Drastic Differences in Glycosylation of Related S-layer Glycoproteins from Moderate and Extreme Halophiles[*]

Reiner Mengele and Manfred Sumper[‡]

From the Lehrstuhl Biochemie I, Universität Regensburg, Universitätsstrasse 31, 8400 Regensburg, Federal Republic of Germany

The outer surface of the moderate halophilic archaebacterium *Haloferax volcanii* (formerly named *Halobacterium volcanii*) is covered with a hexagonally packed surface (S) layer glycoprotein. The polypeptide (794 amino acid residues) contains 7 N-glycosylation sites. Four of these sites were isolated as glycopeptides and the structure of one of the corresponding saccharides was determined. Oligosaccharides consisting of β-1,4-linked glucose residues are attached to the protein via the linkage unit asparaginyl-glucose. In the related glycoprotein from the extreme halophile *Halobacterium halobium*, the glucose residues are replaced by sulfated glucuronic acid residues, causing a drastic increase in surface charge density. This is discussed in terms of a recent model explaining the stability of halophilic proteins.

The first procaryotic glycoprotein was discovered in halobacteria (1) and found to be the main constituent of the cell surface layer (S-layer). Structural work following this discovery lead to the elucidation of the complete chemical structure of the glycoprotein from *Halobacterium halobium* (for a review, see Refs. 2 and 3). This information makes the halobacterial S-layer an attractive object for the elucidation of its three-dimensional structure by electron microscopy and the techniques of image processing. However, data collection from the extreme halophiles like *H. halobium* was hampered by the fact that the S-layers require high salt conditions to maintain integrity. Fortunately, this limitation does not exist with the S-layer of the moderate halophile *Haloferax volcanii* (4). For this reason, we have started to analyze the structure of the corresponding cell surface glycoprotein. After cloning and sequencing the gene (5), the determination of O-glycosylation sites (5), we now report the structure of the saccharides attached to the polypeptide by N-glycosidic linkages. Although the polypeptides from *H. halobium* and *H. volcanii* exhibit sequence homology, their N-glycosidically linked saccharides differ significantly.

MATERIALS AND METHODS

Bacterial Strain—*H. volcanii* was obtained from M. Mevarech (Israel) and cultivated as described previously (6).

Isolation of the S-layer Glycoprotein—The S-layer glycoprotein was prepared as previously described (5); however, the sodium dodecyl sulfate-polyacrylamide gel electrophoresis was replaced by a HPLC[1] purification step. The concentrated crude preparation (after ultrafiltration) was adjusted to 50 mM Tris-HCl, pH 8, 1 mM MgCl$_2$ and incubated with 50 units of benzonase (Merck) for 4 h at 37 °C. After dialysis against water, the solution was brought to 10 mM K$_2$HPO$_4$, filtrated, and applied to a PLRP-S HPLC column (8 µm; 7.5 × 150 mm; 1000 Å Polymer Laboratories). The conditions of the chromatography were as follows: solvent A, 10 mM K$_2$HPO$_4$; solvent B, acetonitrile. A gradient from 40 to 80% B was applied (20 min; 2 ml/min flow rate). The glycoprotein elutes in homogeneous form at 50% acetonitrile.

Preparation of Glycopeptides—5 mg of glycoprotein was dissolved in 2 ml of 0.1 M Tris-HCl, pH 7.5, and 5 mM CaCl$_2$ and digested with 100 µg of L-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin at 37 °C. After 4 h, the same amount of trypsin was added and incubation continued for 14 h. Peptides were fractionated on a preparative reversed phase column (Lichrosorb 10 µm; 10 × 200 mm; Merck). A linear gradient from 5 to 65% acetonitrile in 0.1% trifluoroacetic acid was applied (90 min; flow rate 4 ml/min). Fractions of 6 ml were collected, freeze-dried, and screened for carbohydrates with the orcinol spot test (7). Final purification to homogeneity was achieved on a RP18 column (Lichrosorb 5 µm; 4 × 200 mm; Merck). Optimized conditions for the different glycopeptides were as follows. Solvent A, 0.1% trifluoroacetic acid; solvent B, 70% acetonitrile, 0.1% trifluoroacetic acid. GP12: 15–18% acetonitrile in 5 min, 18–19% in 25 min, and then 10 min at 19%; GP30: isocratic at 30% acetonitrile; GP45, 30–35% acetonitrile in 30 min, 35–40% in 10 min. Flow rates were 1 ml/min. Control of purity was done on a narrow bore column (Vydac 218 TP; 5 µm; 200 × 2.1 mm) with a linear gradient of 10–37% acetonitrile in 40 min (200 µl/min). Peptides were detected at 206 nm.

