Study of the Primary Structures of the Peptide Core of Bovine Estrus Cervical Mucin

POSSIBLE EXISTENCE OF SMALL SIMILAR SUBUNITS*

(Received for publication, July 13, 1976, and in revised form, June 7, 1977)

KANDUKURI S. P. BHUSHANA RAO AND PIERRE L. MASSON

From the Unit of Experimental Medicine, International Institute of Cellular and Molecular Pathology, 75, Avenue Hippocrate, B-1200 Brussels, Belgium

Two populations of tryptic peptides were isolated from bovine estrus cervical mucin (BCM). One contained all the carbohydrate, and was rich in threonine and serine. These glycopeptides had, like the whole mucin, alanine as their NH₂-terminal residues. Their COOH-terminal residues were arginine. The second population of peptides was rich in carboxylic amino acids, contained two cysteinyl residues, and had, like the whole mucin, leucine as COOH-terminal residues. Their NH₂-terminal residues were aspartic acid.

The sum of the residues of one glycopeptide plus one cysteinyl-containing peptide corresponded to the number of residues constituting a putative subunit of BCM. The amino acid sequence of the major cysteinyl peptide was determined. A cluster of hydrophobic residues was found in the COOH-terminal region. The amino acid sequences of two of the glycopeptides were found identical up to the 22nd residue.

The small number of tryptic peptides, as well as the large amount of NH₂- and COOH-terminal amino acids found in BCM indicate that this glycoprotein is made up of similar subunits with a molecular weight of about 22,000, one of the glycopeptides representing the NH₂-terminal part, and one of the cysteinyl peptides, the COOH-terminal part. However, the existence of these subunits was not confirmed by ultracentrifugation of BCM in dithiothreitol and sodium dodecyl sulfate. BCM was polydisperse and had a mean molecular weight of 507,000.

Cervical mucus is generally compared to a sort of valve which, at certain periods during the reproductive cycle, allows the entry of sperm into the uterus and, at other times, bars its admission. Therefore, the study of the glycoprotein which is responsible for the physical and physiological properties of cervical mucus is of great interest in the physiology of reproduction.

Bovine estrus cervical mucin is a glycoprotein displaying a continuous type of variation based mainly on differences in the carbohydrate moiety (1), which represents about 70% of the whole mucin (2, 3). We have previously reported (4) that BCM contains groups of amino acids in a constant ratio, which suggested the existence of similar subunits or of repeating sequences. This latter type of structure has been proposed for bovine submaxillary mucin (5).

Two populations of tryptic peptides were isolated from BCM, and the BCM was purified by filtration on Sepharose 6B as previously described, after reduction with 0.01 M dithiothreitol (4). Neuraminic acid was released from purified BCM by hydrolysis with 0.01 M HCl at 80° for 1 h or by neuraminidase digestion (4). A 4% solution of neuraminic acid-free BCM was further reduced with 0.1 M thioglycolic acid (Calbiochem) in 0.5 M Tris/HCl buffer, pH 8.2, containing 6 M guanidinium chloride, for 90 min at 37°. The material was then alkylated for 2 h at 37° using 0.24 M (¹⁴C)iodoacetic acid (126 µCi) (The Radiochemical Centre, Amersham, U. K.) in 1 M Tris solution containing 4 M guanidinium chloride.

The contaminating amino acids in desicilyl BCM were removed with copper-saturated Chelex. The ligand exchange column (27 × 2.5 cm) was prepared according to the method of Bellinger and Buist (8), using sodium Chelex-100 (50 to 100 mesh) from Bio-Rad.

* This investigation was supported by the Ford Foundation Grant 669-9435 and World Health Organization Grant H9/181/111. 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.
The reduced and alkylated mucin was dialyzed against 0.1 M ammonium hydroxide for 24 h before application to the Chelex, and eluted with the same solution. The free amino acids were eluted with 0.0 M ammonium hydroxide. To remove the copper ions from the mucin, it was passed through a column of sodium Dowex AG 50W-X8 equilibrated with 0.1 M ammonium hydroxide. The BCM was then chromatographed on Sepharose 6B in 0.05 M ammonium bicarbonate to remove unidentified radioactive material of low molecular weight.

**Fractionation of Tryptic Peptides**

Aminex Column - Tryptic peptides from 1 g of BCM were dissolved in 3 ml of 0.2 M pyridine adjusted to pH 3.1 with glacial acetic acid and applied to a column (21 × 2.0 cm) of Aminex Q15S (Bio-Rad) in the same buffer. The peptides were eluted at a flow rate of 30 ml/h at room temperature by a gradient of increasing concentration of pyridine (9). After washing the column for 1 h with pyridine/acetic acid buffer, pH 3.1, a gradient generated by a Varigrad mixer was applied. Each cell of the mixer was filled with 324 ml of pyridine/acetic acid buffer, the molar concentration of pyridine and corresponding pH being, respectively, 0.1 M and pH 3.1 in Cells I to 4; 2 M and pH 5.1 in Cells 5 to 7; 2 M and pH 6.5 in Cells 8 and 9.

Elution was monitored by a Technicon peptide AutoAnalyzer with aliquots being removed at regular intervals for staining with ninhydrin before and after alkaline hydrolysis. In each fraction, the radioactivity was measured in a Packard scintillation counter, and the sugar content determined by the anthrone method.

