Ovine Submaxillary Mucin

PRIMARY STRUCTURE AND PEPTIDE SUBSTRATES OF UDP-N-ACETYLGALACTOSAMINE:MUCIN TRANSFERASE*

(Received for publication, December 13, 1976)

HOYLE D. HILL, Jr.,* MARTIN SCHWYZER,§ HOWARD M. STEINMAN,‖ AND ROBERT L. HILL

From The Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710

Tryptic digests of ovine submaxillary apomucin were fractionated by gel filtration and ion exchange chromatography to give 14 peptide fractions. Three purified tryptic peptides, representing 106 of the 660 residues in apomucin, were submitted to automated sequence analysis. The NH₂-terminal 50 of the 71 residues in one peptide and the entire sequence of the other two hexadecapeptides were established. These studies suggest that purified ovine submaxillary mucin is chemically homogeneous, containing a unique primary structure without substantial repeating sequences in its polypeptide chain.

The sequences adjacent to 28 known O-glycosidically substituted seryl and threoynl residues were compared. No homologies were apparent around the glycosylated seryl and threoynl residues which might define the specificity of the UDP-N-acetylgalactosaminyl:mucin polypeptide transferase that incorporates N-acetylgalactosamine into O-glycosidic linkage in glycoproteins. However, there appears to be a minimum size requirement for glycosylation, because the transferase catalyzes glycosylation of tryptic peptides efficiently, while chymotryptic and thermolytic peptides were much poorer substrates for the transferase.

In the preceding paper (1) evidence is reported that ovine submaxillary mucin is formed by noncovalent aggregation of glycoprotein subunits, with a molecular weight about 154,000. One-third of the residues in the subunit polypeptide chain (molecular weight about 58,300) are threonine or serine, each of which is substituted by the disaccharide group, N-acetylneuraminyl-a2→6-N-acetylgalactosamine in O-glycosidic linkage.

In this paper we report the purification and sequence analysis of three tryptic peptides which represent 100 of the 650 residues in the ovine mucin subunit chain. No evidence was obtained for an internally repeating sequence of about 28 residues, as suggested by others (2) for bovine submaxillary mucin, which has properties very similar to ovine mucin (3). In addition, there were no salient features in the sequences adjacent to serine and threonine residues that might prescribe a substrate acceptor requirement for incorporation of N-acetylgalactosamine into O-glycosidic linkage. Such sequence homologies have been suggested for the N-glycosidic linkage in other glycoproteins (4, 5). These ovine mucin peptides were, however, useful in examining the structural requirements for porcine submaxillary UDP-N-acetylgalactosamine-mucin polypeptide transferase (EC 2.4.1.41).

EXPERIMENTAL PROCEDURES

Materials

Ovine mucin and apomucin were isolated by gel filtration after dansylation to remove contaminating proteins, as described in the preceding paper (1). Porcine mucin was purified by the method of Planter (6). The following were obtained commercially: TPCK-treated trypsin and DFP-treated carboxypeptidase (Worthington Biochemical Corp.); dimethylaminonaphthalene sulfonyl chloride (Pierce Chemical Co.); dansyl amino acids (Calbiochem, Inc.); Staphylococcus aureus (V8) protease (Miles Laboratories, Research Division); Triton X-100 (Rohm and Haas); polyamide thin layer sheets (Gallard-Schlesinger Chemical Manufacturing Co.); UDP-N-acetyl-D-[14C]galactosamine (New England Nuclear); and Instagel (Packard Instruments, Inc.).

UDP-N-acetyl-D-galactosamine was synthesized as described earlier (7). Pyridine, triethylamine, triethanolamine, N-ethylmorpholine, and collidine were distilled from anhydride (1 to 2 g/liter) prior to use and stored protected from light, under N₂, at 4°C.

Protolytic digests of apomucin were prepared as substrates for UDP-N-acetyl-D-galactosaminyl:mucin polypeptide transferase. Apomucin (5 mg/ml in 0.046 M triethanolamine/HCl, pH 8.1, containing 0.012 M CaCl₂) was incubated at 37°C for 24 h with TPCK-trypsin (0.1 mg/ml). Apomucin (5 mg/ml in 0.08 M triethanolamine/HCl, pH 7.6, containing 0.1 M CaCl₂) was incubated at 37°C for 24 h with α-chymotrypsin (0.1 mg/ml) or thermolysin (0.1 mg/ml), respectively. Each digest mixture was heated at 100°C for 5 min, followed by 24-h incubation, to inactive the proteases.

