Purification and Characterization of a Prokaryotic Glycoprotein from the Cell Envelope of *Halobacterium salinarium*

MATTHEW F. MESCHER AND JACK L. STROMINGER
From the Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138

The glycoprotein which accounts for approximately 50% of the protein and all of the nonlipid carbohydrate of the cell envelope of *Halobacterium salinarium* (Mescher, M. F., Strominger, J. L., and Watson, S. W. (1974) *J. Bacteriol.* 120, 945-954) has been purified and partially characterized. The glycoprotein has an apparent molecular weight of 200,000, is extremely acidic, and has a carbohydrate content of approximately 10 to 12%. The carbohydrate included neutral hexoses, amino sugar, and uronic acid.

Information regarding the number, composition, and mode of attachment of the carbohydrate chains was obtained by isolation and examination of the glycopeptides derived from degradation of cell envelope protein with trypsin and pronase. Trypsin digestion resulted in two glycopeptides. One of these was large (approximately 55,000 daltons) and had most of the neutral hexose linked to it. The carbohydrate moieties consisted of di- and trisaccharides of glucosylgalactose and (uronic acid, glucose)-galactose attached via O-glycosidic linkages between galactose and threonine. The other tryptic glycopeptide had a relatively large heterosaccharide attached to it via an alkaline-stable linkage. The heterosaccharide contained 1 glucose, 8 to 9 galactose, 1 mannose, and 10 to 11 glucosamine residues, and approximately 6 residues of an unidentified amino sugar. The alkaline stability of the linkage and the amino acid composition of glycopeptides resulting from Pronase digestion of the tryptic glycopeptide showed that the heterosaccharide was attached to an asparagine residue, presumably via an N-glycosylamine bond to the amide group.

The intact glycoprotein has a single N-linked heterosaccharide, 22 to 24 O-linked disaccharides, and 12 to 14 O-linked trisaccharides per molecule. N- and O-glycosidic linkages are the most common carbohydrate-protein linkages in mammalian glycoproteins but, to our knowledge, this is the first report of either type of linkage in a prokaryotic cell envelope protein.

Complex carbohydrate-containing polymers are an almost universal feature of the cell envelope and cell surface of both prokaryotic and eukaryotic cells. In prokaryotic organisms these cell envelope carbohydrates typically occur as long chain polysaccharides composed of repeating units as in the case of peptidoglycans, lipopolysaccharides, and teichoic acids. In some cases these polymers have been shown to be covalently linked to proteins (1-3). In contrast to prokaryotic organisms, the majority of the cell surface carbohydrates of eukaryotic cells are present in glycoproteins and occur as relatively short heterosaccharide chains covalently linked to the polypeptide. The most common types of carbohydrate-protein linkages in glycoproteins are the N-glycosidic bond between asparagine and N-acetylglucosamine and the O-glycosidic bond between the reducing terminus of the carbohydrate chain and the hydroxyl group of serine and threonine (or hydroxylysine in the case of collagens) (4).

The gram-negative, obligately halophilic bacteria of the genus *Halobacterium* lack the peptidoglycan or other long chain polysaccharides typical of cell walls of prokaryotic organisms (5). The cell envelope of *Halobacterium salinarium* has been shown to possess a high molecular weight carbohydrate-containing protein which accounts for 40 to 50% of the total cell envelope protein and has all of the nonlipid carbohydrate of the cell envelope bound to it (6). The structural studies described in this report show the *H. salinarium* glycoprotein to resemble the cell surface glycoproteins found in eukaryotic organisms in both the composition of the oligosaccharide moieties and in the nature of the carbohydrate-protein linkage sites. Both N- and O-glycosidic linkage sites are present in the *H. salinarium* glycoprotein, and to our knowledge this is the first report of the occurrence of either type of linkage in a prokaryotic cell envelope protein, i.e. it is so far the only known example of a true glycoprotein in prokaryotic cells.

**MATERIALS AND METHODS**

Acrylamide and *N*,*N*-methylenebisacrylamide were purchased from Bio-Rad. Standard amino acid mixtures for use with the Beckman 120C amino acid analyzer were obtained from Calbiochem. Monosaccharides for use as standards for gas chromatography and colorimetric assays were purchased from Sigma. Trimethylchlorosilane and hexamethyldisilazane were purchased from Pierce Chemical Co. NaBHT, was obtained from New England Nuclear. All chemicals used were reagent grade.
**Cell Envelope Preparation**—*Halobacterium salinarium*, strain 1 (ATCC 19700) was grown in 20-liter batches in a New Brunswick fermentor at 37° with vigorous aeration in the medium previously described (6) and harvested at the end of the log phase of growth using a Sharples refrigerated centrifuge. A yield of approximately 5 g/liter of wet cells was routinely obtained. Cells were broken and the cell envelopes isolated as previously described (6). Lipid-free cell envelope protein was obtained by repeated extraction of wet cell envelope pellets with chloroform/methanol, 2/1 at room temperature.

