Occurrence of Sulfate in the Asparagine-linked Complex Carbohydrate Units of Thyroglobulin

IDENTIFICATION AND LOCALIZATION OF GALACTOSE 3-SULFATE AND N-ACETYLGLUCOSAMINE 6-SULFATE RESIDUES IN THE HUMAN AND CALF PROTEINS*

(Received for publication, February 9, 1988)

Robert G. Spiro$ and Vishnu D. Bhoyroo

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

Human thyroglobulin glycopeptides representing the multiple asparagine-linked complex (unit B) carbohydrate units of this protein were found to contain substantial amounts of sulfate (ranging from 0.5 to 2.5 mol/mol of oligosaccharide); this substituent was shown to occur primarily in the form of terminal β-linked Gal-3-SO4 residues which represent novel capping groups occurring alternatively to sialic acid and in comparable amounts. Upon hydrazine/nitrous acid fragmentation and radiolabeling with NaB³H₄, all human unit B DEAE-resolved glycopeptide fractions yielded an acidic disaccharide which was characterized as Gal-3-SO₄β1→4-anhydromannitol. Studies on glycopeptides modified by desialylation, desulfation, and β-galactosidase treatment indicated that the majority (~70%) of the complex carbohydrate units contain sulfate groups and that Gal-3-SO₄ and sialic acid residues can coexist in terminal positions on the same N-linked oligosaccharide. In addition to Gal-3-SO₄, the most acidic unit B variants were found to contain GlcNAc-6-SO₄ which was recovered as Galβ1→4-anhydromannitol-6-SO₄ after hydrazine/nitrous acid treatment and NaB³H₄ reduction. On the basis of chromatography on immobilized concanavalin A, it was determined that whereas the Gal-3-SO₄ groups occur on biantennary as well as more highly branched carbohydrate units, GlcNAc-6-SO₄ is exclusively present in the latter oligosaccharides. In contrast to the N-linked carbohydrate units, the previously described O-linked glycosaminoglycan chain of human thyroglobulin yielded GlcAβ1→3-anhydrotalitol-6-SO₄ upon hydrazine/nitrous acid/NaB³H₄ treatment, indicating that it is a chondroitin 6-sulfate-like polymer.

The distribution of sulfate in the complex oligosaccharides of calf thyroglobulin was quite different from that in the human protein; sulfate was not detectable in most of the glycopeptides and was sequenced in a single multibranched complex-type glycopeptide fraction (1.6 mol of sulfate/mol of oligosaccharide) which contained about equal amounts of Gal-3-SO₄ and GlcNAc-6-SO₄. The difference in galactose sulfation between human and calf thyroglobulins may be related to the substitution in the latter protein of some of the galactose residues by α-D-Gal capping groups.

This work was supported by Grants DK 17325 and DK 17477 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

$ To whom correspondence should be addressed: Elliott P. Joslin Research Lab., One Joslin Place, Boston, MA 02215.

Thyroglobulin, the major secretory product of thyroid follicular cells, is known to undergo a number of co- and post-translational modifications (1) which include oligosaccharide attachment and processing (2, 3), phosphorylation of amino acid and sugar residues (4-6), as well as the iodination and coupling steps which lead to the formation of the peptide-linked tri- and tetraiodothyronines (7, 8). Although all thyroglobulins which have so far been examined contain carbohydrate, predominantly in the form of asparagine-linked polysaccharide (unit A) and complex (unit B) oligosaccharides (9, 10), distinguishing species differences have been noted (10-12), among which is the occurrence in the human protein of a sulfated O-linked glycosaminoglycan chain (13).

This report further delineates the complexity of the human thyroglobulin molecule by demonstrating that the numerous N-linked complex-type carbohydrate units of this protein (22/ molecule, 11/330-kDa subunit) are also extensively sulfated, primarily in the form of a previously unrecognized capping group consisting of Gal-3-SO₄ in β1→4 linkage to GlcNAc. Furthermore, we find that some of the unit B variants also contain more internally located sulfate substituents in a Galβ1→4GlcNAc-6-SO₄ sequence. An examination of thyroglobulin from calf glands indicated that sulfation of its N-linked carbohydrate also occurs but is less widespread than in the human protein, being confined to a multiply branched oligosaccharide in which about equimolar amounts of Gal-3-SO₄ and GlcNAc-6-SO₄ occur.

EXPERIMENTAL PROCEDURES

Preparation of Thyroglobulins—Human thyroids from individuals without known thyroid disease were obtained at autopsy from the Brigham and Women’s Hospital, whereas calf thyroids were obtained from a local slaughterhouse. All glands were stored frozen at −20 °C until they were used for the preparation of thyroglobulin by previously described procedures (14).