Digestion of GP30—30 µg of GP30 in 50 µl of 50 mM Tris-HCl, pH 7.5, 1 mM CaCl$_2$ were digested with 5 µg of proteinase K for 16 h at 37 °C. The expected mixture of two glycopeptides was separated by high performance anion exchange chromatography coupled to electrochemical detection by pulsed amperometry. The Dionex-BioLC-system fitted with a CarboPac PA1 column (4 × 250 mm) and a guard column PA (3 × 25 mm) was operated at a flow rate of 1 ml/min and a sensitivity of 3000 nA. The following potentials and pulse times were used: $E1 = 0.05$ V (120 ms); $E2 = 0.65$ V (60 ms); $E3 = -0.95$ V (180 ms). Eluant A was 100 mM NaOH, eluant B was 100 mM NaOH and 500 mM sodium acetate. Chromatographic conditions were 0% B for 10 min, then 0–100% B in 20 min, and finally 100% B for 40 min.

Permethylation—Permethylation of the glycopeptide was done according to Hakomori (8) with the modifications described by Waeghe *et al.* (9). Sodium dimethylsulfinyl carbanion was prepared as described by Harris *et al.* (10). The methylated product was purified with Sep-Pak C$_{18}$ cartridges (Waters Millipore Inc.) by elution with 70% acetonitrile. After hydrolysis in 4 M trifluoroacetic acid for 2 h at 100 °C, the methylated sugars were reduced with NaBH$_4$, peracetylated, and analyzed by GC-MS.

GC-MS—All GC-MS analyzes were performed with a Hewlett Packard model 5995 GC equipped with a Durabond 1701 fused silica capillary column (30 m, J & W Scientific). The carrier gas (1 ml/min) was high purity helium. The temperature gradient for methylated alditol acetates was 2 °C/min, starting at 140 °C.

Assignment of signals and mass spectra was accomplished by using standards and reference spectra (11).

Digestions with α- and β-Glucosidase—50 mM sodium acetate, pH

[1] The abbreviations used are: HPLC, high performance liquid chromatography; GC-MS, gas chromatography-mass spectrometry.
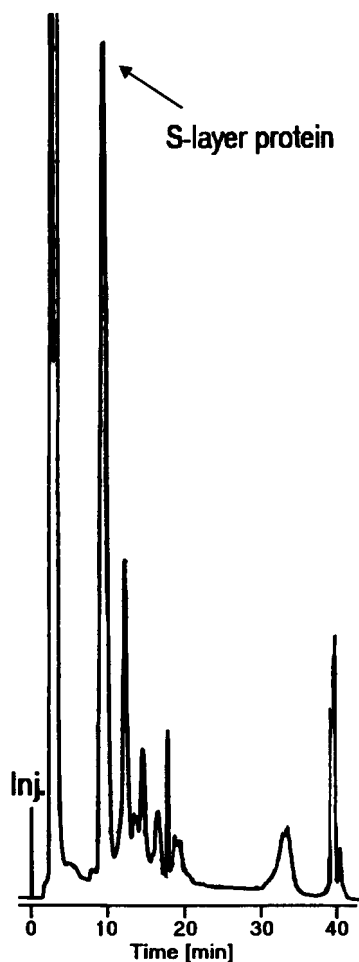
FIG. 1. **Purification of S-layer glycoprotein by reversed phase HPLC under alkaline conditions.** Details are given under "Materials and Methods." Absorbance of the eluant was monitored at 216 nm.

6.6, was used with α-glucosidase, the same buffer at pH 5.0 with β-glucosidase. 0.1 unit of enzyme in 10 μl of buffer were applied. 5-μl aliquots were spotted on silica plates which were developed in 85% acetonitrile for the chromatography of monosaccharides.

RESULTS

Isolation of S-layer Protein—Addition of EDTA to the growth medium converts normal cupped disk-shaped cells of *H. volcanii* to a spherical form. This transition causes massive shedding of the S-layer glycoprotein, which can be recovered from the growth medium in a highly enriched state. Previously, final purification was achieved by preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis (5). For large scale purification, this procedure is tedious and therefore was replaced by reversed phase HPLC on a column designed for applications at alkaline pH. With the eluant of choice, acetonitrile/water buffered with 10 mM K_2HPO_4, this allowed a one-step purification of the S-layer glycoprotein to homogeneity (Fig. 1).