**Polyacrylamide Gel Electrophoresis in Sodium Dodecyl Sulfate**—Sodium dodecyl sulfate gel electrophoresis was done by the procedure of Weber and Osborn (7). Protein was visualized with Coomassie Brilliant Blue and carboxyblue by the periodate-arsenite-Schiff method of Fairbanks et al. (8).

**Amino Acid Analysis**—Amino acid analysis was done on a Beckman 120C amino acid analyzer after hydrolysis of the sample in 6 N HCl for varying times at 100°, in vacuo. Hydrolysates were evaporated to dryness under a stream of N2. Cysteine and cystine were determined as cysteic acid following hydrolysis of parallel samples in 6 N HCl in the presence of 0.1 M dimethylsulfoxide (9). Data presented are based on calculation of cysteic acid content relative to alanine, which is stable under these hydrolysis conditions. a-Aminobutyric acid was determined on the long column of the Beckman 120C amino acid analyzer where it eluted between alanine and valine. Authentic L-a-aminobutyric acid (Sigma) was used as the standard.

**Carbohydrate Analysis**—Quantitative analysis of neutral sugars was done by gas chromatography of their trimethylsilyl ethers (10). Gas chromatography was done on a glass column (6 ft x 2 mm) packed with 3% SE-30 on SO/100 Chromosorb W HP. Initial temperature was 140° and temperature rise was programmed at a rate of 0.5°/min.

Identification of H-labeled alditoles was done by gas chromatography of their trimethylsilyl ethers (10). Samples were hydrolyzed for varying times in 1.5 N HCl, in vacuo and evaporated to dryness under a stream of N2. Derivatization was done by addition of a freshly prepared 1/15 mixture of trimethylchlorosilane, hexamethyldisilazane, and pyridine. Samples were incubated for 30 min at room temperature, evaporated to dryness under N2, and redissolved in heptane. Gas chromatography was done on a glass column (6 ft x 2 mm) packed with 37% SE-30 on 80/100 Chromosorb W HP. Supelcoport (Supelco) run isothermally at 185°. The outlet end of the column was fitted with a 1:10 stream splitter: 0.1 of the sample going to a flame ionization detector and 0.9 of the sample collected in 1-min fractions on glass wool and counted in toluene-Omnifluor as standards.

**Alkaline Borohydride Degradation**—Alkaline borohydride degradation was carried out in 0.1 N NaOH, 1 M NaBH4 for 24 hours at 37° at a peptide concentration of approximately 2.5 mg/ml (17). Following incubation the sample was acidified (to pH 4) with concentrated acetic acid. Samples for amino acid analysis were then dried on a rotary evaporator, washed several times with methanol, and finally hydrolyzed. For further study of the released oligosaccharides the following, following incubation, was diluted to a sodium ion concentration of approximately 1 M with 0.01 N formic acid and passed through a column of Bio-Rad AG-50-X8, 20 to 100 mesh (H+) having 5 times the number of equivalents of sodium ion present. The column was washed with 6 column volumes of 0.01 N formic acid, and the effluent was lyophilized and washed repeatedly with methanol on a rotary evaporator (17). Alkaline borohydride degradation in the presence of a palladium catalyst was done as described by Tanaka and Pugmam (18).

**Gel Filtration and Ion Exchange Chromatography**—Bio-Gel P-10, P-150 and Bio-Gel A-5m were obtained from Bio-Rad. Sephadex G-100 and Whatman DE52 (microgranular) were obtained from Pharmacia. All were prepared according to the manufacturer's directions.

**RESULTS**

**Whole Glycoprotein**

**Purification**—It was previously shown that the *Halobacterium salinarium* cell envelope glycoprotein could be partially purified by aqueous phenol extraction of lipid-free cell envelopes (6). The major contaminants present in the aqueous phase along with the glycoprotein were nucleic acids and degradation products of the glycoprotein which result from proteolysis during preparation of the cell envelopes. All of the remaining envelope proteins were present in the phenol phase following extraction. Following dialysis of the aqueous phase against distilled water, the glycoprotein was further purified by chromatography on DEAE-cellulose (Fig. 1). The phenol remaining in the aqueous phase did not bind to the column and appeared as a large peak of 260 nm absorbing material which gave a positive color reaction in the protein assay of Lowry et al. (16). The glycoprotein eluted as a single peak of protein and carbohydrate in a linear NaCl gradient well ahead of the nucleic acid. Fractions containing the glycoprotein were pooled, and the solution was dialyzed for 24 hours against distilled water and lyophilized. The yield of glycoprotein was approximately 30%. Sodium dodecyl sulfate-acrylamide gels (Fig. 2) of the purified glycoprotein showed it to be free of other envelope proteins and to be present in the high molecular weight form to the extent of greater than 95%. Small amounts of proteolytic degradation products of the glycoprotein could be seen on gels overloaded with protein. Attempts to remove these degradation products from the high molecular weight material by gel filtration were unsuccessful due to aggregation of the glycoprotein in the absence of detergent. This aggregation may explain the relatively poor yield obtained with DEAE-cellulose chromatography.