Preparation of Glycopeptides—Glycopeptides containing the complex carbohydrate units (unit B) were prepared from Pronase digests of human and calf thyroglobulins by DEAE-cellulose chromatography of the Sephadex G-25-excluded material which remained unadsorbed by Dowex 50-X2 at pH 3.0 (10). The human glycopeptides were resolved as previously described (10) on a 2.1 X 75-cm DE52 cellulose column with a sodium phosphate, pH 7.6, concentration gradient starting at 0.5 mM. During this chromatography, unit B glycopeptides were eluted as peaks at the following buffer concentrations: peak 3, 7 mM; peak 4, 9 mM; peak 5, 11 mM; peak 6, 12 mM; peak 8, 13 mM; peak 9, 14 mM; peak 10, 15 mM; peak 11, 16 mM; peak 12, 17 mM; peak 13, 19 mM; and peak 14, 21 mM. An additional glycopeptide fraction (peak 15) was eluted with 200 mM sodium phosphate buffer at pH 7.6. Peaks 1 and 2 contained unit A oligosaccharide and a hybrid glycopeptide, respectively.

Calf thyroglobulin glycopeptides (peaks 2–8) were resolved as previously described (12) by DEAE-cellulose chromatography with 20–200 mM pyridine acetate gradient at pH 5.5, and an additional
glycopeptide was fraction-eluted with 400 mM pyridine acetate at this pH (peak 9).

Glycopeptides containing the O-linked glycosaminoglycan chain from human thyroglobulin were prepared by elution from DEAE-cellulose with 2 mM pyridine acetate, pH 5.1, as previously described (13).

Radiolabeling of Glycopeptides—Reductive alkylation of glycopeptides with [14C]formaldehyde (52 mCi/mmol; Du Pont New England Nuclear) was accomplished as previously reported (12); the radiolabeled glycopeptides were separated from excess reagent by Bio-Gel P-2 filtration and were found to have average specific activities of 1.2 × 10³ dpm/nmol.

Hydrazine/Nitrous Acid Treatment of Glycopeptides and NaBH₄ Reduction of Saccharide Products—To achieve de-N-acetylation of glucosamine residues glycopeptides (20-200 nmol) were heated with freshly redistilled hydrazine (0.5 ml) in sealed tubes under nitrogen for 16 h at 100 °C as previously described (15). After removal of the reagent in vacuo with several additions of toluene and vacuum desiccation over an H₂SO₄ trap, deamination was carried for 2 h at room temperature by the addition of 0.5 ml of 0.2 M NaNO₂ in 0.4 N acetic acid. The saccharide products of this reaction, after passage through columns of Dowex 50-X₄ (H⁺ form), were then reduced with 3 mCi of NaBH₄ (53 Ci/mmol), whereas any remaining glucosamine was quantitated with 0.1 M NaOH for 3 h at room temperature. After decomposition of the borohydride with acetic acid, the samples were passed through columns of Dowex 50-X₄ (H⁺ form), followed by lyophilization and removal of boric acid as methyl borate. Reduction of an N-acetylg glucosamine residue with sodium phosphide produced a product with an average specific activity of 7.5 × 10⁴ dpm/nmol. When the hydrazine/nitrous acid/NaBH₄ procedure was applied to fetuin glycopeptides (20-200 nmol) were heated with galactose oxidase (4 units; Sigma, Type VI) at 37 °C for 40 h in 250 μl of 0.15 M sodium acetate buffer, pH 5.0, in the presence of toluene. The digestions were terminated by heating, and the samples were desalted by passage through coupled columns of Dowex 50-X₄ (200-400 mesh, H⁺ form) and Dowex 1-X₂ (200-400 mesh, acetate form). Alkaline glycopeptides, after desulfation (85 nmol mixed with corresponding [3H]-labeled tracer glycopeptides), were digested with jack bean β-galactosidase (0.25 unit; Sigma, Type VII) at 37 °C for 40 h in 250 μl of 0.15 M sodium acetate buffer, pH 5.0, in the presence of toluene. Subsequently, the glycopeptides were separated from salt and released galactose by filtration on Bio-Gel P-2 equilibrated with 0.1 M pyridine acetate, pH 5.0.

Glycopeptides were also treated with Aspergillus niger β-N-acetylglucosaminidase (0.25 unit; Sigma) in 0.15 M sodium citrate buffer, pH 5.0, in a manner similar to that employed for the β-galactosidase digestion.

Galactose Oxidase/NaBH₄ Treatment—Acidic asialoglycopeptides made neutral by desulfation (25 nmol) as well as Galβ1-3GlcNAcβ1-4Manβ1-3Galβ1-4GlcNAcβ1-O-Me and Galβ1-4GlcNAcβ1-O-Et (purchased from Stockerbolaget) were incubated with galactose oxidase (4 units; Sigma, Type V) at 37 °C for 16 h in 140 μl of 0.1 M sodium phosphate buffer at pH 7.0 in the presence of toluene. The digests were terminated by heating, and the oxidized saccharides were reduced for 2 h by the addition of 3 mCi of NaBH₄ (53 Ci/mmol; Du Pont New England Nuclear) and the HCl was removed by lyophilization. The samples were then acidified with acetic acid, and the glycopeptides were desalted by filtration on Bio-Gel P-2 equilibrated with 0.1 M pyridine acetate, pH 5.0. They were then further purified by chromatography on paper for 16 h in Solvent System B (for definition, see "Chromatographic Procedures"), where they remained at the origin from which they were eluted with water. The alkyl glycodies, on the other hand, were passed through coupled columns of Dowex 50 (H⁺ form) and Dowex 1 (acetate form) from which they were recovered with a 26% ethanol wash; after drying of the samples in a vacuum rotator, the 'boric acid was eluted with 40% pyridine acetate, further resolution was achieved by thin-layer chromatography.