Carbohydrate Composition of the S-layer Glycoprotein—Total sugar composition was analyzed by GC-MS after methanolysis and derivatization of the resulting methyl glycosides with pentafluoropropionic acid anhydride. We could not detect either amino sugars or uronic acids. Neutral sugar analysis (as alditol acetates) revealed that mannose, galactose, glucose, and idose are present in the ratio 1/3/3/0.2.
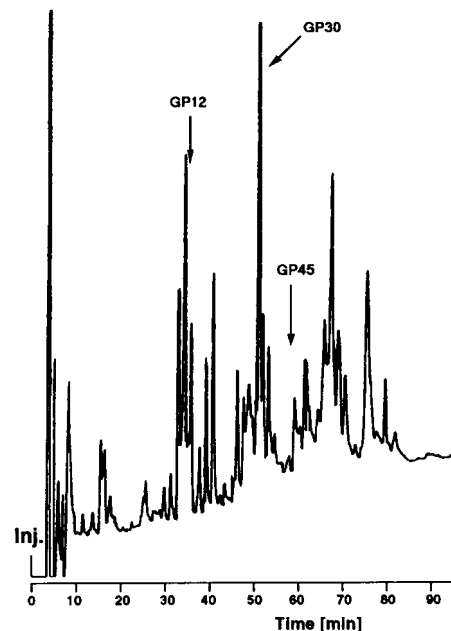
Isolation of Glycopeptides—The polypeptide chain of the

FIG. 2. **Preparative separation of tryptic peptides from the S-layer glycoprotein by HPLC.** Chromatography was performed on a reversed phase C18 column using an acetonitrile gradient in 0.1% trifluoroacetic acid. Details are given under "Materials and Methods." Absorbance was monitored at 206 nm. The positions of carbohydrate-positive material are marked with an *arrow*.

mature glycoprotein contains 794 amino acid residues and exhibits a total of seven potential *N*-glycosylation sites (5). In order to establish the saccharide structure at individual sites, the glycoprotein was digested with trypsin and the resulting peptide mixture fractionated on a preparative reversed phase HPLC column. (Fig. 2). All fractions obtained were screened for the presence of carbohydrate by the orcinol spot test. Carbohydrate-positive material was re-chromatographed on a reversed phase HPLC column, this time on an analytical scale. Three different glycopeptides, named GP12, GP30, and GP45 were purified to homogeneity (Fig. 3). Automated Edman degradation allowed the identification of the following amino acid sequences.

GP12: GNLDADSESF-K
GP30: NIGDTSEVGIA-SSAT-TSGSSTGPTVETA
GP45: VTDQALTGQFTTINGQVAPVETGTVDI-
      GTAASGANSVLVIFVDER

Since GP30 still includes two glycosylation sites, this peptide was further digested with proteinase K, producing two glyco-peptides, which were completely fractionated by HPLC on an anion exchange column under strongly alkaline conditions (Fig. 3). The glycopeptides (named GP-A and GP-B) had the amino acid sequences -TS and -SS, representing both the *N*-glycosylation sites of the original GP30.

Saccharide Structure—The carbohydrate composition of the purified glycopeptides was determined by capillary GC of the peracetylated alditols (Table I). Unexpectedly, glucose turned out to be the main constituent. The saccharides attached to GP12 and GP45 exclusively consist of glucose residues, whereas the saccharides derived from GP30 contain in addition galactose and idose. GP12 was selected for a complete determination of its saccharide structure by permethylation analysis. Unmodified GP12 was insoluble in dimethyl sulfoxide and consequently did not yield the methylated derivative. Therefore, GP12 was digested with chymotrypsin which produced a saccharide linked to the dipeptide
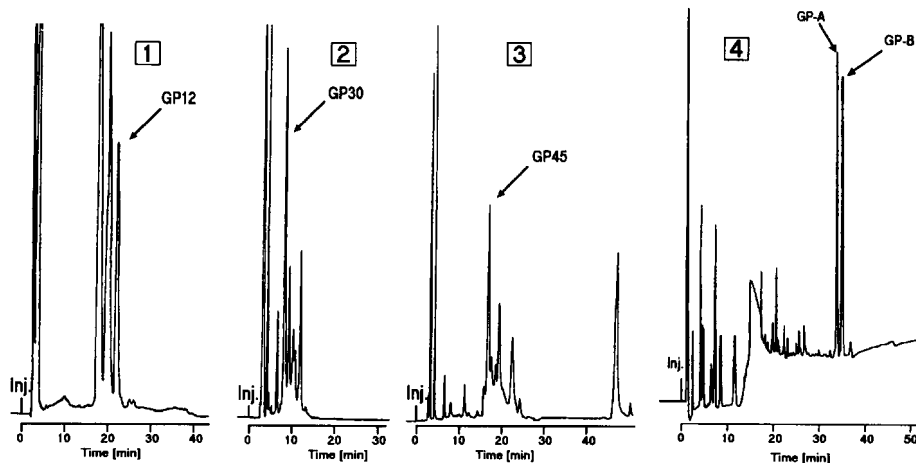
FIG. 3. **Purification of glycopeptides by HPLC.** *1–3*, the carbohydrate positive fractions obtained by preparative HPLC from tryptic digests of the S-layer glycoprotein (Fig. 2) were re-chromatographed on an analytical reversed phase column. Absorbance was monitored at 206 nm. Details are given under "Materials and Methods." *4*, glycopeptide GP30 was digested with proteinase K and applied to anion exchange HPLC column under strongly alkaline conditions. The eluant was monitored by pulsed amperometry. Details are given under "Materials and Methods."