**Molecular Weight of Glycoprotein**—The purified glycoprotein ran almost coincidentally with myosin on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Amino sugars were determined on the short column of a Beckman 120C amino acid analyzer using glucosamine-HCl and galactosamine-HCl as standards.
Structure of a Prokaryotic Glycoprotein from H. salinarium


sulfate-acrylamide gels (Fig. 3), indicating an apparent molecular weight of approximately 200,000. Glycoproteins may give anomalously high molecular weight values on sodium dodecyl sulfate gels and the value of 200,000 should be considered as an upper limit for the true molecular weight. However, the low percentage of carbohydrate in the H. salinarium glycoprotein (see below) would indicate that this molecular weight estimate is not greatly in error (19).

Amino Acid and Carbohydrate Composition—The amino acid and carbohydrate composition of the glycoprotein are shown in Table I. The most notable feature of the amino acid composition is the high content of acidic residues (uncorrected for amides) and the low content of basic residues. If the ammonia peak obtained on amino acid analysis is taken as a maximum estimate of the amide content of the glycoprotein, then it would still have a 12.5 mole per cent excess of acidic over basic amino acids. This value for amide content is probably a considerable overestimate, as no special precautions were taken to exclude ammonia from the hydrolysates.

The glycoprotein had a carbohydrate content of approximately 9% based on colorimetric assays of neutral hexose and amino sugar. Qualitative gas chromatographic analysis of the carbohydrates indicated the presence of approximately equal amounts of glucose and galactose and a trace of mannose. Two amino sugars were found, one of which had a retention time identical with glucosamine.

Separation of Hexose- and Amino Sugar-containing Glycopeptides—It was previously shown that trypsin cleavage of the glycoprotein produced two different glycopeptides, one containing all of the amino sugar and the other having most (approximately 80%) of the neutral hexose (6). These glycopeptides were isolated and purified in order to further study the nature of the carbohydrate-protein linkages in the H. salinarium glycoprotein. The glycopeptides were obtained from trypsin digests of lipid free cell envelope protein to facilitate their large scale preparation. Glycoproteins often have heterogeneity in their carbohydrate moieties, and using cell envelope protein

TABLE I

Amino acid and carbohydrate composition of Halobacterium salinarum envelope glycoprotein

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Hexose</th>
<th>Amino Sugar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>9.7</td>
<td>1.0</td>
</tr>
<tr>
<td>Histidine</td>
<td>2.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Arginine</td>
<td>12.7</td>
<td>2.1</td>
</tr>
<tr>
<td>Aspartic Acid</td>
<td>134.1</td>
<td>21.2</td>
</tr>
<tr>
<td>Threonine</td>
<td>57.8</td>
<td>9.8</td>
</tr>
<tr>
<td>Serine</td>
<td>65.0</td>
<td>10.5</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>74.3</td>
<td>12.0</td>
</tr>
<tr>
<td>Proline</td>
<td>26.7</td>
<td>4.4</td>
</tr>
<tr>
<td>Glycine</td>
<td>56.0</td>
<td>9.2</td>
</tr>
<tr>
<td>Alanine</td>
<td>37.8</td>
<td>6.4</td>
</tr>
<tr>
<td>Valine</td>
<td>46.7</td>
<td>8.0</td>
</tr>
<tr>
<td>Methionine</td>
<td>4.8</td>
<td>0.9</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>26.4</td>
<td>4.4</td>
</tr>
<tr>
<td>Leucine</td>
<td>31.1</td>
<td>5.1</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>15.7</td>
<td>2.6</td>
</tr>
<tr>
<td>Phenyllalanine</td>
<td>10.9</td>
<td>1.7</td>
</tr>
<tr>
<td>Half-cystine</td>
<td>N.D.*</td>
<td>N.D.</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>N.D.</td>
<td></td>
</tr>
<tr>
<td>Hexose</td>
<td>41.7</td>
<td></td>
</tr>
<tr>
<td>Amino sugar</td>
<td>8.6</td>
<td></td>
</tr>
</tbody>
</table>

* N.D., not determined.
as starting material for glycopeptide preparation also avoided
the possibility of selective loss, undetectable by sodium dodecyl
sulfate gel electrophoresis, during purification of the glyco-
protein on DEAE-cellulose. Gel filtration of the digest on
Sephadex G-100 resulted in separation of the amino sugar- and
hexose-containing glycopeptides (Fig. 4), which will subse-
quently be referred to as the HT glycopeptide