Desulfation—To remove sulfate, lyophilized glucoproteins (25 nmol) and NaBH₄-labeled oligosaccharides (5 × 10⁴ dpm), which had been placed into the pyridinium form by elution from Dowex 1 with 1.0 M pyridine acetate, pH 5.0, were heated at 100 °C for 6 h with 0.3 ml of a dimethyl sulfoxide reagent containing 10% methanol titrated with HCl to pH 4.0 (17). The reaction was terminated by adding 1 eq at NH₂OH to the cooled sample and then applying it to a DE52 cellulose column (2.5 ml) equilibrated with 5 mM pyridine acetate, pH 5.0. The neutral products were recovered in the effluent and washed from the acidic component by chromatography on a column (0.8 × 5 cm) of Dowex 1-X₂, 200-400 mesh (acetate), by an elution scheme which was modified from a previous report (15) and consisted in sequence of 35 ml of water (fraction 1); 45 ml of 1 N formic acid (fraction 2); 35 ml of 0.7 M pyridine acetate, pH 5.0 (fraction 3); and 35 ml of 2 M pyridine acetate, pH 5.0 (fraction 4). This procedure permitted the separation of neutral and acidic containing saccharides, present in fractions 1 and 2, respectively, from monosulfated components containing either neutral sugars (fraction 3) or uronic acids (fraction 4); any desulfated saccharide species would elute in fraction 4 (15). After removal of solvent by lyophilization, further resolution was achieved by thin-layer chromatography.

Periodate Oxidation—The NaBH₄-labeled sulfated and desulfated saccharides (~50,000 dpm) were oxidized in 100 μl of 0.1 M sodium metaperiodate in 0.04 M sodium acetate buffer, pH 4.5, at 4 °C for 16 h (15). The reaction was terminated by the addition of a 2-fold molar excess of ethylene glycol; and after 1 h at room temperature, the samples were brought to pH 10 with sodium borate buffer and reduced with 30 μmol of NaBH₄, for 2 h. After decomposition of the NaBH₄ with acetic acid, neutral saccharide samples were desalted by passage through coupled columns of Dowex 50 (H⁺ form) and Dowex 1 (acetate form) from which they were recovered with a 26% ethanol wash; after drying of the samples in a vacuum rotator, the 'boric acid was eluted with 40% pyridine acetate.
room temperature on columns (0.7 x 5.0 cm) of concanavalin A-Sepharose (Pharmacia LKB Biotechnology Inc.) equilibrated with 0.05 M Tris-HCl buffer, pH 7.8, containing 0.1 M sodium chloride, 1 mM calcium chloride, 1 mM magnesium chloride, and 0.02% sodium azide. After application, the samples were allowed to interact with the immobilized lectin for 30 min prior to a 10-mL buffer wash; the bound glycopeptides were then eluted with buffer containing 10 mM methyl-α-D-mannoside (10 mL). For further study, the unbound as well as bound material was concentrated by lyophilization and separated from salt and glycoside by filtration on columns of Bio-Gel P-2 equilibrated with 0.1 mM pyridine acetate, pH 5.0.

Preparation of Radiolabeled Sulfate Standards—For the preparation of [3H]Ga1f1—3AnMan and [3H]Ga1f1—4AnMan, galactose oxidase/NabH4, treatment Ga1f1—3GlcNAc-2-O-Me and Ga1f1—4GlcNAc-0-Et, respectively, were submitted to hydrazinolysis for 72 h at 100 °C. Subsequently, the samples were treated with nitrous acid, followed by preparative thin-layer chromatography in Solvent System B. The preparation of [3H]labeled Ga1f1—4AnManH2, Ga1f1—4AnManH2-6-SO4, GlcA1f1—5αTalH2-4-SO4, and GlcA1f1—3AnTalH2-6-SO4, was accomplished by hydrazine/nitrous acid/NabH4 treatment of asialofetuin, keratan sulfate, chondroitin 4-sulfate and accounted for about 10% of the total unit B hexose.