Bio-Gel P-2 - The peptides eluted by the Bio-Gel P-2 column on Aminex Q15S were applied to a Bio-Gel P-2 column (200 to 400 mesh, Bio-Rad 132 × 2.25 cm) in 0.1 M acetic acid, and eluted at a flow rate of 15 to 20 ml/h with the same solution. The peptide fractions were pooled and lyophilized.

Dowex AG-I-X2 - Fraction I from the Bio-Gel P-2 column was dissolved in 5 ml of a mixture of pyridine/picoline/pyridine buffer adjusted to pH 9.4 with glacial acetic acid, and applied to a column (23 × 1.5 cm) of Dowex AG-I-X2 (200 to 400 mesh, Bio-Rad). After washing the column for 1 h with this buffer, the peptides were eluted at a flow rate of 10 to 15 ml/h with decreasing pH (10). The pH gradient was generated by the Varigrad mixer filled in the following manner: buffer, pH 9.0, in Cell 1; buffer, pH 8.4, in Cells 2 and 3; buffer, pH 6.5, in Cells 4 to 6; 0.5 M acetic acid in Cells 7 and 8 and 2 M acetic acid in Cell 9. Elution was monitored with a Technicon AutoAnalyzer.

**Analytical Ultracentrifugation**

BCM and a tryptic glycopeptide of BCM were centrifuged at 20°C in a Spinco model E ultracentrifuge using a 12 mm double sector

**Determination of COOH-terminal Amino Acids**

The COOH-terminal amino acids of BCM were released by Enziate-agarose-carboxypeptidase A (350 units/ml suspension) from Miles-Seravac Laboratories. BCM (17 mg) was dissolved in 2.83 ml of 0.1 M ammonium acetate buffer, pH 7.5, containing 0.5% LiCl. Then samples of 3 mg of BCM were applied on the Enziate-agarose column and incubated at 37°C for 12, 24, 32 and 48 h. After each incubation, the sample was eluted and the column was washed three times with buffer. The eluates were then lyophilized, and their amino acid composition was analyzed. After a control, the column of insolubilized carboxypeptidase A was incubated with buffer only and the eluates treated in the same way as the samples before amino acid analysis. The amounts of amino acids obtained in these determinations were subtracted from those found in the mucin samples.

The tryptic peptides were digested with carboxypeptidases A and B (Worthington) in a soluble form according to the method described by McKean et al. (17), following pretreatment with diisopropyl phosphorofluoridate.

**Cleavage by Cyanogen Bromide**

In a tightly stoppered vial, 8 mg of glycopeptide I was dissolved in 1.2 ml of 70% formic acid. To this, 8 mg of solid CNBr (Koch-Light Laboratories) was added, and the mixture was stirred at room temperature. Four 400-μl samples were removed at intervals of 0.24, 32, and 48 h, diluted 10 times, and lyophilized. One part of this material was hydrolyzed with 6 N HCl at 110°C for 24 h in an evacuated sealed tube, and submitted to amino acid analysis. The rest was analyzed by electrophoresis at pH 7.4 in a 7.5% polyacrylamide gel containing 0.1% SDS, in 7% polyacrylamide gel. A control experiment was run simultaneously without CNBr under similar conditions. After the electrophoresis the gels were stained with Coomassie blue.

**Amino Acid Sequence Determination**

Sequences from 100 to 200 nmol of peptides were determined by the manual dansyl-monitored Edman method (18). The dansyl amino acids were identified by chromatography on polyamide layers (19). Each amino acid residue was further identified as a phenyl thiohydantoin derivative. To obtain a clean phenyl thiohydantoin derivative at each Edman cycle before the cleavage with trifluoroacetic acid, the phenyl thiocarbamyl peptide was extracted three times with benzene to remove by-products of the coupling reaction. The thiourolione derivatives of the amino acids after extraction with ethyl acetate were converted to phenyl thiohydantoin derivatives by heating 10 min at 90°C in 1 N HCl under nitrogen. The phenyl thiohydantoin derivatives were identified by gas liquid chromatography on a column of Chromosorb WW/dimethyl chlorosilane, 80 to 100 mesh, containing 10% SP-400 as a stationary phase (20), and by chromatography on polyamide sheets (21). The fluorescence scintillation was used in the filter system in place of butyl-PBD (2,4'-d Butylphenyl-5,5'- diphenylyl-1,3,4-oxidiazole).

**High Voltage Paper Electrophoresis and Paper Chromatography**

HVE was performed according to Aslam (11) on 1- or 3-mm thick Whatman paper for 60 to 180 min at pH 1.9 (8% acetic acid, 2% formic acid), pH 3.6 (5% acetic acid, 0.5% pyridine), or pH 9.2 (0.5% sodium tetraborate). The peptides were eluted with 20% acetic acid. Descending paper chromatography was performed with 1-butanol-acetic acid-water-pyridine (15:3:13:10) (12). α-DNP-lysine, Xylene blue cyanol FF and Red Pentol were used as visual markers in addition to the usual amino acid markers. Kodak Blue Brand film was used for autoradiography.

**Amino Acid Analysis**

Samples were hydrolyzed for 18 h in evacuated and sealed ampules at 110°C using 0.5 ml of 5.7 N HCl containing 0.1% phenol. After hydrolysis, the samples were dried in vacuo over NaOH and P2O5, and redissolved in 2 ml of 0.2 M sodium citrate buffer. