, the sequence and monosaccharide linkages were confirmed by HMBC, owing to the inter-residue 1H-13C correlations B-H1/A-C4, B-H1/B-C4 and C-H1/D-C4, as C-1-4-D-1-4-B-1-A (data not shown).

Taken together, the homo- and heteronuclear NMR experiments, as well as the MS results, established the structure of the tetrasaccharide bound to the glycopeptides as Me-O-4-GlcAlβ1-4-GlcAlβ1-4Glcβ1-Asn peptide.

Pentasaccharide
In addition to fractions containing the tri- and tetrasaccharides described above, a third fraction containing a mixture of a major and
Hfx. volcanii N-linked pentasaccharide composition

Fig. 4. $^1$H and $^{13}$C NMR structural analysis of the Me-O-4-GlcAβ1-4-GalAα1-4-GlcAβ1-4-Glcβ1-Asn tetrasaccharide associated with pronase-generated glycopeptides. The spin system of the glycan moiety was established by a combination of (A) $^1$H 1D, (B) $^1$H/$^1$H TOCSY, and (C and D) $^1$H/$^{13}$C HSQC. (E and F) The O-CH$_3$ group was unambiguously associated with the C4 position of the terminal GlcA residue by the observation of intense $^3$J$_{4C-C4}$ correlations on the $^1$H/$^{13}$C HMBC spectrum starting from both C-C4 and O-CH$_3$ $^{13}$C values, coupled with the strong deshielding of C-C4 to 82.9 ppm. The asterisk in (F) indicates HBMC rotation bands of the O-CH$_3$ signal.

a minor compound was obtained. The major structure was identified as Me-O-4-GlcAβ1-4-GalAα1-4-GlcAβ1-4-Glcβ1-Asn-peptide, based on its NMR parameters being almost identical to those of the tetrasaccharide identified in the previous fraction. The minor compound was a pentasaccharide that was not, however, detected in the previous fraction. Because of the low quantity of this oligosaccharide, it could not be separated from the major tetrasaccharide by chromatographic methods and as such, its structure was established from NMR-based analysis of the mixture.

The pentasaccharide was first identified as such based on the presence of a supplemental anemeric signal at 5.201 ppm assigned as residue E, the spin system of which was partially identified by NOESY, COSY and TOCSY experiments (Figure 5A and Table IV). Along with residue E, further analysis of resulting spectral data identified a clear set of minor spin systems (A’, B, C’ and D’) that suggested residue E as being linked to the non-reducing terminal residue of the Me-O-4-GlcAβ1-4-GalAα1-4-GlcAβ1-4-Glcβ1-oligosaccharide (Figure 5B–D; Table IV). Indeed, C’ and D’ showed similar, although slightly different parameters as C and D, demonstrating that the addition of residue E onto the tetrasaccharide induced modifications in the chemical shifts of the penultimate (C) and ante-penultimate (D) residues of the tetrasaccharide. However, A’ and B’ could not be distinguished from A and B, showing that modification of the non-reducing end of the oligosaccharide did not affect the NMR parameters of the two residues found at the reducing end of the oligosaccharide. The spin system of E was established by COSY and TOCSY (Figure 5B). The $^3$J$_{2H-4H}$ coupling constant pattern of E was tentatively established as SiLsLs, which would correspond to a β-idose (Ido) configuration. However, the extreme rarity of this monosaccharide made it difficult to compare the experimental data obtained here with available information from known polysaccharides. Nonetheless, the chemical $^1$H/$^{13}$C shifts of E did not at all fit with those published for free β-Ido (Snyder and Serianni 1986), which suggested that E may be another monosaccharide that adopts a non-classical conformation. Furthermore, compositional analysis of this fraction by gas chromatography did not reveal the presence of idose, which strongly suggested that E is not idose but rather another monosaccharide, the nature and conformation of which have yet to be established by complementary physico-chemical means. Owing to their spin systems, C’ and D’ residues were identified as CH$_3$-O-4-GaLa and GlcA residues, respectively, as in the tetrasaccharide. The NOESY spectrum showed clear NOE correlations E-H1/C-H2 and C-H1/D-H4 that strongly suggested residue E as being linked to C’ in the C2 position and that residue C’ was linked to D’ at the C4 position (Figure 5E). The substitution of C2 in the C2 position was in agreement with the downfield shift observed for C-H2/C2 to δ 3.487/79.7, when compared with (δ +0.133/5.7) C-H2/C2 when C was in the terminal non-reducing position. Similarly, substitution of D’ at the C4 position was in agreement with D-H4/C4 values at 4.062/80.6 ppm.