Table I

<table>
<thead>
<tr>
<th>Glycopeptide</th>
<th>Sulfate</th>
<th>Sialic acid</th>
<th>Galactose</th>
<th>Sulfate + sialic acid/ galactose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycopeptide</td>
<td>m mol/mol glycopeptide</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (12)</td>
<td>0.5</td>
<td>0.6</td>
<td>2.3</td>
<td>0.6</td>
</tr>
<tr>
<td>4 (14)</td>
<td>0.7</td>
<td>1.0</td>
<td>2.2</td>
<td>0.8</td>
</tr>
<tr>
<td>6 (6)</td>
<td>0.8</td>
<td>1.4</td>
<td>2.1</td>
<td>1.0</td>
</tr>
<tr>
<td>7 (7)</td>
<td>0.5</td>
<td>0.5</td>
<td>2.3</td>
<td>0.8</td>
</tr>
<tr>
<td>8 (15)</td>
<td>1.1</td>
<td>1.4</td>
<td>2.4</td>
<td>1.0</td>
</tr>
<tr>
<td>9 (6)</td>
<td>0.7</td>
<td>1.3</td>
<td>2.4</td>
<td>0.8</td>
</tr>
<tr>
<td>11 (7)</td>
<td>1.5</td>
<td>1.0</td>
<td>2.3</td>
<td>1.1</td>
</tr>
<tr>
<td>13 (7)</td>
<td>1.7</td>
<td>1.1</td>
<td>2.8</td>
<td>1.0</td>
</tr>
<tr>
<td>14 (4)</td>
<td>1.6</td>
<td>1.4</td>
<td>2.9</td>
<td>1.0</td>
</tr>
<tr>
<td>15 (15)</td>
<td>2.5</td>
<td>1.5</td>
<td>4.2</td>
<td>1.0</td>
</tr>
</tbody>
</table>

* Glycopeptide fractions containing complex (unit B) carbohydrate units were prepared by DEAE-cellulose chromatography as previously reported (10) and described under "Experimental Procedures." Fractions 1 and 2, which contain unit A (polymannose) oligosaccharides and an uncharacterized glycopeptide, respectively, are not listed; and minor peaks 5, 10, and 12 were not analyzed.

† Values are expressed on the basis of 3 mannose residues/carbohydrate unit; the complete sugar and amino acid compositions of these glycopeptides have been previously reported (10).

§ Values in parentheses represent the molar percentage of total human unit B carbohydrate present in each glycopeptide fraction.

The calf protein, was studied in detail.

Nature of Sulfated Disaccharides Released by Hydrazine/Nitrous Acid Treatment of Unit B Glycopeptides—Thin-layer chromatography of the NabH4-reduced acidic products (Dowex 1 fraction 3) from hydrazine/nitrous acid-treated human thyroglobulin glycopeptides revealed the presence of a major component (component B in Fig. 1) which migrated as a disaccharide with an Rf of 1.83 and 1.72 in Solvent Systems A and B, respectively (Fig. 1). The unfractionated unit B glycopeptides (lane IN) furthermore contained smaller amounts of a slower disaccharide, A, which was recovered exclusively in glycopeptide 15 (Fig. 1) and which migrated with an Rf of 1.41 (Solvent System A) and 1.17 (Solvent System B). A similar chromatographic pattern was observed when the oligosaccharides released by hydrazine from the unit B glycopeptides and isolated by Bio-Gel P-6 filtration after re-N-acetylation were submitted to the hydrazine/nitrous acid/NabH4 procedure (data not shown).

The sulfated calf thyroglobulin glycopeptide fraction was also observed to contain both components, but in quite different relative amounts (Fig. 1), whereas the ratio of disaccharide B to A in human glycopeptide 15 was 5.3, it was 0.85 in the calf carbohydrate unit. Disaccharides B and A were clearly distinct from the product which was produced by hydrazine/nitrous acid treatment of the O-linked sulfated glycosaminoglycan chain of human thyroglobulin (15) which would be expected to elute in Dowex 1 fraction 4. Indeed, treatment of unfractionated human thyroid glycopeptides, as well as the purified glycosaminoglycan-peptide, by hydrazine/nitrous acid yielded, after NabH4 reduction, a component present in this Dowex 1 fraction which migrated (Rf = 0.52) to the position of standard GlcA1f1—3AnTalH2-6-SO4 in Solvent System B and remained close to the origin in Solvent System A (Fig. 1, lane GAG). In contrast, no oligosaccharides were observed upon thin-layer chromatography of Dowex 1 fraction 4 from the human unit B glycopeptides, indicating the absence of monosulfated disaccharides containing uronic acids or disulfated disaccharides containing uronic acids or neutral sugars (data not shown).
Since the sulfated uronic acid-containing disaccharide (Fig. 1, lane GAG) co-migrated in Solvent System B with GlcA(1→3)AnTalH2-6-SO4 and moved at an appreciably slower rate than the GlcA(1→3)AnTalH2-4-SO4 standard, which had an \( R_{gel} \) of 0.75, it is evident that the human glycosaminoglycan chain consists of chondroitin 6-sulfate-like repeating disaccharide units.

To establish that the acidic substituent in the unit B-derived disaccharides is indeed sulfate, the components in Dowex 1 fraction 3 from the unfractionated glycopeptides (Fig. 1, lane IN) were alternatively digested with alkaline phosphatase or submitted to solvolytic desulfation. Thin-layer chromatography in Solvent System C, in which disaccharides A and B are unresolved, indicated that only the latter treatment brought about a change in migration, with the appearance of a saccharide which migrated to the position of the Gal(β1→4)AnManH2 standard (Fig. 2).

**Structure of Sulfated Disaccharides**—Structural studies were carried out on purified disaccharides A and B which migrated as distinct components in their native state (disaccharide A co-migrated with standard Gal(β1→4)AnManH2-6-SO4 from keratan sulfate) but were converted after sulfate removal to a common disaccharide (Fig. 3). Acid hydrolysis (1 N HCl, 4 h, 100 °C) of both disaccharides yielded a single radiolabeled component which was identified as AnManH2 (\( R_{oct} \) = 2.93) by chromatography in Solvent System B (data not shown), in which this monosaccharide is clearly separated from AnTalH2 (\( R_{oct} \) = 2.66).

Although disaccharides A and B were resistant to the action of β-galactosidase, they were readily cleaved by this enzyme after desulfation, yielding, in both instances, a radiolabeled component which migrated to the position of AnManH2 (Fig. 4). On the other hand, β-N-acetylglucosaminidase treatment did not bring about hydrolysis of the sulfate-free disaccharides (data not shown).

Periodate oxidation studies provided information pertaining to the location of the sulfate substituents on the two disaccharides. Chromatographic examination of radiolabeled products indicated that disaccharide A was susceptible to Smith periodate degradation both before and after sulfate removal, yielding AnManH2-6-SO4 (\( R_{oct} \) = 2.06) and An-
The disaccharides (1 × 10^6 dpm) derived from hydrazine/nitrous acid/NaB₃H₄ treatment of unit B glycopeptides were subjected to thin-layer chromatography for 40 h on a cellulose-coated plate in Solvent System A before (−) and after (+) desulfation (deS) by solvolysis as described under "Experimental Procedures." The radiolabeled components were visualized by fluorography; the positions to which standard lactitol without acid/NaB₃H₄ treatment of unit B glycopeptides were subjected to migration of purified disaccharides A and B (Fig. 5). It should be noted that in Solvent System A before desulfation, was readily degraded with the formation of disaccharides A and B (Fig. 5), On the other hand, disaccharide (−) was resistant to periodate oxidation in its native form, but after Smith periodate degradation (NaIO₄) as described under "Experimental Procedures." Chromatography was carried out for 3 h on a cellulose-coated plate in Solvent System B, and the radiolabeled components were detected by fluorography. The positions of migration of standard lactitol (arrow), AnManH₂ (aM), and AnManH₂-6-SO₄ (aM(6S)) are indicated. In this system, standard AnManH₂-3-SO₄ and AnManH₂-4-SO₄ co-migrate with an R_Lact of 2.41 and are clearly resolved from AnManH₂-6-SO₄, which moves with an R_Lact of 2.06; however, separation of native and desulfated disaccharides B is not achieved.

ManH₂, respectively (Fig. 5). On the other hand, disaccharide B was resistant to periodate oxidation in its native form, but after desulfation, was readily degraded with the formation of AnManH₂ (Fig. 5). It should be noted that in Solvent System B (Fig. 5), the native and desulfated forms of disaccharide B are not resolved.

Although the above data indicate that the structures of disaccharides A and B are Galβ₁→4AnManH₂-6-SO₄ and Galβ₁→3AnManH₂-6-SO₄, respectively, a distinction between a galactosyl linkage to C-3 or C-4 of AnManH₂ could not be made because of the periodate resistance which either substitution would confer on this anhydrohexitol. Furthermore, because of the symmetry of the reduced anhydromannoside residue, Galβ₁→3GlcNac and Galβ₁→4GlcNac sequences would be expected to yield the identical Galβ₁→3AnManH₂ disaccharide after hydrazine/nitrous acid/NaB₃H₄ treatment. To circumvent this difficulty, acidic asialoglycopeptides from human thyroglobulin were desulfated and radiolabeled with galactose oxidase NaB₃H₄ prior to hydrazine/nitrous acid degradation. When the unreduced products of this treatment were examined by thin-layer chromatography, a radioactive component was observed which migrated to the position of standard Galβ₁→4AnMan and was clearly distinguished from the more rapidly moving Galβ₁→3AnMan isomer (Fig. 6). After NaBH₄ reduction, as anticipated, the [³H]galactose-labeled disaccharide standards as well as the thyroglobulin
14356 Sulfated N-Linked Oligosaccharides

Gal the neutral aloid fractions of human thyroglobulin unit B glycopeptide 3 (Table I), hydrazine/nitrous acid/NaB₃H₄ treatment of native and various asi-

obtained after sialic removal as well as the latter fraction after

rides released by hydrazine/nitrous acid fragmentation. After
desulfation alone

Gal\(\beta_3\)\(\alpha\)Man\(\beta\)H₂ \((\text{Gal}(\beta_3)-\alpha\text{Man})\),

bution of sulfated and neutral Gal\(\beta_3\)GlcNAc sequences in

acid/NaB₃H₄ treatment of concanavalin A (ConA)-bound

valin-Sepharose chromatography. Following hydrazine/nitrous

acid/NaB₃H₄ treatment of bound and unbound glycopeptides as described under “Experimental Procedures.”

Sulfated refers to the distribution of the Gal-3-\(\text{SO}_4\)\(\beta\)1→4\(\text{AnManH}_2\) unit as measured by hydrazine/nitrous acid/NaB₃H₄ treatment of bound and unbound glycopeptides as described under “Experimental Procedures.”

Total was determined with \(^{14}\text{C}-\text{labeled glycopeptides; average recovery from columns was 88%}.