1 The abbreviations used are: TPCK, L-1-(p-toluene sulfonyl)amido-2-phenylmethylchioromethyl ketone; DFP, diisopropylfluorophosphate; Mes, 2-(N-morpholino)ethanesulfonic acid; PTH, 1-phenyl-2-thiohydantoin; dansyl, 5-dimethylaminonaphthalene-1-sulfonyl.

*sis of three tryptic peptides which represent 100 of the 650 residues in the ovine mucin subunit chain. No evidence was obtained for an internally repeating sequence of about 28 residues, as suggested by others (2) for bovine submaxillary mucin, which has properties very similar to ovine mucin (3). In addition, there were no salient features in the sequences adjacent to serine and threonine residues that might prescribe a substrate acceptor requirement for incorporation of N-acetylgalactosamine into O-glycosidic linkage. Such sequence homologies have been suggested for the N-glycosidic linkage in other glycoproteins (4, 5). These ovine mucin peptides were, however, useful in examining the structural requirements for porcine submaxillary UDP-N-acetylgalactosamine-mucin polypeptide transferase (EC 2.4.1.41).

EXPERIMENTAL PROCEDURES

Materials

Ovine mucin and apomucin were isolated by gel filtration after dansylation to remove contaminating proteins, as described in the preceding paper (1). Porcine mucin was purified by the method of Planter (6). The following were obtained commercially: TPCK-treated trypsin and DFP-treated carboxypeptidase (Worthington Biochemical Corp.); dimethylaminonaphthalene sulfonyl chloride (Pierce Chemical Co.); dansyl amino acids (Calbiochem, Inc.); Staphylococcus aureus (V8) protease (Miles Laboratories, Research Division); Triton X-100 (Rohm and Haas); polyamide thin layer sheets (Gallard-Schlesinger Chemical Manufacturing Co.); UDP-N-acetyl-D-[14C]galactosamine (New England Nuclear); and Instagel (Packard Instruments, Inc.).

UDP-N-acetyl-D-galactosamine was synthesized as described earlier (7). Pyridine, triethylamine, triethanolamine, N-ethylmorpholine, and collidine were distilled from anhydride (1 to 2 g/liter) prior to use and stored protected from light, under N₂, at 4°C.

Protolytic digests of apomucin were prepared as substrates for UDP-N-acetyl-D-galactosaminyl:mucin polypeptide transferase. Apomucin (5 mg/ml in 0.046 M triethanolamine/HCl, pH 8.1, containing 0.012 M CaCl₂) was incubated at 37°C for 24 h with TPCK-trypsin (0.1 mg/ml). Apomucin (5 mg/ml in 0.08 M triethanolamine/HCl, pH 7.6, containing 0.1 M CaCl₂) was incubated at 37°C for 24 h with α-chymotrypsin (0.1 mg/ml) or thermolysin (0.1 mg/ml), respectively. Each digest mixture was heated at 100°C for 5 min, followed by 24-h incubation, to inactive the proteases.

1 The abbreviations used are: TPCK, L-1-(p-toluene sulfonyl)amido-2-phenylmethylchloromethyl ketone; DFP, diisopropylfluorophosphate; Mes, 2-(N-morpholino)ethanesulfonic acid; PTH, 1-phenyl-2-thiohydantoin; dansyl, 5-dimethylaminonaphthalene-1-sulfonyl.
Isolation of Tryptic Peptides from Apomucin

Aponbucin (5 mg/ml) was incubated at 37°C with TPCK-trypsin (50 μg/ml) and the mixture maintained at pH 7.5 to 8.2 by periodic addition of 1 M NaOH. After 2 h, the same amount of TPCK-trypsin was added and hydrolysis allowed to proceed for an additional 5 to 8 h and the mixture lyophilized. The mixture was dissolved in 10% acetic acid (5 mg/ml) and applied to a column (25 × 36.5 cm) of Sephadex G-50 (superfine) equilibrated with 10% acetic acid. The column was developed at 25°C with 10% acetic acid and fractions (2 ml) collected at 6-min intervals. Aliquots (50 μl) were analyzed manually with ninhydrin reagent after alkaline hydrolysis (7). Fractions were pooled and further purified by gel filtration (Fraction T1) on the same column in 10% acetic acid or by ion exchange chromatography (Fraction T4) on Dowex 50-X2 (200 to 400 mesh). Peptides (1 to 5 μmol) were applied to columns equilibrated with 0.01 M pyridine phosphate, pH 7, and the columns developed with the same solvent system as described earlier (Method 1 in Ref. 11). Elution patterns were monitored automatically with a diode array monitor (FDP-5) at 200 to 400 nm and the fractions containing the enzymatic activity were collected and those containing the enzymatic activity were collected and those containing the enzymatic activity were pooled, dried and stored frozen.