TABLE I

Sugar compositions (mol %) of isolated glycopeptides

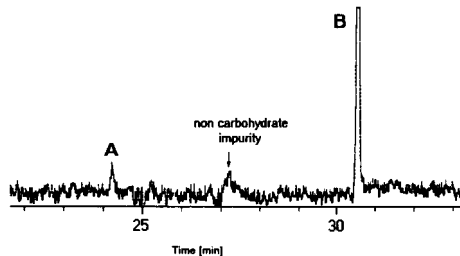|	Glucose	Idose	Galactose
GP12	100		
GP-A	90	10	
GP-B	67	16	17
GP45	100		

FIG. 4. **GC-MS (total ion profile) of the partially methylated alditol acetates derived from the permethylated saccharide attached to the glycopeptide GP12.** *A*, 1,5-di-*O*-acetyl-2,3,4,6-tetra-*O*-methylglucitol; *B*, 1,4,5-tri-*O*-acetyl-2,3,6-tri-*O*-methylglucitol. Details are given under "Materials and Methods."

TABLE II

Comparison of the structures of N-glycosidically linked saccharides from the glycoproteins of H. volcanii and H. halobium

H. volcanii	Glc-[β1 →4Glc]₈-β1→4Glc-Asn
H. halobium	GlcA-[β1→4GlcA]₂-β1→4Glc-Asn
	| |
	OSO₃⁻ OSO₃⁻

Asn-Lys. Dansylation of this derivative greatly enhanced its solubility and allowed complete methylation. After acid hydrolysis, reduction, and peracetylation, the resulting mixture of partially methylated alditol acetates was analyzed by GC-MS (Fig. 4). Two products were obtained and identified by their mass spectra: 1,5-di-*O*-acetyl-2,3,4,6-tetra-*O*-methylglucitol and 1,4,5-tri-*O*-acetyl-2,3,6-tri-*O*-methylglucitol in a ratio of approximately 1:9. This result proves the existence of a linear chain of 1–4-linked glucose residues with a chain length of 10. The saccharide is completely digested by β-glucosidase, indicating the presence of β-1,4-linkages.

Structure of the N-Glycosidic Linkage Unit—In the cell surface glycoprotein of *H. halobium*, hexuronic acid-containing oligosaccharides (12) are attached to the protein via the novel *N*-glycosidic linkage unit asparaginyl-glucose (13). With the above structural data for the glycopeptide GP12, it is established that the same linkage unit is present in the S-layer glycoprotein of *H. volcanii*. In fact, asparaginyl-glucose appears to be the linkage unit of all the remaining *N*-glycosylation sites of this glycoprotein. Treatment of a glycoprotein with anhydrous hydrogen fluoride cleaves all *O*-glycosidic bonds but leaves intact the *N*-glycosidic bond in the linkage unit (14). Only glucose was found to be attached to the S-layer glycoprotein of *H. volcanii* after this treatment confirming the existence of asparaginyl-glucose linkage units.

DISCUSSION

In this study we report the structure of the saccharides linked *N*-glycosidically to the cell surface glycoprotein from *H. volcanii*. Together with our previous work (5, 15), the complete primary structures of the S-layer glycoproteins from a moderate (*H. volcanii*) and an extreme halophile (*H. halobium*) are now available for further studies.

