The Carbohydrate Units of Thyroglobulin*

ROBERT G. SPIRO†

From the Department of Biological Chemistry, Harvard Medical School, and the Elliott P. Joslin Research Laboratory, Boston, Massachusetts 02215

(Received for publication, September 25, 1964)

Thyroglobulin is a high molecular weight glycoprotein which has previously been shown to contain approximately 300 monosaccharide residues per molecule in the form of galactose, mannose, glucosamine, sialic acid, and fucose (2).

The purpose of the present investigation was to initiate a study of the structure of the carbohydrate portion of this glycoprotein by determining the size, number, and composition of the units among which these monosaccharides are distributed. For this purpose, the peptide portion of native calf thyroglobulin was subjected to extensive digestion with the enzyme Pronase, making it possible to obtain the carbohydrate units of this molecule with only a few amino acids attached. The glycopeptides prepared in this manner were fractionated and then characterized by chemical and physical means.

These studies indicated that thyroglobulin contains two distinct types of carbohydrate units differing both in molecular weight and composition, with a total of approximately 23 units in each molecule.

EXPERIMENTAL PROCEDURE

Preparation of Thyroglobulin—The thyroglobulin* used in this study was prepared from calf thyroid glands as previously described (3) by phosphate buffer fractionation (4). Some of the physical properties and the carbohydrate analyses of this preparation have been described (2, 3).

Proteolytic Digestion of Thyroglobulin—Streptomyces griseus protease (Pronase) was obtained from the California Corporation for Biochemical Research. The digests were carried out in 0.2 M sodium phosphate buffer, pH 7.8, at 37° in the presence of 0.0015 M CaCl2. Thyroglobulin was present at a concentration of 2.5%. Pronase was added initially to equal 1% of the thyroglobulin on a weight basis, and at 48 and 72 hours further additions of the enzyme equal to 0.5% of the substrate were made. The incubations were carried out for a total period of 96 to 120 hours. Small amounts of toluene were added to prevent bacterial growth. The course of the digestion was followed by the ninhydrin reaction (5) with leucine as a standard.

At the end of the incubation, a small amount of insoluble material representing denatured enzyme was removed by centrifugation. The clear supernatant solution was then dialyzed in 18/32 Visking cellophane tubing at 2–4°, first against 0.1 M NaCl and then against distilled water, for a total dialysis period varying from 64 to 216 hours. The nondialyzable material was lyophilized and passed through a small column of Dowex 50-X10, 20 to 50 mesh (II+ form).

Ultracentrifugal Measurements—Ultracentrifugal analyses were performed in a Spinco Model E instrument at 20°. The determinations were carried out in potassium phosphate buffer, pH 6.5, 0.1 ionic strength, with NaCl making up 80% of the ionic strength. For the determination of sedimentation constants a synthetic boundary cell was employed at 56,100 r.p.m.

Molecular weights were determined by the short column sedimentation equilibrium method (6). A speed of 20,410 r.p.m. was employed and the angle of the schlieren diaphragm was 50°.

Chromatography of Glycopeptides on Diethylaminoethyl Cellulose Columns—Columns of DEAE-cellulose (2.1 X 65 cm) were prepared and equilibrated with 0.0005 M sodium phosphate buffer at pH 7.6, as previously described (7). These columns were employed to chromatograph the nondialyzable fraction of Pronase digests from approximately 1500 mg of thyroglobulin. Samples were titrated to pH 7.6 and placed on the columns in the 0.0005 M sodium phosphate buffer at that pH. An additional 350 ml of this buffer was then run through the columns. Chromatography was continued with a gradual linear concentration gradient consisting of 1400 ml of 0.0005 M sodium phosphate buffer at pH 7.6 in the mixing chamber and an equal volume of 0.03 M phosphate buffer at pH 7.6 in the reservoir. A flow rate of 22 ml per hour was maintained and fractions of 14 ml were collected. The elution was followed by analysis of aliquota by the ninhydrin (5), anthrone (8), and resorcinol (9) reactions. The fractions making up various peaks were pooled and salt was removed by dialysis for 8 hours in 18/32 Visking cellophane tubing against distilled water.

Chromatography of Glycopeptides on Charcoal-Celite Columns—The glycopeptides were fractionated on charcoal-Celite (10) columns of 2.5 cm diameter with a column volume of 1 ml per mg of hormone present in the sample. After application of the sample in water or buffer, the column was washed with distilled water (8 ml per column volume). Elution was accomplished in a stepwise manner with solutions of increasing ethanol concentration in the following order: 17% ethanol (25 ml per column volume); 20% (20 ml per ml); 25% (10 ml per ml); 30% ethanol (10 ml per ml); and 50% ethanol (25 ml per column volume).