In summary, NMR analysis established the presence of a minor pentasaccharide that resulted from a substitution of the previously identified tetrasaccharide at the C2 position of its terminal monosaccharide. As earlier studies revealed the terminal sugar of the N-linked pentasaccharide to be Man (Guan et al. 2010; Calo et al. 2011; Kaminski et al. 2012), the structure of the pentasaccharide bound to the glycopeptides is proposed to be Man-1,2-[Me-4-GlcA], 4GalAα1,4-GlcAβ1,4-Glcβ1-Asn-peptide. Still, the origin of the unusual NMR parameters of the terminal hexose residue should be further investigated.

Discussion

A series of genetic and biochemical studies have served to delineate the pathway employed by Hfx. volcanii for the assembly and attachment of an N-linked pentasaccharide decorating protein targets in this organism, including the S-layer glycoprotein (Eichler et al. 2013). In the Agl pathway, the first four sugars of the N-linked pentasaccharide are added to a common DolP carrier by the sequential actions of the glycosyltransferases AglJ, AglG, AglI and AglE (Abu-Qarn et al. 2008; Arbib et al. 2013). Once the lipid-linked tetrasaccharide is assembled, it is flipped to face the cell exterior by an unknown mechanism, at which point the glycan is transferred to target Asn residues by the oligosaccharyltransferase AglB. At the same time, the fifth and final pentasaccharide sugar is added to its own DolP carrier by the glycosyltransferase AglD (Abu-Qarn et al. 2007; Guan et al. 2010). The resulting lipid-linked monosaccharide is flipped to face the cell
et al. 2011; Kaminski et al. 2012), as well as those demonstrating Man at the terminal position of the N-linked pentasaccharide (Calo et al. 2010), the structure of the same position as being derived from Man-charged DolP (Guan et al. 2010). In the context of earlier results showing the incorporation of [2-3H]GalA in the trisaccharide and pentasaccharide, the disaccharide core is extended by a 1→4-linked methylated GlcA in α anomery and a 1→4-linked methylated GlcA in β anomery and a 1→2-linked Man. Such structural diversity is reminiscent of what occurs in eukaryotic cells, where Golgi-generated N-linked glycan diversity is routinely seen (e.g., see Knežević et al. 2009). Such diversity in glycan composition could also reflect the ability of Hfx. volcanii to modify its N-glycan profile as a function of environmental salinity (Kaminski, Guan et al. 2013).

The detection of two different sugars at position three of the N-linked glycans could reflect one of two possible biosynthetic processes. In the first scenario, two different glycosyltransferases could add the sugar found at position three of the DolP-linked precursor, with extension of the glycan only occurring when GalA is added. Gene deletion-based studies revealed AglI as responsible for adding a HexA to the DolP-bound disaccharide that serves as precursor of the N-linked pentasaccharide biosynthesis except the gene encoding AglD, responsible for charging DolP with the same precursor as gives rise to the N-linked trisaccharide. As such, it would appear that the Hfx. volcanii Agl pathway generates two different N-linked glycans, namely the trisaccharide and the pentasaccharide (and pentasaccharide) reported here, all sharing the common GlcAβ1,4-Glcβ1 disaccharide core. In the trisaccharide, a 1→4-linked GlcA is added in the β anomery, whereas in the higher-ordered glycans (i.e. the tetrasaccharide and pentasaccharide), the disaccharide core is extended by a 1→4-linked GalA in α anomery, a 1→4-linked methylated GlcA in β anomery and a 1→2-linked Man. Such structural diversity is reminiscent of what occurs in eukaryotic cells, where Golgi-generated N-linked glycan diversity is routinely seen (e.g., see Knežević et al. 2009). Such diversity in glycan composition could also reflect the ability of Hfx. volcanii to modify its N-glycan profile as a function of environmental salinity (Kaminski, Guan et al. 2013).