The S-layer glycoproteins are in immediate contact with the ionic milieu of the growth medium. Structural adaptations to high salt conditions should therefore be more evident than in the case of intracellularly located proteins with unknown micro-environments. The S-layer glycoproteins from the moderate and the extreme halophilic archaebacterium are closely related: a comparison of the entire two amino acid sequences shows 40.5% identity. Moreover, the elucidation of the corresponding three-dimensional structures from electron micrographs revealed no structural differences at least to the 2.5-nm level attained so far (4). The stabilization of halophilic proteins in solvents containing high salt concentrations has been discussed in terms of apparent peculiarities in their amino acid composition, *i.e.* higher proportions of acidic amino acid residues (16). A comparison of both S-layer glycoproteins indeed reveals a slight increase in the number of acidic residues (156 for the moderate halophile and 184 for the extreme halophile). However, it is the pattern and the chemical nature of the *N*-linked saccharides which exhibits a drastic change at the transition from moderate to extreme

halophily. Only neutral oligosaccharides (mainly β-1, 4-linked glucoses) are synthesized in the moderate halophile. These glucose residues are replaced by glucuronic acid residues in the extreme halophile. In addition, each glucuronic acid is sulfated at its 3-position, introducing another negative charge per monosaccharide unit (Table II). The S-layer glycoprotein of *H. volcanii* contains seven *N*-glycosylation sites. This number is increased to 12 sites in *H. halobium*. The latter organism uses one of these sites (amino acid position 2) to attach even a novel type of saccharide, which is a repeating unit saccharide introducing again four negative charges per repeat (17). This different pattern of glycosylation introduces at least 120 additional negative charges into the glycoprotein.

From a recent study of halophilic malate dehydrogenase a model was proposed for the stabilization of halophilic proteins (18). It was assumed that the halophilic protein has a core structure similar to the non-halophilic protein, but has in addition protruding highly negatively charged loops required for stabilization in high salt concentrations. The loops are assumed to organize an energetically favored protein-water-salt hydration network. Highly negatively charged saccharides exposed at the protein surface are likely to fulfill a similar function and may allow even higher charge densities than polypeptide chains.

REFERENCES

1. Mescher, M. F., and Strominger, J. L. (1976) *J. Biol. Chem.* **251**, 2005-2014
2. Sumper, M. (1987) *Biochim. Biophys. Acta* **906**, 69-79
3. Lechner, J., and Wieland, F. (1989) *Annu. Rev. Biochem.* **58**, 173-194
4. Kessel, M., Wildhaber, I., Cohen, S., and Baumeister, W. (1988) *EMBO J.* **7**, 1549-1554
5. Sumper, M., Berg, E., Mengele, R., and Strobel, I. (1990) *J. Bacteriol.* **172**, 7111-7118
6. Mevarech, M., and Werczberger, R. (1985) *J. Bacteriol.* **162**, 461-462
7. Vaskovsky, V. E., Kostetsky, E. J., Svetashe, V. I., Zhukova, I. G., and Smirnova, I. G. (1970) *Comp. Biochem. Physiol.* **34**, 163
8. Hakomori, S. (1964) *J. Biochem.* (*Tokyo*) **55**, 205-208
9. Waeghe, T. J., Darvill, A. G., McNeil, M., and Albersheim, P. (1983) *Carbohydr. Res.* **123**, 281-304
10. Harris, P. J., Henry, R. J., Blakeney, A. B., and Stone, B. A. (1984) *Carbohydr. Res.* **127**, 59-72
11. Jannson, E., Kenne, M., Liedgren, M., Lindberg, B., and Loengren, J. (1976) *Chem. Commun.* **8**, 1-74
12. Lechner, J., Wieland, F., and Sumper, M. (1985) *J. Biol. Chem.* **260**, 860-866
13. Wieland, F., Heitzer, R., and Schaefer, W. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 5470-5474
14. Mort, A. J., and Lamport, D. T. A. (1977) *Anal. Biochem.* **82**, 289-309
15. Lechner, J., and Sumper, M. (1987) *J. Biol. Chem.* **262**, 9724-9729
16. Rao, J. K. M., and Argos, P. (1981) *Biochemistry* **20**, 6536-6543
17. Paul, G., and Wieland, F. (1987) *J. Biol. Chem.* **262**, 9587-9593
18. Zaccai, G., Cendrin, F., Haik, Y., Borochov, N., and Eisenberg, H. (1989) *J. Mol. Biol.* **208**, 491-500