**HT Glycopeptide**

**Purification of HT Glycopeptide**—Sodium dodecyl sulfate
gels of the glycopeptide excluded from the G-100 column and
containing most of the hexose showed the presence of a single
band with both protein and carbohydrate staining (6). The
high 260 nm absorption indicated contamination with nucleic
acid, and the glycopeptide was further purified by chromato-
graphy on DEAE-cellulose (Fig. 5). Approximately 80 to 90% of
the hexose applied to the column was recovered in the eluant.
The glycopeptide eluted at 0.03 M NaCl and accounted for at
least 95% of the recovered hexose. The small anthrone-positive
peak eluting with the nucleic acid was present in all prepara-
tions and may result from the glycoprotein degradation prod-
ucts present in the isolated cell envelopes. It was not further
investigated. Fractions containing the major hexose peak were
pooled and dialyzed overnight against distilled water.

**Estimated Molecular Weight**—The HT glycopeptide had a
mobility on sodium dodecyl sulfate-acrylamide gel electropho-
resis which indicated a molecular weight of approximately
55,000 (Fig. 6). The glycopeptide contained approximately 20%
carbohydrate (Table II), suggesting that the molecular weight
as determined by sodium dodecyl sulfate gels was not overesti-
ated by more than 10,000 (19). Results obtained upon gel
filtration of the glycopeptide were consistent with it being a
large peptide. As in the case of the intact glycoprotein, the
glycopeptide aggregated in aqueous buffer and eluted in the
void volume on a Bio-Gel A-5m column run in 0.05 M Tris
buffer, pH 7.2. When run in buffer containing 0.1 M EDTA the
aggregation was partially eliminated (Fig. 7A) and dialysis of
the glycopeptide for 24 hours against 0.1 M EDTA prior to gel
filtration on a column equilibrated in 0.1 M EDTA resulted in a
single included peak (Fig. 7B). The glycopeptide eluted
slightly behind phosphorylase A (93,000 daltons) on the gel
filtration column. Lack of adequate standards prevented an
accurate estimate of molecular weight by this method but the
gels result support the conclusion that the glycopeptide is large
and that the value of 55,000 obtained by sodium dodecyl
sulfate-acrylamide gel electrophoresis is not greatly in error.

**Amino Acid and Carbohydrate Composition**—The amino
acid and carbohydrate composition of the HT glycopeptide are
shown in Table II. The amino acid composition of the
glycopeptide showed that a number of potential tryptic cleav-
age sites remained intact following trypsin digestion. This was
probably a result of their being inaccessible to the enzyme due
to blocking by the carbohydrate moieties (20). Glucose and
galactose were present in approximately equal amounts and
were the only neutral hexoses detected by gas chromatography.
No pentoses were found. No amino sugars were detected by gas
chromatography or on the amino acid analyzer. Hexuronic
acid, measured by the colorimetric method of Blumenkranz

1The abbreviations used are: HT glycopeptide, hexose-containing
tryptic glycopeptide; AST glycopeptide, amino sugar-containing tryptic
glycopeptide; ACP glycopeptide, amino sugar-containing tryptic
peptide after pronase digestion; FAS stain, periodate-arsenite-Schiff
stain.

![Fig. 4. Gel filtration of trypsin-digested whole envelope protein.](http://www.jbc.org/)

![Fig. 5. DEAE-cellulose chromatography of the HT glycopeptide. A sample containing approximately 24 amol of hexose was applied to a DE-52 column (0.8 x 23 cm) in 10 ml of 0.05 M Tris, pH 7.2. Elution was done with 60 ml of 0.05 M Tris, pH 7.2, followed by a linear gradient (beginning at Fraction 23) consisting of 60 ml of starting buffer in the mixing chamber and 0.5 M NaCl in starting buffer in the reservoir. Following the gradient the column was eluted with 60 ml of 1.0 M NaCl in 0.05 M Tris, pH 7.2. Fractions (3 ml) were collected and aliquots were assayed for hexose by the anthrone method (13).](http://www.jbc.org/)
TABLE II