* This work was supported by a grant from the American Heart Association and by Grant AM 05363 from the National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, United States Public Health Service. A preliminary report has been published (1).
† This work was done during the tenure of an Established Investigatorship from the American Heart Association.
* Kindly made available by Dr. M. J. Spiro.
Fractionation of Glycopeptides by Gel Filtration—For gel filtration Sephadex G-25 (100 to 200 mesh) or Sephadex G-50 (100 to 200 mesh) was packed to a height of 80 cm in glass columns 2.2 cm in diameter, fitted with sintered glass plates. The Sephadex was equilibrated with 0.1 M sodium phosphate buffer at pH 7.0 saturated with toluene to prevent bacterial growth. The sample, consisting of the nondialyzable fraction of a Pronase digest of approximately 300 mg of thyroglobulin, was applied to the column in 4.0 ml of this buffer. Elution was achieved with the same buffer at a rate of 15 ml per hour, and 5.2-ml fractions were collected. Aliquots of each fraction were analyzed by the anthrone and resorcinol reactions. The fractions comprising the elution peaks were pooled and freed of salt by dialysis against distilled water for 8 hours in 18/32 Visking tubing.

Paper Electrophoresis of Glycopeptides—Electrophoresis was performed at a potential of 12 to 20 volts per cm on Whatman No. 1 paper in a water-cooled enclosed strip electrophoresis apparatus. The buffer employed was pyridine-acetic acid-water (25:1:228) at pH 6.4, 0.25 N acetic acid at pH 2.6, and sodium phosphate of 0.05 ionic strength at pH 9.0. The glycopeptides were detected by the ninhydrin stain as previously employed (7).

Carbohydrate Analyses—The sugar components were identified and estimated as previously described (11, 12).

Sialic Acid Removal—Sialic acid-free glycopeptides were prepared by hydrolysis with 0.05 N sulfuric acid at 80° for 1 hour, followed by passage through Dowex 1 columns as previously described (7).

Amino Acid Analyses—Amino acid analyses were performed on the hydrolyzed glycopeptides by quantitative paper chromatography as previously described in detail (7). The acidic amino acids were separated from the basic and neutral amino acids on Dowex 1 columns (7). After this separation all of the amino acids present in the glycopeptide could be resolved by paper chromatography in butanol-acetic acid-water (4:1:5) and pyridine-ethyl acetate-water-acetic acid (5:5:3:1) (13). The second system achieved a good separation of the serine, glycine, and glutamic acid present in the neutral amino acid fraction.

Polypeptide measurements were made by the method of Lowry, Rosebrough, Farr, and Randall (14).

Amino-terminal Analyses—NH₂-terminal residues of the glycopeptides were determined with fluorodinitrobenzene as previously described (7). Because of the presence of NH₂-terminal glycine in several glycopeptides, the dinitrophenylated glycopeptides were hydrolyzed in constant boiling HCl for 4 hours as well as 16 hours at 105°. All values were corrected for destruction during acid hydrolysis.

Digestion of Glycopeptides with Leucine Aminopeptidase and Carboxypeptidase—Leucine aminopeptidase (Worthington Biochemical Corporation) with a C₅ of approximately 10 was preincubated in 0.05 M barbital buffer, pH 8.0, with 0.01 M MnSO₄ for 20 minutes at 40°. Carboxypeptidase (Mann Research Laboratories) was treated before use as previously described (7). The glycopeptides (1.0 to 1.25 amoles per ml) were incubated with these enzymes in 0.05 M barbital buffer at pH 8.0 for 45 hours. The digestion with leucine aminopeptidase was carried out at 40° in the presence of 0.002 M MnSO₄. Incubation with carboxypeptidase was at 37° with 0.5 to 0.75% LiCl present. When a sample was incubated with both enzymes in succession, the leucine aminopeptidase was added first, and after 45 hours of incubation the carboxypeptidase was introduced into the incubation mixture. Each enzyme was employed in an amount equal to approximately 25% of the weight of the peptide portion of the glycopeptide. The released amino acids were separated from the remaining glycopeptide by passage through Dowex 50-X16, 20 to 50 mesh (H⁺ form) and were identified and estimated as the free amino acids after elution, as previously described (7). The glycopeptides were recovered from the Dowex 50 effluent and wash by adsorption on small charcoal-Cellite columns followed by elution with 50% ethanol (30 ml per ml of column volume).

RESULTS

Proteolytic Digestion of Thyroglobulin—Extensive proteolytic digestion of native calf thyroglobulin was achieved with the enzyme Pronase. In 96 hours of digestion, approximately 2000 bonds per molecule of thyroglobulin were cleaved, representing 45% of the total peptide bonds. In 90 minutes about 850 bonds per molecule were cleaved, and by 24 hours about 2100 linkages had been cleaved. Incubation of thyroglobulin with pepsin at pH 6.5 was nearly as effective, resulting in the cleavage of about 2500 peptide bonds in 96 hours.