Sequence Analysis

Amino acid compositions were determined as described earlier (8). NH-terminal residues were qualitatively identified by the dansyl method of Gray (10), using chromatography on polywoloid thin layer sheets. Peptides were sequenced automatically with an updated Beckman 890B Sequencer, using either N,N-dimethyloctylamine or N,N-dimethylallylamine buffers programs. The phenylthiohydantoin derivatives were analyzed by gas chromatography and thin layer chromatography (9).

Preparation of UDP-N-Acetylgalactosaminy1-Mucin Polypeptide Transferase

A Triton X-100 extract was prepared from porcine submaxillary glands and chromatographed on UDP-hexaholamine-agarose as previously described (11). UDP-GalNAc mucin polypeptide transferase did not bind to UDP-agarose and was partially purified from the column effluent as follows. All steps were performed at 4°C. The transferase solution (2600 ml) was dialyzed for 2 days against two changes of 12 liters of water; 70 ml of 1 M sodium acetate, pH 5, were added and the solution centrifuged for 40 min at 5,400 rpm in a Sorvall RC-3 centrifuge (HG-4L rotor). The resulting supernatant was applied to a column (8 × 5 cm; 250 ml) of SE-Sephadex C-50 equilibrated with 0.025 M sodium acetate, pH 5, containing 0.65 M NaCl and 0.1% Triton X-100. The column was developed with a linear gradient formed with 1 liter of 0.025 M sodium acetate, pH 5, containing 0.1% Triton X-100, and 0.95 M sodium chloride and 1 liter of the same solvent containing 1 M sodium chloride. Fractions of 300 ml were collected and those containing enzymatic activity were brought to pH 7 by the addition of 1 ml imidazole/HCl, pH 8. Fractions 2 and 3 contained most of the transferease and were concentrated 10-fold by ultratritration through an Amicon PM-10 membrane. A portion (7 ml) of the concentrated solution was applied to a column (1.5 × 57 cm) of Sephadex G-100 equilibrated in 0.025 M imidazole/HCl, pH 7, containing 0.1 M sodium chloride and 0.1% TritonX-100 and developed in the same buffer at 4°C. Fractions (3.5 ml) were collected at a flow rate of 13 ml/h. The transferase emerged in Fractions 13 to 17. Fraction 12 contained 30-fold purified transferase (specific activity 0.06 unit/mg of protein) and was used in the studies described here. The transferease contained the following relative enzyme activities (each assayed under optimal conditions): UDP-N-acetylgalactosamine mucin polypeptide transferase (12), 100%; UDP-N-acetylgalactosamine:galactose transferase (11), 1%; UDP-N-acetylgalactosamine hydrolase, 2%; UDP-galactose:UDP-N-acetylgalactosamine transferase (13), <0.1%; CMP-sialyl-N-acetylgalactosamine transferase (14), <0.1%; CMP-sialylgalactose transferase (15), <0.1%.

Under assay conditions described below, transfer of N-acetylgalactosamine to apomucin was proportional to enzyme concentration and time of incubation. Linear Michaelis-Menten kinetic was observed with respect to the three substrates and the following K values were obtained from double reciprocal plots: 4 × 10⁻⁸ M for Man₄, 8 × 10⁻⁸ M for UDP-N-acetylgalactosamine, and 5 × 10⁻⁸ M (concentration of potential threonine and serine acceptor side) for apomucin. After removal of Triton X-100 with aid of Bio-Beads SM2 (16), the enzyme retained 100% activity in the standard assay.

Assay of UDP-GalNAc-Mucin Polypeptide Transferase

Method 1 (Standard Assay)—The incubation mixture (50 μl), prepared at 0°C, contained 8 μmol of MesiNai, pH 6.9, 1 μmol of MnCl₂, 0.5 mg of Triton X-100, 5 nmol of UDP-(1-³²P)GalNAc, and 0.15 mg of ovine apomucin, and up to 50 microunits of UDP-GalNAc-mucin polypeptide transferase. After incubation for 10 min at 37°C, the reaction product was isolated by Sephadex G-50 chromatography and counted as described earlier (Method 1 in Ref. 11). One enzyme unit is defined as that amount of UDP GalNAc-mucin polypeptide transferase catalyzing the transfer of 1 μmol of GalNAc to apomucin per min under standard assay conditions.