Composition of HT peptide before and after alkaline borohydride treatment

<table>
<thead>
<tr>
<th>Amino acid composition</th>
<th>Before β-elimination</th>
<th>After β-elimination</th>
<th>∆</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>4.0</td>
<td>3.6</td>
<td>-0.4</td>
</tr>
<tr>
<td>Histidine</td>
<td>4.5</td>
<td>4.9</td>
<td>+0.4</td>
</tr>
<tr>
<td>Arginine</td>
<td>6.2</td>
<td>6.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>45.5</td>
<td>44.5</td>
<td>-1.2</td>
</tr>
<tr>
<td>Threonine</td>
<td>71.6</td>
<td>74.2</td>
<td>+2.6</td>
</tr>
<tr>
<td>Serine</td>
<td>29.8</td>
<td>28.6</td>
<td>-1.2</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>70.2</td>
<td>67.9</td>
<td>-2.3</td>
</tr>
<tr>
<td>Proline</td>
<td>23.3</td>
<td>19.5</td>
<td>-3.8</td>
</tr>
<tr>
<td>Glycine</td>
<td>52.2</td>
<td>51.2</td>
<td>-1.0</td>
</tr>
<tr>
<td>Alanine</td>
<td>38.0</td>
<td>39.8</td>
<td>+1.8</td>
</tr>
<tr>
<td>Valine</td>
<td>43.2</td>
<td>45.4</td>
<td>+2.2</td>
</tr>
<tr>
<td>Methionine</td>
<td>14.4</td>
<td>13.7</td>
<td>-0.7</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>23.6</td>
<td>23.8</td>
<td>+0.2</td>
</tr>
<tr>
<td>Leucine</td>
<td>40.6</td>
<td>40.8</td>
<td>+0.2</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>7.8</td>
<td>8.4</td>
<td>+0.6</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>12.8</td>
<td>15.6</td>
<td>+2.8</td>
</tr>
<tr>
<td>Cysteic acid</td>
<td>4.4</td>
<td>4.8</td>
<td>+0.4</td>
</tr>
<tr>
<td>α-Aminobutyric acid</td>
<td>0.0</td>
<td>31.2</td>
<td>+31.2</td>
</tr>
<tr>
<td>Glucose*</td>
<td>34.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Galactose**</td>
<td>33.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uronic acid*</td>
<td>12.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Determined by hydrolyzing in 1.5 N HCl at 110° for varying times followed by gas chromatography as the trimethylsilyl derivatives (see "Materials and Methods"). Correction was made for loss during hydrolysis.

** Determined colorimetrically as described under "Materials and Methods."

production of 31 residues of α-aminobutyric acid. (Approximately 1 residue of serine was also lost, but the accuracy of the analysis did not allow a conclusion to be made as to the significance of this.) The 83% yield of α-aminobutyric acid relative to the amount of threonine lost was somewhat higher than has been observed by others using a lower concentration of NaBH₄ (18). When alkaline borohydride treatment of the HT glycopeptide was done using 0.1 N NaOH/0.3 M NaBH₄, the yield of α-aminobutyric acid was reduced to 66%. Borohydride reduction in the presence of palladium chloride as a catalyst by the method of Tanaka and Pigman (18) resulted in a yield of 102%.

The carbohydrate content of the HT glycopeptide (Table II) indicated the presence of approximately 1 glucose, 1 galactose, and 0.33 hexuronic acid residues per threonine residue lost upon alkaline borohydride treatment. This suggested that the O-linked oligosaccharides are small, having an average of 2 to 3 carbohydrate residues per chain.

Identification of Linkage Sugars—Alkaline borohydride cleavage of O-glycosidic linkages allows identification of the linkage sugar as the corresponding alditol following acid hydrolysis. Alkaline reductive cleavage of the HT glycopeptide in the presence of NaBH₄ resulted in the release of oligosaccharides having H-labeled alditols in the terminal position. When the released oligosaccharides were examined by gel filtration the majority of the label eluted in the region expected for di-
and tri-saccharides (Fig. 9). There was, in addition, a small amount of higher molecular weight material. The column effluent was pooled in three fractions, as shown in Fig. 9, and lyophilized. Following acid hydrolysis, the 3H-labeled alditols of each pooled fraction were identified by gas chromatography (Fig. 10). Galactitol was the only labeled alditol found in the low molecular weight fractions (pools II and III). No labeled alditol was detectable in pool I following hydrolysis. The labeled material in this region of the column probably resulted from partial degradation of the peptide portion of the glycopeptide during alkaline treatment. These results showed that galactose is the only hexose involved in O-glycosidic linkages to threonine and support the suggestion that the oligosaccharides are present as di- and trisaccharides of galactose, glucose and hexuronic acid.

**Separation and Composition of O-linked Oligosaccharides**

O-linked oligosaccharides were prepared from the HT glycopeptide by alkaline borohydride degradation as described under “Materials and Methods,” and chromatographed on a Bio-Rad AG-IX4 formate column (Fig. 11). Oligosaccharides containing only alditol and neutral hexose did not bind to the column (Fig. 11, peak I), while those containing, in addition, hexuronic acid bound and were eluted with 2 N formic acid (Fig. 11, peak II). Gas chromatographic analysis at peak I showed this material to have equal amounts of galactitol and glucose. Peak II material had equal amounts of galactitol and glucose as determined by gas chromatography, and a hexuronic acid:glucose ratio of 0.85 based on colorimetric assays. The identity of the hexuronic acid was not determined.

These results showed the O-linked oligosaccharides to be of two types: glucosylgalactose disaccharides and trisaccharides of hexuronic acid, glucose, and galactose. This conclusion is...
consistent with the carbohydrate composition and number of threonine linkage sites of the HT glycopeptide, and with the finding that galactose is the only sugar involved in O-glycosidic linkages to threonine.