The Gal\(\beta\)1→4\(\text{AnManH}_2\)6-\(\text{SO}_4\) disaccharide unit which is present in these glycopeptide fractions (see Fig. 1) was found solely in the unbound material.

component migrated to the same position, which moreover coincided with that of the desulfated disaccharide in which the radiolabel was introduced into the anhydromannitol by NaB₃H₄ treatment (Fig. 6).

Localization of Sulfated Saccharides in Carbohydrate Units—After digestion of several human glycopeptide peaks with \(C.\ perfringens\) neuraminidase, the acidic character was retained by a substantial portion as judged by DEAE-cellulose chromatography (peak 3, 81%; peak 4, 54%; peak 8, 72%; and peak 13, 90%). Furthermore after such treatment of the total human unit B glycopeptides, 69% were still retained by the anion exchange. Similar results were obtained when sialic acid removal was achieved by mild acid hydrolysis under conditions in which no desulfation took place. Together these data indicated that most of the human unit B carbohydrate units contain one or more sulfated saccharide residues. Since the analyses of glycopeptide peaks 8 and 13 suggested that the galactose is completely substituted by either sialic acid or sulfate residues (Table I), the retention of acidity noted after desialylation was so great as to suggest the presence of both of these capping groups on the same carbohydrate unit. In these glycopeptide fractions, if the sialic acid and sulfate had been sequenced on distinct oligosaccharides, only 42 and 61% should have remained acidic after desialylation, respectively.

Information about the localization of the Gal-3-\(\text{SO}_4\) group on the human saccharide chains was obtained by chromatographic examination of the disaccharide products resulting from hydrazine/nitrous acid treatment of a unit B glycopeptide fraction after sequential modification of its carbohydrate moiety. The native glycopeptide yielded the Gal-3-\(\text{SO}_4\)\(\beta\)1→4\(\text{AnManH}_2\) disaccharide as well as Gal\(\beta\)1→4\(\text{AnManH}_2\) itself, the latter arising from unsubstituted galactose residues, whereas the neutral glycopeptide fraction (19% of total) obtained after sialic acid removal contained exclusively the unsulfated disaccharide, as would be expected (Fig. 7).

More importantly, the acidic asialoglycopeptide gave rise to both Gal\(\beta\)1→4\(\text{AnManH}_2\) and Gal-3-\(\text{SO}_4\)\(\beta\)1→4\(\text{AnManH}_2\) in a ratio of 0.66, indicating that siaIyated and/or unsubstituted galac-
tose residues do coexist with Gal-3-\(\text{SO}_4\) on the same carbohydrate unit. Following desulfation, the acidic asialoglycopep-

![Fig. 7. Effect of sequential modification of human thyro-
globulin unit B glycopeptide on neutral and acidic disaccha-
rades released by hydrazine/nitrous acid fragmentation. After

hydrolyzine/nitrous acid/NaB₃H₄ treatment of native and various asi-
alo fractions of human thyroglobulin unit B glycopeptide 3 (Table I), the neutral (lanes N) and acidic (lanes A) saccharide fractions (Dowex 1 fractions 1 and 3, respectively) were applied to a cellulose-coated plate. The asialoglycopeptides include the neutral and acidic fractions obtained after sialic removal as well as the latter fraction after desulfation alone (des) and after sequential desulfation and 3-galactosidase treatment (des/Gal). Chromatography was carried out for 53 h in Solvent System A, and the components were visualized by fluorography. The identities of the neutral and acidic components, Gal\(\beta\)1→4\(\text{AnManH}_2\) (Gal-\(\alpha\)M) and Gal-3-\(\text{SO}_4\)\(\beta\)1→4\(\text{AnManH}_2\) (Gal(3S)-\(\alpha\)M), respectively, are indicated.

![Fig. 8. Thin-layer chromatographic assessment of distribu-
tion of sulfated and neutral Gal\(\beta\)1→4\(\text{GlcNAc}\) sequences in human thyroglobulin unit B glycopeptides after concana-
alain-Sepharose chromatography. Following hydrazine/nitrous acid/NaB₃H₄ treatment of concanavalin A (ConA)-bound (+) and unbound (−) unit B glycopeptides, the neutral (lanes N) and acidic (lanes A) saccharide fractions (Dowex 1 fractions 1 and 3, respectively) were applied to a cellulose-coated plate. Chromatography was performed for 40 h in Solvent System A, and the radioactive components were detected by fluorography. The identities of the various disaccharides, Gal\(\beta\)1→4\(\text{AnManH}_2\) (Gal-\(\alpha\)M), Gal-3-\(\text{SO}_4\)\(\beta\)1→4\(\text{AnManH}_2\) (Gal(3S)-\(\alpha\)M), and Gal\(\beta\)1→4\(\text{AnManH}_2\)6-\(\text{SO}_4\) (Gal-
\(\alpha\)M(6S)), are indicated. The unfraccionated human unit B glycopeptides (Total Glycopept) and glycopeptide 3 (Glycopept 3) (Table I) were examined.