Method 2—The reaction mixture was prepared and incubated as in Method 1, but the reaction product was isolated by chromatography on Dowex 1-C₄, which did not bind apomucin, as described earlier (Method 3 in Ref. 11). Similar results were obtained with Methods 1 or 2 or with paper electrophoresis (12).

RESULTS

Isolation of Tryptic Peptides from Apomucin

Tryptic digests were expected to contain 21 different peptides since the apomucin subunit polypeptide chain contains 20 arginine residues. Separation of digests by gel filtration on Sephadex G-50, as shown in Fig. 1, gave four discrete peptide fractions, labeled T1, T2, T3, and T4.

Peptide T1—T1 was fractionated on Sephadex G-50 under the same conditions as in Fig. 1 to give a single peptide in 88% yield, with the amino acid composition listed in Table I. It contained a single NH-terminal serine residue as judged by dansyl coupling. Automated Edman degradation on a peptide sequenator for 34 cycles provided the sequence data in Table II. Hydrolysis of T1 with the Staphylococcus aureus protease, followed by gel filtration on Sephadex G-50 in 10% acetic acid, yielded two peptides (not shown), T1A and T1B, whose compositions are listed in Table I and sequence data in Table II. It is evident from Table II that a single Glu-Ser peptide bond has been selectively cleaved by the S. aureus protease, yielding an NH-terminal peptide of 18 residues, T1A, and a COOH-terminal peptide of 20 residues, T1B. From the combined sequence data, the NH-terminal 50 residues of T1 were established (Table III).

Peptide T2—This peptide gave one major peak on chromato-
Structure and Glycosylation of Ovine Submaxillary Mucin

Amino acids.

Glycosylation of Apomucin, Proteolytic Digests of Apomucin

and Pure Peptides from Apomucin by UDP-N-Acetylgalactosamine

Mucin Polypeptide Transferase

Quantitative ninhydrin determination of the appearance of

togon on Dowex 50 as described in Fig. 2 for T4 and contained only NH₂-terminal glycine by the dansyl method. However, analysis on a sequenator through 13 cycles indicated two peptide sequences, in approximately equal amounts, each commencing with NH₂-terminal glycine. These two peptides have not been separated and further characterized.

Peptide T3 — This fraction gave several peaks when fractionated on Dowex 50 as described for peptide T4 (Fig. 2), and contained at least 13 different peptides, but none were observed in pure form as judged by dansyl end group analysis.

Peptide T4 — This fraction was further purified on Dowex 50 as shown in Fig. 2 to yield two hexadecapeptides, T4c and T4f.

Comparison of Sequences around Each Seryl and Threonyl

Residue in T1, T4c, and T4f

Tables V and VI compare the amino acid sequences adjacent to each of the threonine and serine residues identified in peptides T1, T4c, and T4f. The composite sequence data from these three peptides contain 100 of the 650 residues and almost 15% of the hydroxyamino acids in the apomucin subunit chain.

An average of one in every 3 residues in apomucin is a serine or threonine, and each of these hydroxyamino acids is glycosylated on the basis of the composition of mucin and asialomucin (1). Inspection of Tables V and VI discloses no common neighboring sequence which might serve as a recognition complex for glycosylation, analogous to that found in N-glycosylated glycoproteins (4, 5), Asn-X-Thr/Ser, where Asn is glycosylated and X has been observed to be any one of a number of amino acids.

**Glycosylation of Apomucin, Proteolytic Digests of Apomucin**

and Pure Peptides from Apomucin by UDP-N-Acetylgalactosamine. Mucin Polypeptide Transferase

Quantitative ninhydrin determination of the appearance of

amino groups during proteolysis of apomucin indicated that the average length of peptides was 5 residues in the thermolytic, 15 residues in the chymotryptic, and 70 residues in the tryptic digests. This latter value is considerably greater than the average length of 31 residues for tryptic peptides predicted from the 20 arginine residues per 650 residues in the

| Amino acid | T1 Residues/molecule |
|------------|----------------------|
| Aspartic acid | 9.8 (10) |
| Threonine | 12.9 (13) |
| Serine | 5.4 (5) |
| Glutamic acid | 12.2 (12) |
| Proline | 14.8 (15) |
| Glycine | 6.2 (6) |
| Alanine | 4.6 (5) |
| Glutamine | 4.9 (5) |
| Asparagine | 1.8 (2) |
| Arginine | 1.0 (1) |