On dialysis of the Pronase digests there was a rapid loss of the peptide material from inside the dialysis tubing (Table I). After only 6 hours of dialysis, it was possible to obtain a nondialyzable fraction containing less than 4% of the peptide portion of the molecule, but more than 90% of the carbohydrate. With more prolonged dialysis there was a preferential loss of both mannose and glucosamine, which was more striking in the case of the mannose component. More than 50% of the mannose present in the original protein dialyzed out in a period of 8 to 9 days.

<table>
<thead>
<tr>
<th>Length of</th>
<th>Mannose</th>
<th>Glucosamine</th>
<th>Sialic acid</th>
<th>Fucose</th>
<th>Galactose</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>dialysis</td>
<td>mg/g protein digested</td>
<td>mg/g protein</td>
<td>mg/g protein</td>
<td>mg/g protein</td>
<td>mg/g protein</td>
<td>mg/g protein</td>
</tr>
<tr>
<td>hrs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0*</td>
<td>23.2 (100)†</td>
<td>28.0 (100)</td>
<td>13.7 (100)</td>
<td>4.14 (100)</td>
<td>13.4 (100)</td>
<td>920.0 (100)</td>
</tr>
<tr>
<td>6½</td>
<td>21.3 (92)</td>
<td>23.1 (89)</td>
<td>13.5 (99)</td>
<td>3.63 (88)</td>
<td>13.8 (103)</td>
<td>25.0 (2.7)</td>
</tr>
<tr>
<td>23</td>
<td>19.3 (83)</td>
<td>22.0 (85)</td>
<td>13.7 (100)</td>
<td>3.71 (90)</td>
<td>12.7 (95)</td>
<td>37.1 (4.0)</td>
</tr>
<tr>
<td>45</td>
<td>14.8 (64)</td>
<td>21.5 (81)</td>
<td>13.4 (98)</td>
<td>3.63 (88)</td>
<td>11.0 (82)</td>
<td>31.5 (3.4)</td>
</tr>
<tr>
<td>96 5</td>
<td>12.7 (55)</td>
<td>19.6 (75)</td>
<td>12.6 (92)</td>
<td>3.75 (91)</td>
<td>10.2 (76)</td>
<td>21.7 (2.4)</td>
</tr>
<tr>
<td>192</td>
<td>10.0 (45)</td>
<td>17.9 (69)</td>
<td>19.7 (93)</td>
<td>3.34 (81)</td>
<td>11.4 (85)</td>
<td>15.8 (1.7)</td>
</tr>
<tr>
<td>216</td>
<td>10.2 (44)</td>
<td>18.6 (72)</td>
<td>12.6 (92)</td>
<td>3.67 (89)</td>
<td>11.2 (84)</td>
<td>16.3 (1.8)</td>
</tr>
</tbody>
</table>

* Native protein.
† Figures in parentheses refer to the percentage of the component recovered.
In contrast, even after such prolonged dialysis, there were only slight decreases in the amounts of the sialic acid, fucose, and galactose present in the nondialyzable fraction. These results of dialysis suggested the presence of more than one type of carbohydrate unit in the thyroglobulin molecule, with part of the mannose and glucosamine being in a unit different from that which contained the other sugars. Subsequent studies were directed toward achieving a separation of these carbohydrate units.

**Chromatography of Glycopeptides on DEAE-cellulose**—In order to separate the glycopeptide components present in the nondialyzable fraction of the Pronase digest of thyroglobulin, chromatography of this material on DEAE-cellulose columns was performed. In Fig. 1 a representative elution diagram is shown. All of the glycopeptides were eluted when a concentration of approximately 0.014 M phosphate buffer had been reached. Reproducible elution patterns were achieved. Recovery of the carbohydrate material placed on the DEAE-cellulose columns averaged 96% on the basis of the hexose measurements.

The amount of carbohydrate in Peak 1, which represents the material not adsorbed on the DEAE-cellulose column, was inversely related to the length of prior dialysis of the digest, while the amounts present in the other peaks remained essentially constant regardless of the length of the dialysis. After 23 hours of prior dialysis, Peak 1 contained 10.7 mg of hexose per g of thyroglobulin digested; after 48 hours, 6.8 mg of hexose per g; and after 8 days, 2.1 mg per g of thyroglobulin. Since recovery of the carbohydrate material placed on the DEAE-cellulose columns was essentially complete, chromatography of a digest after only brief dialysis (Table I) made it possible to account for more than 90% of the carbohydrate present in the thyroglobulin molecule in the DEAE-cellulose elution peaks.

The carbohydrate composition of the glycopeptides obtained by DEAE-cellulose chromatography, expressed as molar ratios either to glucosamine or fucose, is given in Table II. From these analyses it is evident that two types of carbohydrate units of markedly different composition are present. Peak 1 contained only mannose and glucosamine in significant amounts. The other peaks which were eluted with the buffer gradient contained all of the five sugars present in thyroglobulin. The carbohydrate composition of these peaks was quite similar, except for a progressive increase in the amount of sialic acid (Table II). A small decrease, of questionable significance, in the galactose of the later peaks was also observed.

**Chromatography of Glycopeptides on Charcoal-Celite**—The nondialyzable fraction of the Pronase digests of calf thyroglobulin could also be fractionated into two types of carbohydrate units by charcoal-Celite chromatography. In Table II, the carbohydrate composition of the glycopeptide fractions eluted with varying concentrations of ethanol are presented. The fraction eluted with 17% ethanol contained only mannose and glucosamine and was entirely free of galactose, sialic acid, and fucose. Its composition was similar to that of Peak 1 from the DEAE-cellulose column. The fractions eluted with concentrations of ethanol of 20% and upward had a carbohydrate composition similar to the other DEAE-cellulose peaks. The 20% ethanol fraction represented a mixture of these two types of units. Recovery of the carbohydrate material placed on the charcoal-Celite columns averaged 80% on the basis of hexose measurements when elution was carried up to 50% ethanol. By increasing the ethanol concentration to 70% a small additional amount of glycopeptide material with a higher sialic acid content, similar to that found in the late DEAE-cellulose peaks, could be eluted.

The data in Table II are presented as glycopeptides 1 through 9. Only those peaks which contained at least one significant amount of glucosamine or galactose are numbered. The other peaks which were eluted with the buffer gradient contained all of the five sugars present in thyroglobulin. The carbohydrate composition of the glycopeptides obtained by DEAE-cellulose chromatography, expressed as molar ratios to glucosamine or fucose, is given in Table II. From these analyses it is evident that two types of carbohydrate units of markedly different composition are present. Peak 1 contained only mannose and glucosamine in significant amounts. The other peaks which were eluted with the buffer gradient contained all of the five sugars present in thyroglobulin. The carbohydrate composition of these peaks was quite similar, except for a progressive increase in the amount of sialic acid (Table II). A small decrease, of questionable significance, in the galactose of the later peaks was also observed.

**Chromatography of Glycopeptides on Charcoal-Celite**—The nondialyzable fraction of the Pronase digests of calf thyroglobulin could also be fractionated into two types of carbohydrate units by charcoal-Celite chromatography. In Table II, the carbohydrate composition of the glycopeptide fractions eluted with varying concentrations of ethanol are presented. The fraction eluted with 17% ethanol contained only mannose and glucosamine and was entirely free of galactose, sialic acid, and fucose. Its composition was similar to that of Peak 1 from the DEAE-cellulose column. The fractions eluted with concentrations of ethanol of 20% and upward had a carbohydrate composition similar to the other DEAE-cellulose peaks. The 20% ethanol fraction represented a mixture of these two types of units. Recovery of the carbohydrate material placed on the charcoal-Celite columns averaged 80% on the basis of hexose measurements when elution was carried up to 50% ethanol. By increasing the ethanol concentration to 70% a small additional amount of glycopeptide material with a higher sialic acid content, similar to that found in the late DEAE-cellulose peaks, could be eluted.

The amount of carbohydrate in the 17 and 20% ethanol fractions varied inversely with the length of prior dialysis of the digest, in a manner similar to Peak 1 of the DEAE-cellulose columns. The amount of carbohydrate in the other fractions remained essentially constant regardless of the length of dialysis. Chromatography of Peak 1 from the DEAE-cellulose column on charcoal-Celite showed that it consisted primarily of material eluting with 17% ethanol and contained only small amounts of

![Fig. 1. Elution diagram from DEAE cellulose column chromatography of the nondialyzable fraction, after 48 hours of dialysis, of a Pronase digest from 1385 mg of calf thyroglobulin. Column size was 2.1 X 65 cm. The sample was applied in 0.0005 M sodium phosphate buffer, pH 7.6, and a linear gradient from 0.0005 M to 0.03 M phosphate, pH 7.6, was employed as described in the text. The solid bars indicate the tubes which were pooled from the numbered peaks. These peaks will be referred to in subsequent discussion as glycopeptides 1 through 9.](image-url)
the fractions eluting with higher ethanol concentrations. The highest degree of purification of glycopeptides containing the mannose-glucosamine unit was achieved by a combination of the DEAE-cellulose and charcoal-Celite chromatography, and the material obtained by this combined procedure was used for further study of this carbohydrate unit, including the determination of its molecular weight.