**Amino Sugar-containing Glycopeptide**

**Purification of AST and ASP Glycopeptides**—The amino sugar-containing glycopeptide isolated by gel filtration from a trypsin digest of whole cell envelope protein (Fig. 4) was further purified by aqueous phenol extraction followed by gel filtration of the aqueous phase (Fig. 12, A and B). Further degradation of the peptide was done by digestion of the purified AST glycopeptide with pronase. Fig. 12, C and D shows the gel filtration column profiles obtained for ASP1; the glycopeptide resulting from pronase digestion at 37°C for 48 hours. A glycopeptide (ASP1II) resulting from more extensive digestion of AST with pronase was also isolated. AST was digested with pronase (1/2, pronase/amo amino sugar) for 24 hours at 55°C (21) and isolated by gel filtration on Bio-Gel P-10. This procedure was repeated twice and the glycopeptide was finally purified by running twice on Bio-Gel P-10. The column profiles were identical with those shown for ASP1 (Fig. 12, C and D).

**Composition**—The results of amino acid analysis of acid hydrolysates of the purified AST and ASP glycopeptides are shown in Table III. The number of residues shown for the tryptic peptide is based on the composition of the smaller peptides resulting from pronase digestion. The tryptic glycopeptide gave a single ninhydrin-positive spot on high voltage electrophoresis, and Edman degradation (22) gave the PTH-derivative of alanine with a yield consistent with the peptide composition shown in Table III. A potential trypsin cleavage site was intact in the AST peptide, probably due to masking by the carbohydrate moiety (20). Pronase digestion resulted in removal of most of the amino acids present in the AST peptide and yielded glycopeptides whose amino acid composition was dependent on the conditions used for the digestion. Further treatment of the isolated glycopeptides with either leucine aminopeptidase or carboxypeptidase A resulted in no further degradation. The inability to degrade the glycopeptide to a single amino acid residue is probably a result of the large carbohydrate moiety blocking access of the proteolytic enzymes to the peptide bonds near the carbohydrate linkage site. The presence of a single aspartic acid residue in the ASP1I glycopeptide and the stability of the carbohydrate linkage to alkaline borohydride degradation strongly suggest that the carbohydrate is attached via an N-glycosylamine bond to the amide group of asparagine.

Carbohydrate analysis (Table III) of the isolated glycopeptides by gas chromatography showed the carbohydrate moiety to be a relatively large heterosaccharide. Approximately 27 mol

![Fig. 12. Purification by gel filtration of tryptic and pronase amino sugar-containing glycopeptides. A, amino sugar-containing fractions from the Sephadex G-100 column of a trypsin digest (Fig. 4) were pooled, lyophilized, and redisolved in 10 mM NH₄-bicarbonate. A 2.2-ml sample (11 μmol of amino sugar) was applied to a Bio-Gel P-150 column (1.5 x 43 cm) and eluted with 10 mM NH₄ bicarbonate. Fractions (2.7 ml) were collected and assayed for amino sugar and Aₛeurop. Amino sugar containing fractions were pooled and lyophilized. B, glycopeptide from A above was phenol-extracted as previously described with a yield of 90 to 100% and rerun on the same Bio-Gel P-150 column following dialysis. A 3-ml sample (4.5 μmol of amino sugar) was applied and 2.7-ml fractions were collected and assayed. C, AST glycopeptide (3.5 μmol of amino sugar) was digested with pronase (Calbiochem) for 48 hours at 37°C in 5 ml of 0.05 Tris, pH 7.2. Pronase, at a ratio of enzyme to amino sugar of 1:2, was added at 24 and 36 hours. The 3-ml sample was applied to a Bio-Gel P-10 column (1.5 x 43 cm) and eluted with 10 mM NH₄ bicarbonate. Fractions (1.6 ml) were collected and assayed. Amino sugar-containing fractions were pooled, lyophilized, and redisolved in 10 mM NH₄ bicarbonate. D, glycopeptide from C above was rerun on the same Bio-Gel P-10 column. A 1-ml sample (3 μmol of amino sugar) was applied and 1.1-ml fractions were collected and assayed.](http://www.jbc.org/)
The occurrence of 0-glycosidic links between mannose and threonine was determined by gas chromatography as described under "Materials and Methods." Glucosamine was determined on the short column of a Beckman 120C amino acid analyzer following hydrolysis of the glycopeptide in 6 N HCl at 110° C in vacuo for varying times, and correction was made for loss during hydrolysis. A quantitative estimate of the unidentified amino sugar was made by subtracting the contribution due to glucosamine from the value for total amino sugar obtained by the modified Morgan-Elson assay (14).