**TABLE I**

Characterization of mucin peptides

| Amino acid | Peptide | T1 | T1A | T1B | T4C | T4F |
|------------|---------|----|-----|-----|-----|-----|
| Aspartic acid | 9.8 (10) | 1.0 (1) |
| Threonine | 12.9 (13) | 3.9 (4) |
| Serine | 5.4 (5) | 3.3 (3) |
| Glutamic acid | 12.2 (12) | 3.5 (3) |
| Proline | 14.8 (15) | 3.2 (3) |
| Glycine | 6.2 (6) | 2.2 (2) |
| Alanine | 4.6 (5) | 1.0 (1) |
| Glutamine | 4.9 (5) | 1.1 (1) |
| Asparagine | 1.8 (2) | 2.0 (2) |
| Arginine | 1.0 (1) | 1.2 (1) |

**TABLE II**

Sequence analysis of T1, T1A, and T1B

| Cycle | Amino acid | T1 Residues/molecule |
|-------|------------|----------------------|
| 1 | Ser | Q |
| 2 | Ser | Q |
| 3 | Thr | 230 (+) |
| 4 | Pro | 223 (+) |
| 5 | Gly | 151 (+) |
| 6 | Gly | 175 (+) |
| 7 | Ser | Q |
| 8 | Ala | 169 (+) |
| 9 | Thr | Q (+) |
| 10 | Pro | 152 (+) |
| 11 | Gln | Q (+) |
| 12 | Gln | Q (+) |
| 13 | Pro | 123 (+) |
| 14 | Gly | 82 (+) |
| 15 | Ala | 120 |
| 16 | Leu | 65 |
| 17 | Ser | 8 |
| 18 | Gln | 81 (+) |
| 19 | Ser | Q |
| 20 | Thr | Q (+) |
| 21 | Thr | Q (+) |
| 22 | Gln | Q (+) |
| 23 | Leu | 64 |
| 24 | Pro | 47 (+) |
| 25 | Gly | 51 (+) |
| 26 | Val | 53 |
| 27 | Thr | Q (+) |
| 28 | Gly | 32 (+) |
| 29 | Thr | Q (+) |
| 30 | Ser | Q |
| 31 | Ala | 41 |
| 32 | Val | 31 |
| 33 | ? | |
| 34 | Gly | 22 |

**TABLE III**

Sequence analysis of T1, T1A, and T1B

| Cycle | Amino acid | T1A Residues/molecule |
|-------|------------|----------------------|
| 1 | Ser | Q |
| 2 | Ser | Q |
| 3 | Val | 230 (+) |
| 4 | Pro | 223 (+) |
| 5 | Gly | 151 (+) |
| 6 | Gly | 175 (+) |
| 7 | Ser | Q |
| 8 | Ala | 169 (+) |
| 9 | Thr | Q (+) |
| 10 | Pro | 152 (+) |
| 11 | Gln | Q (+) |
| 12 | Gln | Q (+) |
| 13 | Pro | 123 (+) |
| 14 | Gly | 82 (+) |
| 15 | Ala | 120 |
| 16 | Leu | 65 |
| 17 | Ser | 8 |
| 18 | Gln | 81 (+) |
| 19 | Ser | Q |
| 20 | Thr | Q (+) |
| 21 | Thr | Q (+) |
| 22 | Gln | Q (+) |
| 23 | Leu | 64 |
| 24 | Pro | 47 (+) |
| 25 | Gly | 51 (+) |
| 26 | Val | 53 |
| 27 | Thr | Q (+) |
| 28 | Gly | 32 (+) |
| 29 | Thr | Q (+) |
| 30 | Ser | Q |
| 31 | Ala | 41 |
| 32 | Val | 31 |
| 33 | ? | |
| 34 | Gly | 22 |
mucin subunit chain), a value consistent with the observed size distribution of the tryptic digests and the lengths of purified tryptic peptides (Fig. 1, Tables I and IV). This discrepancy for the tryptic peptides may reflect a systematic error in the estimation of average peptide length by ninhydrin reaction; for the tryptic peptides may reflect a systematic error in the estimation of average peptide length by ninhydrin reaction; for the tryptic peptides may reflect a systematic error in the estimation of average peptide length by ninhydrin reaction; for the tryptic peptides may reflect a systematic error in the estimation of average peptide length by ninhydrin reaction; for the tryptic peptides may reflect a systematic error in the estimation of average peptide length by ninhydrin reaction.

### Table III

| Peptide | Sequence |
|---------|----------|
| T1      | Ser-Ser-Val-Pro-Gly-Glu-Ser-Ala-Thr-Pro-Gly-Ser-Ala-Leu-50 |
|         | Ser-Glu-Thr-Thr-Gly-Leu-Pro-Gly-Val-Thr-Ser-Thr-Ala-Val-40 |
|         | Thr-Gly-Ser-Glu-Pro-Gly-Leu-Pro-Thr-Gly-Ser-Ala-Leu-Pro-60 |
|         | Gly-Thr(Thr, Ser, Gly, Ala, Val, Leu, Phe) - Arg |
| T4F     | Phe-Arg-Ser-Ala-Gly-Tle-Pro-Ala-Thr-Pro-Gly-Thr-Gly-Arg |

### Table IV

| Cycle | Amino acid | Yield | Amino acid | Yield |
|-------|------------|-------|------------|-------|
|       | nmol       |       | nmol       |       |
| 1     | Ser        | Q     | Phe        | 103 (+) |
| 2     | Ala        | 95    | Ser        | Q     |
| 3     | Gly        | 62 (+) | Ser        | Q     |
| 4     | Ala        | 122   | Ala        | 108   |
| 5     | Gly        | 53 (+) | Gly        | 62 (+) |
| 6     | Thr        | Q (+)  | Ile        | 52    |
| 7     | Ala        | 63    | Pro        | 49 (+) |
| 8     | Gly        | 33 (+) | Ala        | 61    |
| 9     | Val        | 40    | Thr        | Q (+)  |
| 10    | Asp        | 27 (+) | Pro        | 19 (+) |
| 11    | Ser        | 49    | Gly        | 15 (+) |
| 12    | Glu        | Q (+)  | Ser        | Q     |
| 13    | Glu        | Q (+)  | Thr        | Q (+)  |
| 14    | ?          |       | Thr        | Q (+)  |
| 15    | Ala        | 6     | Gly        | 12 (+) |

### Table V

| Cycle | Amino acid | Yield |
|-------|------------|-------|
|       | nmol       |       |
| 1     | Ser        | Q     |
| 2     | Ala        | 95    |
| 3     | Gly        | 62 (+) |
| 4     | Ala        | 122   |
| 5     | Gly        | 53 (+) |
| 6     | Thr        | Q (+)  |
| 7     | Ala        | 63    |
| 8     | Gly        | 33 (+) |
| 9     | Val        | 40    |
| 10    | Asp        | 27 (+) |
| 11    | Ser        | 49    |
| 12    | Glu        | Q (+)  |
| 13    | Glu        | Q (+)  |
| 14    | ?          |       |
| 15    | Ala        | 6     |

The pH of the assay mixture was unaffected by addition of the peptide.

The starting amounts of T4C and T4F were 150 and 140 nmol, respectively.

### Figure 2

Fractionation of T4 on Dowex 50. T4, obtained as shown in Fig. 1 from 20 mg of apomucin, was applied in 2.0 ml of 0.001 M pyridine/acetic acid, pH 1.75, to a column (0.9 × 59 cm) equilibrated with the same solvent. The column was developed at 60 ml/h at 50°C with 100 ml of equilibration buffer, then with the following linear gradients: 500 ml of 0.01 m pyridine/acetic acid, pH 1.75, starting buffer and 500 ml of 0.05 m pyridine/acetic acid, pH 3.25, starting buffer and 500 ml of 0.01 m pyridine/acetic acid, pH 3.25, starting buffer and 500 ml of 0.05 m pyridine/acetic acid, pH 5.0, starting buffer and 500 ml of 0.01 m pyridine/acetic acid, pH 5.0, starting buffer and 500 ml of 0.05 m pyridine/acetic acid, pH 5.0, starting buffer. The column was monitored automatically as described earlier (8). Fractions were pooled as indicated by the bars.
extent of about 1% of the potential acceptor sites, cleaved with V8 protease and peptides T1A and T1B separated chromatographically by the gel filtration system of Fig. 1, and the 14C radioactivity associated with each determined. Peptide T1A, containing 22% of the serine and threonine residues found in T1, had 10% of the radioactivity, and peptide T1B had 90%. This difference may reflect an inherent decreased reactivity of the residues near the NH2 terminus toward glycosylation, or it may indicate that the amino terminus is less accessible in the structural state of isolated T1 under assay conditions. It has been established that peptides T4C and T4F are acceptors, but the kinetics of glycosylation has not been examined thoroughly.