<table>
<thead>
<tr>
<th>Glycopeptide</th>
<th>Glycopeptide bound to concanavalin A°</th>
<th>Sulfated°</th>
<th>Total°</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>69°</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>63</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>58</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>60</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>16°</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>

° Glycopeptides containing complex (unit B) oligosaccharides; individual DEAE-cellulose peaks (see Table I) and unfraccionated unit B glycopeptides (initial) were analyzed.

° Concanavalin-Sepharose chromatography was carried out as described under “Experimental Procedures.”

° Total was determined with \(^{14}\text{C}-\text{labeled glycopeptides; average recovery from columns was 88%}.

° The Gal\(\beta\)1→4\(\text{AnManH}_2\)6-\(\text{SO}_4\) disaccharide unit which is present in these glycopeptide fractions (see Fig. 1) was found solely in the unbound material.
tide was converted to neutrality and yielded Galβ1→4AnManH₂ as the only disaccharide (Fig. 7). Upon treatment of this desulfated glycopeptide with β-galactosidase, the disaccharide could no longer be observed after hydrazine/nitrous acid treatment (Fig. 7), clearly indicating that the Gal-3-SO₄ residues are located in terminal positions.

In order to explore the distribution of sulfated saccharides between biantennary and more highly branched carbohydrate units, several unit B glycopeptide fractions were submitted to concanavalin A-Sepharose affinity chromatography and then analyzed after hydrazine/nitrous acid treatment. Chromatographic examination indicated that both Gal-3-SO₄β1→4AnManH₂ and Galβ1→4AnManH₂ were present in the bound as well as the unbound concanavalin A fractions (Fig. 8); and upon quantitation, it became evident that in each case the sulfated carbohydrate interacted with this lectin to a similar extent as the total carbohydrate (Table II). In contrast, glycopeptides containing the Galβ1→4AnManH₂-6-SO₄ sequence were not bound to concanavalin A, as noted when the entire human unit B glycopeptide fraction was examined (Fig. 8). Indeed glycopeptide 15, in which this disaccharide unit is primarily bound in human thyroglobulin (Fig. 1), was retained only to a slight extent by the lectin column (Table II). A similar observation was made with the sulfated unit B glycopeptide from the calf protein in which Galβ1→4AnManH₂-6-SO₄ is present in substantial amounts (Fig. 1); only 15% of this glycopeptide was bound upon concanavalin A chromatography.

**DISCUSSION**

This study, which relied to a large measure on a characterization of the NaB₃H₄-reduced products of hydrazine/nitrous acid fragmentation (15), as summarized in Table III, has demonstrated that the complex (unit B) carbohydrate units of human thyroglobulin are extensively sulfated and that this substituent occurs primarily in the form of terminal β-linked Gal-3-SO₄ residues which represent capping groups not previously described on N-linked oligosaccharides. The sulfated galactose residues were found to be present in human thyroglobulin in amounts similar to sialic acid and appear to represent a structural alternative to the latter sugar as an acidic chain-terminating constituent. An examination of the DEAE-cellulose-resolved glycopeptide fractions showed that Gal-3-SO₄ groups are distributed among most of the multiple complex-type oligosaccharides (11/330,000-kDa subunit) of the human thyroglobulin molecule, that they are present on biantennary as well as more highly branched structures, and that they can coexist with sialic acid on the same carbohydrate unit. Although extensive variation is apparent in the degree to which terminal galactose becomes sulfated, sialylated, or, in some instances, remains unsubstituted, the scope of the heterogeneity at each one of the 11 unit B glycosylation sites remains to be delineated.

In addition to the predominant Gal-3-SO₄ residues, a smaller number of GlcNAc-6-SO₄ groups were observed in the human thyroglobulin unit B glycopeptides; both sulfated monosaccharides were found to be located in N-acetylgalactosamine sequences, appearing after hydrazine/nitrous acid treatment as distinct acidic disaccharides which could be converted to the same neutral component after desulfation. In contrast to the Gal-3-SO₄β1→4GlcNAc sequence which was observed in all glycopeptides, Galβ1→4GlcNAc-6-SO₄ was only detected in the most acidic fraction (Table I, peak 15) which accounted for 15% of the unit B glycopeptides and therefore no more than two complex N-linked oligosaccharides/human thyroglobulin 330-kDa subunit. Studies with concanavalin A indicated that the Galβ1→4GlcNAc-6-SO₄ sequence is exclusively found in oligosaccharides which contain three or more branches. Although Gal-3-SO₄ groups were found in glycopeptides containing GlcNAc-6-SO₄ residues, it is unlikely that both sulfated sugars occur in the same Galβ1→4GlcNAc sequence as no disulfated disaccharides were detected in the hydrazine/nitrous acid products. The relatively high galactose content of the glycopeptide containing GlcNAc-6-SO₄ (peak 15) suggests the presence of an internally located N-acetyllactosamine sequence which becomes sulfated on the GlcNAc residue.

Our studies confirm the occurrence in human thyroglobulin of a distinct sulfated O-linked glycosaminoglycan chain consisting of repeating glucuronic acid/N-acetylgalactosamine units (13) which is retained on DEAE-cellulose at salt concentrations which elute all of the unit B-containing glycopeptides. Hydrazine/nitrous acid treatment of unfractionated human thyroglobulin glycopeptides as well as the purified peptide-linked glycosaminoglycan chain yielded a single NaB₃H₄-reduced sulfated glucuronic acid-containing disaccharide, the identity of which (GlcAβ1→3AnTalH₂-6-SO₄) indicated the existence of a chondroitin 6-sulfate-type sequence.