<table>
<thead>
<tr>
<th>Amino Sugar</th>
<th>ASPT</th>
<th>ASP II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucosamine</td>
<td>10.6</td>
<td>8.0</td>
</tr>
<tr>
<td>Mannose</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Galactose</td>
<td>8.5</td>
<td>8.0</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>Glucosamine</td>
<td>10.6</td>
<td>9.9</td>
</tr>
<tr>
<td>Uronic acid</td>
<td>N.D.</td>
<td>N.D.</td>
</tr>
</tbody>
</table>

* Determined colorimetrically as described under "Materials and Methods."

of carbohydrate were found/mol of peptide, and included glucose, galactose, mannose, (in a ratio of approximately 1/8/1) and hexosamines. Glucosamine was present along with a second amino sugar which eluted from the short column of a Beckman 120C amino acid analyzer in a position different from that of galactosamine or mannosamine. The hexosamine nature of this component was confirmed (Fig. 13) by assaying the column effluent for reducing power and for hexosamine by the Morgan-Elson procedure (12) which is specific for 2-amino-2-deoxyhexoses. Attempts to isolate and identify this sugar by chromatography on a Dowex 50-H + column (23) were unsuccessful. Only glucosamine could be detected in the column effluent. We are currently attempting to identify the amino sugar by gas chromatography-mass spectrometry.

**DISCUSSION**

The results of the structural studies reported here show the high molecular weight 22,000 dalton tryptic peptide containing all of the amino sugar attached to it and its further degradation to a glycopeptide (ASPT) having a single aspartate residue without loss of carbohydrate indicates that there is a single N-glycosylamine linkage site per protein molecule. This conclusion is consistent with the molecular weight and amino sugar content of the purified glycopeptide. A protein of 200,000 daltons having 8.6 μmol of amino sugar/100 mg dry weight (Table I) would have 10.7 μmol of amino sugar/mole of protein. This value is in good agreement with the value of 16.6 residues of amino sugar per heterosaccharide chain (average value from Table III) and supports the conclusion that there is one N-glycosylamine linkage site per protein molecule. In the case of the O-glycosidic linkage sites, the data of Table I show that there are 81 mol of hexose/200,000 dalton protein, or 69 mol after subtracting the contribution due to the N-linked heterosaccharide. This gives a value of 34 to 35 glucosylgalactose disaccharide units/protein molecule. The isolation of a 55,000 dalton tryptic peptide (HT glycopeptide) containing all of the O-linked units shows them to be localized in approximately one-quarter of the protein. The data of Table II indicate that there are 34 (based on the amount of glucose and galactose) to 38 (based on the amount of threonine lost) sites/55,000 dalton peptide, i.e., 34 to 38 O-linked units/molecule. This estimate is in good agreement with that obtained from the whole protein analytical data. The results obtained for the structure of the H. salinarium glycopeptide are summarized in Fig. 14.

Gottschalk (25) has proposed that glycoproteins are best defined as "... conjugated proteins containing as prosthetic group(s) one or more heterosaccharide(s), usually branched
Glycoproteins are distinguished from heteropolysaccharide-protein complexes in that the latter have carbohydrate moieties which contain serially repeating units, a relatively high number of sugar residues per chain, and a linear structure (e.g. chondroitin sulfate-protein and keratin sulfate-protein). The structural studies described in this report clearly show the 200,000 dalton polypeptide of *H. salinarium* cell envelopes to be a glycoprotein by these criteria.

Whether other species of bacteria possess such glycoproteins is at present an open question. Covalently linked heteropolysaccharide-protein complexes have been shown to be present in the cell envelopes of a variety of bacteria. These include the murein-lipoprotein of *Escherichia coli* (3, 29) and the lipopolysaccharide-protein found in *Serratia marcescens* (1), *E. coli* (1, 2), and *Nitrosooccus oceanus*. Braun and Bosch (26) have shown the lipoprotein to be attached to murein via a linkage between the 6-amino group of its COOH-terminal lysine and the dianominopine acid of the murein. The nature of the lipopolysaccharide-protein linkage has not been determined. There are reports of "glycoproteins" in several different kinds of bacteria, including *E. coli* (27), Group A streptococci (28, 29) and *Streptomyces albus* (30). The results obtained in these studies show that there are carbohydrate-protein complexes present in the cell envelopes of these bacteria and that they are probably covalently linked. Further purification and characterization of these components and determination of the nature of the carbohydrate-protein linkage will be necessary in order to determine whether they are true glycoproteins as found in *H. salinarium* and eukaryotic cells or heteropolysaccharide-protein complexes (or possibly precursors or degradation products of these complexes). It appears, however, that glycoproteins occur rarely in prokaryotic cells, and if this is confirmed by continued studies, then the evolutionary significance of the occurrence of glycoproteins in eukaryotic organisms and in halobacteria needs to be explained.