Electrophoresis of Glycopeptides—The glycopeptides obtained by DEAE-cellulose chromatography appeared as single spots on electrophoresis at pH 6.4. At that pH they were all anionic, except for glycopeptide 1 which was neutral. Their relative mobility toward the anode was in the same order as their elution from the DEAE-cellulose column, the glycopeptides which emerged last from the column showing the most rapid migration. For example, in an electrophoretic separation in which leucine migrated 1.3 cm and aspartic acid 14.0 cm toward the anode, glycopeptide 1 moved 1.3 cm and glycopeptide 9, 8.6 cm toward this pole.

Electrophoresis of the mannose-glucosamine glycopeptide obtained in the 17% ethanol fraction after charcoal-Celite chromatography showed a single spot at pH 2.6, 6.4, and 7.8. This glycopeptide migrated rapidly toward the cathode at pH 2.6, while it was neutral at pH 6.4 and 7.8. The 20% ethanol fraction from the charcoal-Celite column showed at pH 6.4 this neutral glycopeptide as well as an anionic component. The fractions eluted in the higher ethanol concentrations (25 to 50%) showed a mixture of anionic components, with migrations similar to those of glycopeptides 2 through 9.

All of the glycopeptides from DEAE-cellulose or charcoal-Celite columns stayed at the origin when chromatographed in the systems used in this study for the resolution of amino acids or sugars (7, 11). However, migration of the glycopeptides was observed in the ethyl acetate-water-pyridine-acetone system (15) in which movement of ovalbumin glycopeptides has also been reported (16).

Molecular Weight Determinations of Glycopeptides—The determinations of the molecular weight of the two types of carbohydrate units present in thyroglobulin were performed on purified glycopeptide fractions containing one or the other of these two units. The mannose-glucosamine unit, which will be referred to as carbohydrate unit A, was purified by a combination of DEAE-cellulose and charcoal-Celite chromatography. For the studies of the second type of unit, carbohydrate unit B, which contains sialic acid, fucose, galactose, mannose, and glucosamine, the nondialyzable fraction after prolonged dialysis (8 to 9 days) was used.

Fig. 2 shows the ultracentrifugal pattern of the glycopeptides from the two different types of units. Both show a single symmetrical peak indicating that the components present in each fraction are of similar molecular weight.

The sedimentation constant, $S_{20,w}$, of the glycopeptides from unit A was 0.67 S for a 0.7% solution, while the sedimentation constant for the glycopeptides from unit B was 0.73 S for a 1.2% solution.

A partial specific volume of 0.634 ml per g was calculated for the glycopeptide fraction of unit A from its composition, with the use of the volumes of the amino acid residues and basing the volumes of the sugar residues on the volumes of their constituent atoms (17). For the glycopeptides of unit B, a partial specific volume of 0.646 ml per g was employed in the molecular weight calculations. This value has been experimentally obtained for the glycopeptides of fetuin (7), which have a composition very similar to that of the glycopeptides from unit B of thyroglobulin.

Molecular weight determinations by sedimentation-equilibrium, with the use of these partial specific volumes, gave a value of 1250 for the glycopeptides of unit A, and 4100 for the glycopeptides of unit B (Table III).

The composition and molecular weight of unit A suggest that there are 5 residues of mannose and 1 of N-acetylglucosamine in each unit. This would represent a unit weight of 1013, which compares favorably with the 1050 obtained from the physical measurements (Table III).
Although there are minor variations in the composition of the units of type B, especially in the sialic acid content (Table II), the average number of residues present can be estimated from the composition of the glycopeptides containing this type of unit resolved by DEAE-cellulose chromatography. This would indicate an average composition of 3 mannose, 5 N-acetylglucosamine, 4 galactose, 2 sialic acid, and 1 fucose residues per unit, which would represent a unit weight of 2977. This is in good agreement with the molecular weight of 3200 determined experimentally for this unit (Table III).

A calculation of the approximate number of each type of carbohydrate unit present in the thyroglobulin molecule can be made from the data obtained in the dialysis studies (Table I) and the separations by gel filtration to be presented. These studies indicate that approximately 52% (12 mg per g) of the mannose in thyroglobulin is present in carbohydrate unit A. This would represent approximately 45 mannose residues per mole of thyroglobulin which is equivalent to 9 carbohydrate units of type A per molecule. By subtracting from the weight of the total carbohydrate present in each thyroglobulin molecule (53,000) the amount of carbohydrate present in all of the units of type A, a value of 43,550 is obtained, which represents the carbohydrate present in the units of type B. Dividing this amount by 3200, which is the average molecular weight of unit B (Table III), gives a value of approximately 14 units of this type per molecule of thyroglobulin.