Table III. 1H, 13C chemical shifts (in p.p.m.) and Jα,γ coupling constants (in Hz) of the tetrasaccharide CH2-O-4-GlcAβ1-4GalAα1-4GlcAβ1-4Glcβ1-Asn (C-D-B-A)

<table>
<thead>
<tr>
<th>Residue</th>
<th>1H</th>
<th>13C</th>
<th>1H</th>
<th>13C</th>
<th>1H</th>
<th>13C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.023–5.010</td>
<td>4.536</td>
<td>4.492</td>
<td>4.891</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J1,2</td>
<td>80.2</td>
<td>103.4</td>
<td>104.6</td>
<td>102.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J2,3</td>
<td>72.7</td>
<td>74.4</td>
<td>74.1</td>
<td>69.84</td>
<td>37.1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3.702</td>
<td>3.616</td>
<td>3.55</td>
<td>4.104</td>
<td>4.091</td>
<td></td>
</tr>
<tr>
<td>J3,4</td>
<td>76.0</td>
<td>75.2</td>
<td>76.0</td>
<td>70.5</td>
<td>52.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.642</td>
<td>3.727</td>
<td>3.296</td>
<td>4.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J4,5</td>
<td>79.4</td>
<td>80.1</td>
<td>82.9</td>
<td>79.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.662</td>
<td>3.938</td>
<td>3.848</td>
<td>5.023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J5,6</td>
<td>77.6</td>
<td>76.2</td>
<td>76.1</td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/6'</td>
<td>8</td>
<td>6</td>
<td>8</td>
<td>&lt;2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-O-CH3</td>
<td>3.811–3.921</td>
<td>174.0</td>
<td>175.5</td>
<td>174.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values are given in ppm for 1H and 13C chemical shifts and in Hz for coupling constants. Values in bold correspond to positions of glycosidic substitution. n.d., not determined.
Hfx. volcanii N-linked pentasaccharide composition

In the second possible biosynthetic scenario, the GlcA at position three of the DolP-linked precursor is replaced by GalA at the level of the DolP-bound trisaccharide. Isomerization of a sugar found as part of a higher-ordered structure has been shown in heparin/heparin sulfate biosynthesis, where GlcA is isomerized to iduronic acid in the polymeric stage (Hagner-McWhirter et al. 2004). In this context, it is of note that in the absence of AglQ, a predicted isomerase, only a trisaccharide was attached to the DolP carrier or N-linked to the S-layer glycoprotein (Arbiv et al. 2013). AglQ may thus be responsible for converting the GlcA found at position three of the DolP-bound trisaccharide into GalA, so that glycan extension can occur. Such isomerization for adding the fourth sugar to the DolP-bound trisaccharide (Abu-Qarn et al. 2008). Indeed, the two genes are found adjacent to each other in the Hfx. volcanii genome (Yurist-Doutcht and Eichler 2009). It is also possible that DolP-linked GalA-GlcA-Glc corresponds to a poor substrate of the presently unknown Hfx. volcanii flipase that translocates lipid-linked precursors across the membrane, or of AglB, the archaeal oligosaccharyltransferase (Abu-Qarn et al. 2007).

In this study, the identities of the sugars comprising the pentasaccharide linked to at least four of the seven putative N-glycosylation sites of the Hfx. volcanii S-layer glycoprotein were discerned. Of these, Asn-13 and Asn-83 were previously known to bear this glycan and its precursors (Abu-Qarn et al. 2007), whereas Asn-274 and Asn-279 were only now shown to be modified by the same glycan (s). Having previously delineated the pathway responsible for processing the component sugars of the pentasaccharide into a DolP-linked and, ultimately, an N-linked glycan (Abu-Qarn et al. 2007; Yurist-Doutcht et al. 2008; Kaminski et al. 2010; Magidovich et al. 2010; Cohen-Rosenzweig et al. 2012; Kaminski et al. 2012; Arbiv et al. 2013), it will now be possible to obtain detailed information on the workings of the various pathway enzymes. Such efforts will not only advance our understanding of N-glycosylation in archaea but will also offer new insight into this universal protein-processing event across evolutionary lines.

**Materials and methods**

**S-layer glycoprotein isolation**

*Halofex volcanii* cells were grown in complete medium containing 3.4 M NaCl, 0.15 M MgSO₄, 7H₂O, 1 mM MnCl₂, 4 mM KCl, 3 mM CaCl₂, 0.3% (w/v) yeast extract, 0.5% (w/v) tryptone, 50 mM Tris–HCl, pH 7.2, at 37°C (Mevarech and Werczberger 1985). The S-layer glycoprotein was isolated as previously described (Sumper et al. 1990). Briefly, 6 L of cells were grown to stationary phase (OD₆₀₀ = 3.0) and harvested by centrifugation (7000 × g, 30 min). Cell pellets (from 800 mL aliquots of culture) were resuspended in 200 mL of complete medium lacking yeast extract and tryptone, to which 60 mL of 0.5 M EDTA (pH 6.8) was added. The cells were shaken for 30 min at 37°C and the resulting spheroplasts were removed by three centrifugations (3000 × g, 15 min; 7000 × g, 5 min; 13,000 × g, 10 min). The supernatant containing the S-layer glycoprotein was dialyzed against 50 mM Tris–HCl, pH 7.2, and concentrated by success and ultrafiltration to the final volume of 2 mL. The resulting S-layer glycoprotein-containing concentrate was lyophilized and stored at −20°C until used.