The majority of glycoproteins contain only one type of glycopeptide bond, but both N- and O-glycosidic linkages are present in some, including the glomerular basement membrane (31), a yeast cell wall glycoprotein (32) and the major glycoprotein of human erythrocyte membranes (33), as well as a number of nonmembrane glycoproteins (4). These membrane glycoproteins are similar to the *H. salinarium* glycoprotein in that all have one, or a small number of relatively large N-linked heterosaccharides and a larger number of small (1 to 4 sugar residues) O-linked units. The asymmetric distribution of the carbohydrate moieties within a portion of the polypeptide chain as seen for the O-linked units in the *H. salinarium* glycoprotein may be a common feature of membrane glycoproteins. All of the carbohydrate of the major erythrocyte glycoprotein is located in the NH2-terminal one-half of the polypeptide chain (33). This similarity in number, size, and distribution of carbohydrate units may reflect a similarity in their functional role in these proteins.

The O-linked units of the *H. salinarium* glycoprotein most closely resemble the hydroxylysine-linked units found in members of the collagen family. The disaccharide units of glomerular basement membrane and several vertebrate collagens are 2-O-α-D-glucopyranosyl-D-galactose (31). It will be of interest to determine whether the same 1,2-glycosidic linkage is present between the glucose and galactose of the threonine-linked disaccharides in the *H. salinarium* glycoprotein. Earthworm cuticle collagen has di- and trisaccharide units of galactose (34), in this case O-glycosidically linked to serine and threonine residues (35). Uronic acid-containing disaccharides have also recently been found in the cuticle collagen of *Nereis* (clamworm) (36). The disaccharide was shown to be 6-O-α-D-gluconorosyl-D-mannose and was attached to the protein via O-glycosidic linkages to threonine.

The similarity of the *H. salinarium* glycoprotein to the cell surface glycoproteins found in eukaryotic cells suggests that halobacteria may provide a simple model system for studying the function of the glycan chains and their biosynthesis and attachment to the protein. The observation that the antibiotic bacitracin inhibits the growth of halobacteria and causes a morphological conversion from normal rod-shaped cells to spherical cells (37) has provided a means of approaching these problems. Bacitracin kills normal bacteria by associating with the polyprenol pyrophosphate released at the end of the lipid carrier cycle in peptidoglycan synthesis. Complex formation prevents enzymatic dephosphorylation of the lipid, thus preventing its reentry into the cycle and stopping peptidoglycan synthesis (38). There is now considerable evidence that protein glycosylation in eukaryotic cells may occur, at least partially, via polyprenol phosphate-linked intermediates (39). The effect of bacitracin on the growth of *H. salinarium* suggests that glycosylation of its cell envelope glycoprotein occurs via a cyclic pathway of lipid-linked intermediates. Some direct evidence for this has been obtained and these studies are being continued 3. In addition, the possible structural role of the glycoprotein is being investigated. Halobacteria lack the peptidoglycan layer necessary to most prokaryotic organisms for normal growth and cell division and for maintenance of a stable morphology. Despite the lack of a rigid peptidoglycan, halobacteria maintain a stable rod-shaped morphology when grown at optimum salt concentration. The glycoprotein is the predominant envelope component, accounting for 40 to 50% of the total protein and all of the nonlipid carbohydrate, and may be the main structural component (6). Work is in progress to determine whether this is the case and, if so, to determine whether the glycan chains play a role in the function of the glycoprotein. The ability of bacitracin to cause a morphological change from normal rod-shaped cells to spherical forms, probably by blocking glycosylation, suggests that this may be true (37). The glycoprotein may also be largely responsible for the requirement for high monovalent cation concentrations for maintenance of the structural integrity of the cell envelope.

Amino acid analyses of halobacteria cell envelope hydrolysates have shown the cell envelope proteins to be very acidic (40).

---

1M. F. Mescher, and S. W. Watson, manuscript in preparation.

8Preliminary results were presented at the Third International Symposium of Glycoconjugates, Brighton, England, July 1975.
and it has been suggested that monovalent cations of sufficiently high concentrations might effectively neutralize the excess negative charges on the envelope and prevent its dispersion (41, 42). The high glycoprotein content of the cell envelope and the extremely high acidity of the glycoprotein (33 mol per cent glutamic + aspartic acids, uncorrected for amides, as compared to 4 mole per cent basic amino acids, Table I) suggest that it may make a major contribution to the high salt requirement.

REFERENCES

Purification and characterization of a prokaryotic glucoprotein from the cell envelope of *Halobacterium salinarium.*

M F Mescher and J L Strominger


Access the most updated version of this article at [http://www.jbc.org/content/251/7/2005](http://www.jbc.org/content/251/7/2005)

Alerts:
- When this article is cited
- When a correction for this article is posted

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

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/251/7/2005.full.html#ref-list-1](http://www.jbc.org/content/251/7/2005.full.html#ref-list-1)