**DISCUSSION**

Ovine and bovine submaxillary mucins have similar physical and chemical properties (3). Each is a glycoprotein containing N-acetylneuraminyl α2 → 6-N-acetylgalactosamine prothetic groups in O-glycosidic linkage with serine and threonine residues in the protein backbone. Both contain about 60% carbohydrate and have molecular weights ranging from about 375,000 to well over 1,000,000 (3), the exact value depending primarily on ionic strength. The preceding paper (1) indicates that ovine mucin is formed by noncovalent self-association of subunits with a molecular weight of about 154,000, and that oligomer formation is dependent upon carbohydrate content, ionic strength, and mucin concentration. In contrast, apomucin does not self-associate and behaves as a monodisperse species on ultracentrifugation with a molecular weight of about 58,300. The studies reported here suggest that the polypeptide chain of ovine mucin (apomucin) has a unique amino acid sequence, and does not contain a regularly repeating sequence of about 28 residues as suggested for bovine mucin (2). This is supported by analysis of tryptic peptides of apomucin. The tryptic peptides of apomucin cannot be analyzed by peptide mapping on paper, since they have similar charges and streak on paper chromatography in the usual chromatographic solvents. By a combination of gel filtration and ion exchange chromatography (Figs. 1 and 2), tryptic digests could be resolved into at least 18 peptide fractions, close to the 21 expected from the arginine content of apomucin. Three peptides (T1, T4C, T4F) were purified, representing a total of 106 residues. Others, e.g. T2 (about 80 residues) and T3 appear to contain the remaining residues in the molecule. A total of 82 residues in the three pure peptides were placed in exact sequence (Table III). There was no indication of extensive sequence homologies among the peptides or of an internally repeating sequence of any significant length. Serine or threo-

![Table VI](image)

**Table VI**

*Comparison of sequences adjacent to threonyl residues in T1, T4C, and T4F*

- Gly-Gly-Ser-Ala-Thr-Pro-Gln-Gln-Pro-
- Leu-Ser-Glu-Ser-Thr-Thr-Gln-Leu-Pro-
- Ser-Glu-Ser-Thr-Thr-Gln-Leu-Pro-Gly-
- Leu-Pro-Gly-Val-Thr-Gly-Thr-Ser-Ala-
- Gly-Val-Thr-Gly-Thr-Ser-Ala-Val-Thr-
- Thr-Ser-Ala-Val-Thr-Gly-Ser-Pro-
- Gly-Leu-Pro-Ser-Thr-Val-Ser-Gly-
- Leu-Pro-Gly-Thr- - - - -
- Ala-Gly-Ala-Gly-Thr-Val-Ala-Asp-
- Asp-Ser-Gln-Thr-Ala-Arg- - -
- Gly-Ile-Pro-Ala-Thr-Pro-Gly-Thr-Ser-
- Thr-Pro-Gly-Thr-Thr-Gly-Arg- - -
- Pro-Gly-Ser-Thr-Thr-Gly-Arg- - -

![Fig. 3 (left)](image)

Fig. 3 (left). Glycosylation of apomucin and proteolytic digests of apomucin. Assay mixtures contained apomucin previously heated for 5 min at 100°C (●), trypptic digests of apomucin (△), chymotryptic digests of apomucin (●), and thermolysin digests of apomucin (○) at the concentrations indicated. To make all assays comparable to one another, CaCl2 and triethanolamine/HCl, pH 8, were added to final concentrations of 0.01 M and 0.008 M, respectively. After incubation for 30 min at 37°C, transfer of GalNAc was determined by Method 2.

![Fig. 4 (right)](image)

Fig. 4 (right). Glycosylation of peptides T1, T1A, and T1B from apomucin. Assay mixtures (10 μl) contained peptide T1 (●), peptide T1A (○), or peptide T1B (●), at the concentrations indicated. To make assays comparable to one another, NaCl and sodium cacodylate, pH 6, were added to final concentrations of 0.18 M and 0.09 M, respectively. All assay components other than acceptor were present in the concentrations given in Method 1. UDP-GalNAc had a specific activity of 15,000 cpm/nmol. After incubation for 30 min at 37°C, transfer of GalNAc was determined by Method 2.
nine residues were found on an average of about 1 in every 3
residues, as anticipated from the amino acid composition (1).
No evidence was found for sequence microheterogeneity in
that part of ovine apomucin whose partial sequence was exam-
ined.

Comparison of the sequences adjacent to each serine and
threonine residue in known sequences (Tables V and VI) did
not suggest a primary structural explanation of why each of
these hydroxyamino acids is glycosylated. There must be
structural requirements for acceptors of the UDP-N-acetylga-
lactosamine:mucin polypeptide transferase, which adds N-ace-
ty galactosamine to each serine and threonine hydroxyl group,
but these are not readily evident from the sequences available.