**MS sample preparation**

Isolated *Hfx. volcanii* S-layer glycoprotein was prepared for MS according to the Filter Assistant Sample Preparation method, as previously described (Wiśniewski et al. 2009). Briefly, the protein was dissolved in 400 μL of 0.1 M Tris–HCl, pH 8.3 containing 8 M urea and loaded into a Microcon YM-30 filtration device (Millipore), which was centrifuged twice (10,000 × g, 20°C, 10 min). All subsequent centrifugation steps employed the same conditions to allow for maximal concentration. The urea concentrate was diluted with 400 μL of 1 M urea in 0.1 M Tris–HCl, pH 8.3, concentrated (three times) and diluted with 400 μL of 0.1 M Tris–HCl, pH 8.3 containing 0.1 M urea and concentrated (three times). The final concentrated protein solution was subjected to endopeptidase Glu-C (Roche) digestion (37°C, overnight). The digested samples were collected by centrifugation, and the filter

---

**Fig. 5.** ¹H and ¹³C NMR structural analysis of the minor pentasaccharide in mixture with Me-4-GlcAβ1-4-GlcAβ1-4-GlcAβ1-3m ACCl₂, 0.3% (w/v) yeast extract, 0.5% (w/v) tryptone, 50 mM Tris–HCl, pH 7.2, at 37°C (Mevarech and Werczberger 1985). The S-layer glycoprotein was isolated as previously described (Sumper et al. 1990). Briefly, 6 L of cells were grown to stationary phase (OD₆₀₀ = 3.0) and harvested by centrifugation (7000 × g, 30 min). Cell pellets (from 800 mL aliquots of culture) were resuspended in 200 mL of complete medium lacking yeast extract and tryptone, to which 60 mL of 0.5 M EDTA (pH 6.8) was added. The cells were shaken for 30 min at 37°C and the resulting spheroplasts were removed by three centrifugations (3000 × g, 15 min; 7000 × g, 5 min; 13,000 × g, 10 min). The supernatant containing the S-layer glycoprotein was dialyzed against 50 mM Tris–HCl, pH 7.2, and concentrated by success and ultrafiltration to the final volume of 2 mL. The resulting S-layer glycoprotein-containing concentrate was lyophilized and stored at −20°C until used.

**Materials and methods**

**S-layer glycoprotein isolation**

*Halofex volcanii* cells were grown in complete medium containing 3.4 M NaCl, 0.15 M MgSO₄, 7H₂O, 1 mM MnCl₂, 4 mM KCl, 3 mM CaCl₂, 0.3% (w/v) yeast extract, 0.5% (w/v) tryptone, 50 mM Tris–HCl, pH 7.2, at 37°C (Mevarech and Werczberger 1985). The S-layer glycoprotein was isolated as previously described (Sumper et al. 1990). Briefly, 6 L of cells were grown to stationary phase (OD₆₀₀ = 3.0) and harvested by centrifugation (7000 × g, 30 min). Cell pellets (from 800 mL aliquots of culture) were resuspended in 200 mL of complete medium lacking yeast extract and tryptone, to which 60 mL of 0.5 M EDTA (pH 6.8) was added. The cells were shaken for 30 min at 37°C and the resulting spheroplasts were removed by three centrifugations (3000 × g, 15 min; 7000 × g, 5 min; 13,000 × g, 10 min). The supernatant containing the S-layer glycoprotein was dialyzed against 50 mM Tris–HCl, pH 7.2, and concentrated by success and ultrafiltration to the final volume of 2 mL. The resulting S-layer glycoprotein-containing concentrate was lyophilized and stored at −20°C until used.