Fractionation of Glycopeptides by Gel Filtration—Since this study indicated the occurrence in thyroglobulin of two types of carbohydrate units differing in molecular weight, an attempt was made to demonstrate their presence by gel filtration. Briefly dialyzed (6 to 8 hours) Pronase digests, representing over 90% of the carbohydrate of the thyroglobulin molecule, were fractionated on Sephadex G-25 and Sephadex G-50 columns. The recovery of carbohydrate from these columns averaged 96%.

In Figs. 3 and 4 representative elution diagrams for Sephadex G-25 and G-50, respectively, are shown. In both types of column it was apparent from the asymmetry of the plot of the hexose determinations that two incompletely resolved components were present. The sialic acid determinations confirmed this, indicating that this sugar occurred only in the early fractions of eluted material. Analysis of the sugar components of several fractions indicated that those emerging first consisted of sialic acid, fucose, galactose, mannose, and glucosamine, in the same ratios as in the glycopeptides from unit B, while those emerging last consisted solely of mannose and glucosamine in the ratio of 1:1, as in the glycopeptides of unit A. In the middle portion of the elution curve, a mixture of the two types of units was found. Since unit A contains no sialic acid, the plot of the sialic acid values represents the elution diagram for unit B. The elution
Carbohydrate Units of Thyroglobulin

Amino acid composition of glycopeptides from carbohydrate unit B

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Glycopeptide residues/mole glycopeptide</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>1.4</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.54</td>
</tr>
<tr>
<td>Serine</td>
<td>1.1</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>0.30</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.25</td>
</tr>
<tr>
<td>Threonine</td>
<td>0.23</td>
</tr>
<tr>
<td>Proline</td>
<td>0.37</td>
</tr>
<tr>
<td>Valine</td>
<td>0.21</td>
</tr>
<tr>
<td>Leucine</td>
<td>0.11</td>
</tr>
</tbody>
</table>

**NH2-terminal residues of glycopeptides from carbohydrate unit B by dinitrofluorobenzene method**

<table>
<thead>
<tr>
<th>Terminal amino acid</th>
<th>Glycopeptide residues/mole glycopeptide</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>0.19</td>
</tr>
<tr>
<td>Serine</td>
<td>0.23</td>
</tr>
<tr>
<td>Glycine</td>
<td>0.24</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.20</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>0.20</td>
</tr>
<tr>
<td>Valine</td>
<td>0.14</td>
</tr>
<tr>
<td>Leucine</td>
<td>0.10</td>
</tr>
<tr>
<td>Total residues</td>
<td>0.82</td>
</tr>
</tbody>
</table>

* Numbers refer to peaks from DEAE-cellulose column (Fig. 1).

**Table IV**

**Table V**

!!! You can add more content here if needed. The text above is a snapshot of the content and does not contain any images or tables. Please use the provided data to format and present the content as accurately as possible. 

---

Diagram for unit A was obtained by subtracting from the total hexose values of each fraction the hexose contribution of unit B, calculated from the sialic acid to hexose ratio of this unit. The elution patterns of the two units plotted in terms of these calculated hexose values are shown in Figs. 3B and 4B. As would be expected from the molecular weights, the glycopeptides of unit A emerged after those of unit B from both Sephadex columns. Although complete separation of the two types of units was not achieved with either gel, a somewhat better resolution was observed with Sephadex G-50 than with G-25.

Fig. 3 indicates that the glycopeptides from unit B passed through Sephadex G-25 almost unretarded, while those of unit A penetrated the gel to a greater extent. On Sephadex G-50 (Fig. 4) the glycopeptides from both units were retarded to an appreciable extent. The behavior of the glycopeptides is consistent both with their molecular weight (Table III) and the known properties of the Sephadex types employed.

An estimation of the amount of hexose present in both types of carbohydrate units can be made from this type of fractionation. It was calculated from the Sephadex G-50 elution curve (Fig. 4B) that 34% of the total hexose (12.5 mg per g of thyroglobulin) or 54% of the total mannose, is present in the peak containing the glycopeptides of unit A. This is in close agreement with the amount of mannose lost by prolonged dialysis of the Pronase digest (Table I) and the mannose found in the 17% ethanol fraction from the charcoal-Celite chromatography of a briefly dialyzed Pronase digest.