There is no common sequence in the 4 adjacent residues either
side of the glycosylated serine and threonine residues and no
indication of a sequence analogous to Asn-X-Thr/Ser, which is
found in all glycoproteins with carbohydrate prosthetic groups
N-glycosidically linked to asparagine (4, 5). The O-glycosyl-
ated serine and threonine residues appear in clusters of 3 to 9
residues in which at least every other residue is serine or
threonine and the clusters are interspersed by segments of 4 to
7 residues containing neither serine nor threonine. Notewor-
thy, however, is the sequence of the 50 residues in T1, the
longest stretch of sequence established. Based upon the rules
of Chou and Fasman (17) for predicting secondary structures,
an average of 1 out of every 3 to 4 residues throughout the
sequence is likely to break or destabilize either α helices or β
structures. This suggests that apomucin as well as mucin may
resemble random coils, in accord with earlier physical studies
by Gottschalk and McKenzie (18). Thus, O-glycosylation of
mucin and perhaps other glycoproteins (5) may well occur if
the serine and threonine acceptor residues are in regions of the
molecule with little secondary structure and are readily ex-
posed on the surface of the molecule. It may be that accessibility
rather than recognition of amino acid sequences is the key to
the specificity of glycosylation of ovine mucin. Studies with
apomucin, proteolytic digests of apomucin and peptides T1,
T4C, and T4F, support this view, since each was a substrate
for UDP-N-acetylgalactosamine:mucin polypeptide transfer-
ase. T1 (74 residues) was as good as an acceptor as apomucin.
T1B (56 residues) was a better acceptor than either T1A (18
residues) or T4C and T4F (each containing 16 residues), but
was not as readily glycosylated as T1. In accord with earlier
findings that small peptides containing threonine and serine
were not acceptors (12), thermolysin and chymotryptic digests
were poorer acceptor substrates than peptides in tryptic di-
gests. Since the smallest peptides in a tryptic digest probably
contain no more than 16 residues, it is feasible that a mini-
mum size is required for a good acceptor, although studies
with model peptides will be required to establish this size more
exactly.

REFERENCES
1. Hill, H. D., Jr., Reynolds, J. A., and Hill, R. L. (1977) J. Biol.
Chem. 252, 3791–3798
2. Pigman, W., Mosher, A. J., Weiss, M., and Tettamanti, G.
(1973) Eur. J. Biochem. 32, 148–154
3. Gottschalk, A., and Bhargava, A. S. (1972) in Glycoproteins:
Their Composition, Structure and Function (Gottschalk, A.,
ed) pp. 810–829, Kisevler, New York
4. Eylar, E. H. (1965) J. Theor. Biol. 10, 89–113
5. Aubert, J.-P., Biserte, G., and Loucheux-Lefebvre, M.-H. (1976)
Arch. Biochem. Biophys. 175, 410–418
6. Plantner, J. J. (1974) Ph.D. thesis, Case Western Reserve Uni-
versity
7. Hirs, C. H. W. (1961 Methods Enzymol. 11, 320–329
8. Hill, R. L., and Delaney, R. (1967 Methods Enzymol. 11, 339–351
9. Evans, H. J., Steinman, H. M., and Hill, R. L. (1974) J. Biol.
Chem. 249, 1316–1325
10. Gray, W. R. (1972 Methods Enzymol. 23, 121–138
11. Schwyzer, M., and Hill, R. L. (1977) J. Biol. Chem. 252, 2338–
2345
12. McGuire, E. J., and Roseman, S. (1967) J. Biol. Chem. 242, 3745–
3747
13. Schachter, H., McGuire, E. J., and Roseman, S. (1971) J. Biol.
Chem. 246, 5321–5328
14. Carlson, D. M., McGuire, E. J., Jeourdian, G. W., and Roseman,
S. (1973) J. Biol. Chem. 248, 5763–5773
15. Paulson, J. C., Beranek, W. E., and Hill, R. L. (1977) J. Biol.
Chem. 252, 2356–2362
16. Holloway, P. W. (1973) Anal. Biochem. 53, 304–308
17. Chou, P. Y., and Fasman, G. D. (1974) Biochemistry 13, 211–245
18. Gottschalk, A., and McKenzie, H. A. (1961) Biochem. Biophys.
Acta 54, 226–235
Ovine submaxillary mucin. Primary structure and peptide substrates of UDP-N-acetylgalactosamine:mucin transferase.
H D Hill, Jr, M Schwyzer, H M Steinman and R L Hill

J. Biol. Chem. 1977, 252:3799-3804.

Access the most updated version of this article at http://www.jbc.org/content/252/11/3799

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/252/11/3799.full.html#ref-list-1