**MS sample preparation**

Isolated *Hfx. volcanii* S-layer glycoprotein was prepared for MS according to the Filter Assistant Sample Preparation method, as previously described (Wiśniewski et al. 2009). Briefly, the protein was dissolved in 400 μL of 0.1 M Tris–HCl, pH 8.3 containing 8 M urea and loaded into a Microcon YM-30 filtration device (Millipore), which was centrifuged twice (10,000 × g, 20°C, 10 min). All subsequent centrifugation steps employed the same conditions to allow for maximal concentration. The urea concentrate was diluted with 400 μL of 1 M urea in 0.1 M Tris–HCl, pH 8.3, concentrated (three times) and diluted with 400 μL of 0.1 M Tris–HCl, pH 8.3 containing 0.1 M urea and concentrated (three times). The final concentrated protein solution was subjected to endopeptidase Glu-C (Roche) digestion (37°C, overnight). The digested samples were collected by centrifugation, and the filter
device was rinsed with 10% acetonitrile and again centrifuged. The combined filtrates were desalted by passage through a C18 Zip-tip (Millipore) and dried using a SpeedVac evaporator. The dried peptide mixture was resolved in 20 μL of 2% acetonitrile containing 0.1% formic acid. A 4 μL aliquot was used for MS analysis.

MS analysis
Samples were analyzed on an LTQ-Orbitrap Velos mass spectrometer (Thermo Fischer Scientific, Bremen, Germany) coupled to an Eksigent-Nano-HPLC system (Eksigent Technologies, Dublin, CA). Solvent composition at the two channels was 0.1% formic acid in deionized water (channel A) and 0.1% formic acid in acetonitrile (channel B). Peptides were loaded onto a home-made tip column (75 μm x 80 mm) packed with reverse phase C18 resin (AQ, 3 μm, 200 Å; Bischoff, Leonberg, Germany) and eluted with a 30–30% gradient of channel B solvent over a 50 min period, 30% gradient of channel B solvent for 14 min, at a flow rate of 250 nL per min.

For glycopeptide analysis, full-scan MS spectra (800–2000 m/z) and MS/MS spectra from HCD were acquired with resolutions of 60,000 and 15,000 at 400 m/z after accumulation to a target value of 1,000,000 and 100,000, respectively. Both CID and HCD MS/MS spectra of the same precursor ion were recorded in the data-dependent mode of the ion trap and Orbitrap in parallel from the four most intense signals above a threshold of 2000 using a normalized collision energy of 35% and an activation time of 30 ms. Charge state screening was enabled and singly charge states were rejected. Precursor masses already selected for MS/MS were excluded for further selection for 60 s and the exclusion window was set to 20 ppm. The size of the exclusion list was set to a maximum of 500 entries.

Mild alkaline hydrolysis
The Hfx. volcanii S-layer glycoprotein was prepared for analysis by LC–ESI MS as described (Kaminski et al. 2010); however, following trypsin treatment, the dried glycopeptide-containing material was incubated with 0.1 M NaOH for 2 h at 37°C, neutralized using diluted (1:200) trifluoroacetic acid and passed through a C18 Zip-tip. In control experiments, water was added place of 0.1 M NaOH. LC–ESI MS was then performed as described (Kaminski et al. 2010).

NMR sample preparation
The Hfx. volcanii S-layer glycoprotein was digested by S. griseus pronase (Sigma-Aldrich) in 10 mM calcium acetate, pH 8, at 37°C for 24 h. The reaction was terminated by addition of acetic acid until pH 4 was reached. Nine volumes of cold ethanol were added and the solution was incubated overnight at 4°C. The solution was centrifuged at 5,000 x g for 30 min at 4°C. After removal of the supernatant, the pellet was dried, digested by S. griseus pronase a second time and processed as before. The two supernatants were pooled and concentrated by rotary evaporator before desalting on a Biogel P6 (Bio-Rad) filtration column in water. All fractions were screened for the presence of neutral sugars by a colorimetric method (Dubois et al. 1956). Carbohydrate-containing fractions were pooled and subjected to anion exchange chromatography on a DEAE fast flow (GE Healthcare Life Sciences) column equilibrated in water and eluted by a linear NaCl gradient (0–0.6 M). Fractions were assayed for total sugars by 1H NMR. Fractions corresponding to glycopeptides were pooled, concentrated and further purified by gel filtration on a Toyopearl HW40 (Tosoh Bioscience) column and eluted with water.