**Peptide Portion of Glycopeptides from Carbohydrate Unit B**—The amino acid composition of the glycopeptides of unit B from the DEAE-cellulose column is given in Table IV. These glycopeptides were quite similar, each having only a limited number of amino acids and an average chain length of 3 to 5 residues. Aspartic acid was the only amino acid present as a full residue or more in all of the glycopeptides. No aromatic, sulfur-containing, or basic amino acids were found in any of these glycopeptides. The non stoichiometric ratios of the amino acids suggest that each peak represents a group of closely related molecules differing slightly in the length and composition of the peptide chain in the vicinity of the carbohydrate-peptide linkage, as has been found in the glycopeptides from fetuin (7). This type of variation is to be expected in view of the broad specificity of the proteolytic enzyme used and the fact that there are 14 carbohydrate units of this type linked to as many different portions of the peptide chains of thyroglobulin.

The glycopeptides had free NH2-terminal groups, as shown in Table V. The total of the terminal residues determined in each case was close to 1 residue per mole of carbohydrate unit, confirming the physical molecular weight measurement of this polysaccharide unit. The appearance of more than one type of terminal amino acid is consistent with the non stoichiometric ratios of the amino acids in these glycopeptides.

Digestion of the individual glycopeptides with carboxypeptidase or leucine aminopeptidase further shortened the peptide chains, releasing substantial amounts of amino acids, including serine and glutamic acid. The release of amino acids by each of these enzymes indicates the presence of free COOH-terminal as well as NH2-terminal amino acids on these glycopeptides.

In order to obtain all of the carbohydrate present in units of type B with a minimum number of amino acids attached, the non dialyzable fraction of the Pronase digest of thyroglobulin obtained after 8 days of dialysis was studied. After selective removal of the sialic acid, this fraction was digested in sequence with leucine aminopeptidase and carboxypeptidase. The amino acid composition per mole of carbohydrate after such treatment was as follows: aspartic acid, 0.91; glycine, 0.32; serine, 0.32; glutamic acid, 0.26; alanine, 0.30; threonine, 0.16; proline, 0.23; valine, 0.23; and leucine, 0.10. It may be noted that aspartic acid is the only amino acid present in an amount close to 1 residue per mole of carbohydrate unit.

**Peptide Portion of Glycopeptides from Carbohydrate Unit A**—The composition of the glycopeptide fraction from carbohydrate unit A purified by DEAE-cellulose and charcoal-Celite chromatography is shown in Table VI. The values both before and after treatment with a combination of leucine aminopeptidase and carboxypeptidase are given. Aspartic acid is again the predominant amino acid present. NH2-terminal analyses by dinitrophenylation indicated the presence of 0.30 moles of glycine, 0.22 moles of alanine, and about 0.1 mole each of aspartic acid and valine.

The presence of less than 1 residue of aspartic acid per mole of glycopeptide and also less than a total of 1 residue of terminal amino acids suggests that Pronase may cause some cleavage of glycopeptide bonds involving this carbohydrate unit. This possibility is supported by the fact that when the glycopeptides...
containing unit A were fractionated on Dowex 50-X2, 200 to 400 mesh (H+ form), the adsorbed material which contained the major part of the carbohydrate had close to 1 residue of aspartic acid per mole of carbohydrate unit, while the carbohydrate material which was not adsorbed by this resin had only trace amounts of amino acids associated with it. It has been observed previously that Pronase and papain may cause some cleaving of the carbohydrate-peptide bonds in orosomucoid (18) and fetuin (19).

**Table VI**

<table>
<thead>
<tr>
<th>Residues per mole of glycopeptide</th>
<th>Residues per mole after leucine aminopeptidase and carboxypeptidase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mannose</td>
<td>5.1</td>
</tr>
<tr>
<td>N-Acetylglucosamine</td>
<td>1.0</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>0.76</td>
</tr>
<tr>
<td>Alanine</td>
<td>0.47</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.36</td>
</tr>
<tr>
<td>Proline</td>
<td>0.33</td>
</tr>
<tr>
<td>Serine</td>
<td>0.17</td>
</tr>
<tr>
<td>Valine</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>N-Acetylglucosamine</strong></td>
<td><strong>1.0</strong></td>
</tr>
<tr>
<td><strong>Aspartic acid</strong></td>
<td><strong>0.71</strong></td>
</tr>
</tbody>
</table>

* Purified by DEAE-cellulose and charcoal-Celite chromatography as described in text.

After extensive digestion of the peptide portion of thyroglobulin, glycopeptides containing these two types of units could be separated by several techniques because of their markedly different properties. The difference in the molecular weight of the two carbohydrate units made it possible to separate the glycopeptides containing them by dialysis as well as by gel filtration. Chromatographic separation by DEAE-cellulose was facilitated by the fact that one unit has negatively charged sialic acid residues, while the other unit contains only neutral sugars.

Separation on charcoal-Celite was made possible by the smaller carbohydrate portion of the mannosé-glucosamine unit and also by the increased affinity which sialic acid-containing polysaccharides have for this adsorbant.