The distribution of sulfate groups in the complex oligosaccharides of calf thyroglobulin was found to be quite different from that occurring in the thyroglobulin isolated from human thyroid glands. In the calf protein, the sulfate was sequestered in a single late-eluting glycopeptide fraction which accounted for only a small portion (<10%) of the total unit B carbohydrate. These sulfated glycopeptides contained the Gal-3-SO₄β1→4GlcNAc and Galβ1→4GlcNAc-6-SO₄ sequences in about equivalent amounts and, like the human GlcNAc-6-SO₄-containing components, were unsorbed by concanavalin A. The limited distribution of sulfate in calf thyroglobulin is supported by [³⁵S]sulfate labeling studies which indicate that sulfated carbohydrate, in contrast to total unit B

**TABLE III**

*Preparation of radiolabeled sulfated disaccharides from human thyroglobulin glycopeptides by the hydrazine/nitrous acid/NaB₃H₄ procedure*

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Disaccharide B</th>
<th>Disaccharide A</th>
<th>Glycosaminoglycan chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated</td>
<td>Gal-3-SO₄β1→4GlcNAcβ1→4GlcNHz</td>
<td>Galβ1→4GlcNAc-6-SO₄β1→4GlcNHz</td>
<td>(GlcAβ1→3AnTalH₂-6-SO₄β1→4GlcNHz)</td>
</tr>
<tr>
<td>Hydrazine</td>
<td>Gal-3-SO₄β1→4GlcNHz</td>
<td>Galβ1→4GlcNHz-6-SO₄β1→4GlcNHz</td>
<td>(GlcAβ1→3AnTalH₂-6-SO₄β1→4GlcNHz)</td>
</tr>
<tr>
<td>Nitrous acid</td>
<td>Gal-3-SO₄β1→4AnMan + R</td>
<td>Galβ1→4AnMan-6-SO₄ + R</td>
<td>n GlcAβ1→3AnTalH₂-6-SO₄ + R</td>
</tr>
<tr>
<td>Reduction*</td>
<td>Gal-3-SO₄β1→4AnManH₂</td>
<td>Galβ1→4AnManH₂-6-SO₄</td>
<td>n GlcAβ1→3AnTalH₂-6-SO₄</td>
</tr>
</tbody>
</table>

*Glycopeptides were treated in sequence as described under "Experimental Procedures"; the untreated sulfated saccharide chains and the products from each reaction are shown. R represents the remainder of the N-linked complex carbohydrate unit, while R' refers to the linkage region of the glycosaminoglycan chain.*

*The reactions of the sulfated sugar sequences leading to disaccharides B and A (see Fig. 1) are tabulated.

*Radiolabel is introduced into the disaccharide by the use of NaB₃H₄.
oligosaccharides (seven/330,000-kDa subunit), is restricted to a small segment of the polypeptide chain.²

Whereas the Galβ₁→4GlcNAc-6-SO₄ sequence found in thyroglobulin has long been recognized to occur in keratan sulfate (20), the chain-terminating β-linked Gal-3-SO₄ residues represent a novel structural feature of N-linked carbohydrate units which clearly differs from that of the sulfated GalNAc residues which are positioned at the periphery of the asparagine-linked oligosaccharides in pituitary glycoprotein hormones (21). In both thyroglobulin and pituitary hormones, the sulfate-bearing saccharide residue is in β₁→4 linkage to GlcNAc; but in the latter glycoproteins, sulfation is dependent on the prior attachment of a special determinant, the GalNAc residue (21). In contrast, sulfate addition to thyroglobulin can occur on galactose residues which are part of the prevailing N-acetyllactosamine sequences of complex N-linked units.

The striking difference in the extent of galactose sulfation which we observed between human and calf thyroglobulins merits consideration. Recent studies from our laboratory³ have indicated that the limited occurrence of Gal-3-SO₄ groups in the calf protein cannot be attributed to a deficient sulfotransferase or the absence of an appropriate recognition signal on the complex carbohydrate units. The presence in calf thyroglobulin of chain-terminating α-D-galactosyl residues in Galα₁→3Galβ₁→4GlcNAc sequences, which do not exist in human proteins (12), provides a reasonable explanation for the species dissimilarity in the sulfation patterns. Although sialylation of the complex carbohydrate units of both thyroglobulins must restrict the number of β-D-galactosyl residues available for sulfation, the additional attachment of α-D-galactosyl groups in the calf protein would be expected to further limit their number. Indeed, the constellation of the capping groups might, to a large measure, be determined by the relative activities of transferases involved in the attachment of sialic acid, sulfate, and α-D-galactose residues, respectively, to the β-D-galactose termini of the oligosaccharide branches.

The biological role of the sulfated sugars, both Gal-3-SO₄ and GlcNAc-6-SO₄, in the thyroglobulin complex carbohy-


REFERENCES