NMR experiments
After two exchanges with D2O (Eurisotop-Saclay France), samples were dissolved in pure D2O. The pD was kept neutral. Experiments were recorded at 300 K on Bruker spectrometers at 18.8 and 21.6 T and were 1H-resonated at 800.12 and 900.11 MHz and 13C-resonated at 200.3 and 220.0 MHz, respectively. The pulse programs used were extracted from the Bruker pulse program library, with pulses and delays being

Table IV. 1H-13C chemical shifts (in p.p.m.) and J_{H,C} coupling constants (in Hz) of the pentasaccharide Hex1-2[CH2-O-4-] GlcA[1-4GalAα1-4GlcA[1-4Galβ1-3GlcNac-Asn (E-C-D’-B-A)

<table>
<thead>
<tr>
<th>Residue A</th>
<th>1H</th>
<th>13C</th>
<th>Residue B</th>
<th>1H</th>
<th>13C</th>
<th>Residue C</th>
<th>1H</th>
<th>13C</th>
<th>Residue D’</th>
<th>1H</th>
<th>13C</th>
<th>Residue E</th>
<th>1H</th>
<th>13C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.021–5.010</td>
<td>4.546</td>
<td>4.665</td>
<td>4.861</td>
<td>5.021</td>
<td>2</td>
<td>103.4</td>
<td>105.4</td>
<td>102.2</td>
<td>n.d.</td>
<td>J_{H,C}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.649</td>
<td>3.727</td>
<td>3.369</td>
<td>4.062</td>
<td>3.868</td>
<td>n.d.</td>
<td>82.3</td>
<td>80.6</td>
<td>75.0</td>
<td>75.0</td>
<td>75.0</td>
<td>80.6</td>
<td>75.0</td>
<td>75.0</td>
</tr>
<tr>
<td>5</td>
<td>3.652</td>
<td>3.938</td>
<td>3.925</td>
<td>4.957</td>
<td>4.957</td>
<td>n.d.</td>
<td>69.3</td>
<td>69.3</td>
<td>69.3</td>
<td>69.3</td>
<td>69.3</td>
<td>n.d.</td>
<td>69.3</td>
<td>69.3</td>
</tr>
<tr>
<td>4-O-CH3</td>
<td>3.499</td>
<td>61.1</td>
<td>3.499</td>
<td>61.1</td>
<td>61.1</td>
<td>n.d.</td>
<td>61.1</td>
<td>61.1</td>
<td>61.1</td>
<td>61.1</td>
<td>61.1</td>
<td>n.d.</td>
<td>61.1</td>
<td>61.1</td>
</tr>
</tbody>
</table>

Values are given in ppm for 1H and 13C chemical shifts and in Hz for coupling constants. Values in bold correspond to positions of glycosidic substitution. n.d., not determined.
optimized for each experiment. For TOCSY experiments, mixing times of 40, 60, 80 and 100 ms were used for spectra recordings.

To differentiate methyl-esterified HexA residues from O-methylated HexA residues, 1 mg of commercial GalA was incubated in 0.5 M MeOH/HCl for 1 h at 80°C. The methanol was evaporated under a stream of nitrogen and the sample was exchanged twice in D2O before analysis by NMR.

Supplementary data
Supplementary data for this article are available online at http://glycob.oxfordjournals.org/.

Funding
This work was supported by a Negev-Zin Associates Fellowship to L.K., by ETH Zürich to C.-W.L., by the Swiss National Science Foundation (Grant 310030_162636) to M.A., by the Israel Science Foundation (Grant 8/11) to J.E. and by Infrastructures de Recherche Foundation (Grant 14:743–753).

Acknowledgements
We thank the Functional Genomics Center Zurich for help with mass spectrometry. The 900 MHz NMR spectrometer was funded by Région Nord-Pas de Calais, European Union (FEDER), Ministère Français de la Recherche, Université Lille 1 and CNRS.

Conflict of interest statement
None declared.

Abbreviations
Agl, archaeal glycosylation; CID, collision-induced dissociation; DolP, dolichyl phosphate; GalAN, galacturonamide; GalA, galacturonic acid; Glc, glucose; Hex, hexose; HexA, hexuronic acid; HCD, higher-energy collisional dissociation; Ido, idose; LC–ESI MS, liquid chromatography–electrospray ionization mass spectrometry; Man, mannose; NMR, nuclear magnetic resonance; S-layer, surface layer; MS/MS, tandem MS.

References


Kaminski L, Guan Z, Yurist-Doutsch S, Eichler J. 2013. Two distinct N-glycosylation pathways process the Haloferax volcanii S-layer glycoprotein upon changes in environmental salinity. mBio. 4:e00716–e00713.


Hfx. volcanii N-linked pentasaccharide composition