The dialysis studies (Table I) permitted the choice of a dialysis time at which more than 90% of the carbohydrate and yet less than 4% of the peptide present in thyroglobulin could be obtained for further fractionation on DEAE-cellulose, charcoal-

---

* Purified by DEAE-cellulose and charcoal-Celite chromatography as described in text.

---

It is of interest to compare the information concerning the carbohydrate units of calf thyroglobulin with studies on other glycoproteins. Although several types of carbohydrate units have been described, each of the glycoproteins studied previously has been found to contain only a single distinct type of unit. However, evidence for small variations on a basic structural pattern has been reported (18, 20), representing the type of microheterogeneity which apparently exists in the carbohydrate unit B of thyroglobulin in regard to its sialic acid content.

The smaller carbohydrate unit of thyroglobulin, containing only mannose and glucosamine, bears a close resemblance to the carbohydrate unit of ovalbumin in composition as well as molecular weight (21). Its behavior on Sephadex also resembled that of the ovalbumin unit (22). The larger unit, B, has a striking similarity to the carbohydrate units of fetuin, containing the same sugars, except for fucose, and with almost the same molecular weight (7). The glycopeptides from fetuin are similar to those from carbohydrate unit B of thyroglobulin in their non-dialyzability as well as their behavior on DEAE-cellulose chromatography (7).

The complexity of the carbohydrate of thyroglobulin resides not only in the occurrence of two types of carbohydrate units, but also in the presence of a large number of these units in each molecule. The total of approximately 23 units per thyroglobulin molecule may be compared to the single carbohydrate unit present in ovalbumin (21), the 3 units occurring in the fetuin molecule (7), and the 5 to 7 units believed to be present in orosomucoid (18).

A study of the amino acids present in the glycopeptides containing each of the 2 carbohydrate units indicated that only a limited number of amino acid types occurred in the vicinity of the glycopeptide bond despite the fact that there must be approximately 23 different points in the peptide chains at which the carbohydrate units are attached.

The occurrence of free NH₂- and COOH-terminal amino acids in the glycopeptides suggests that the carbohydrate-peptide linkage is through a functional group of an amino acid in the peptide chain. Sodium borohydride reduction of the glycopeptides, followed by acid hydrolysis, showed no chromatographic evidence of hexitols or glucosaminitol, indicating that no free reducing groups are present. This suggests that the carbohydrate-peptide linkage involves C-1 of the terminal sugar as in ovalbumin (16) and fetuin (23).

After exopeptidase digestion of the glycopeptides from both types of units the carbohydrate was obtained with aspartic acid present as the predominant amino acid. The glycopeptides of carbohydrate unit A contained, in addition to aspartic acid, only trace amounts of serine as an amino acid which could be involved in the glycopeptide bond. In the glycopeptides of unit B, in addition to the residue of aspartic acid, small amounts of serine, glutamic acid, and threonine were still present after exopeptidase digestion, and these could be involved in some of the carbohydrate-protein attachments. These findings suggest that aspartic acid is the amino acid primarily involved in the many glycopeptide linkages occurring in the thyroglobulin molecule, as has been reported for several other glycoproteins (16, 24).
SUMMARY

Calf thyroglobulin in its native state was susceptible to extensive proteolytic digestion by the enzyme Pronase. After such treatment, more than 96% of the peptide portion could be removed by brief dialysis.

From such digests the carbohydrate units of thyroglobulin were isolated with only a few amino acids attached. Two distinct types of carbohydrate units differing markedly in size and composition were demonstrated and shown to account for essentially all of the carbohydrate of the thyroglobulin molecule. The glycopeptides containing these two types of units were separated by several techniques, including dialysis and chromatography on DEAE-cellulose or charcoal-Celite columns. In addition, they were partially resolved by gel filtration on Sephadex columns.

One type of unit had a molecular weight of 1050 and consisted of 5 residues of mannose and 1 residue of glucosamine, while the other type of unit had an average molecular weight of 3200 and consisted of sialic acid, fucose, galactose, mannose, and glucosamine. Approximately 9 of the smaller type of unit and 14 of the larger type are present in each thyroglobulin molecule.

A study of the peptide portion of the glycopeptides indicated the presence of amino- and carboxyl-terminal residues. After digestion of the glycopeptides from the two types of carbohydrate units with exopeptidases, the composition of the peptide material remaining attached to the carbohydrate indicated that aspartic acid is the amino acid involved in at least a large number of the approximately 23 glycopeptide linkages of the thyroglobulin molecule.

Acknowledgment—The author would like to thank Miss Margaret Hines for valuable technical assistance.

REFERENCES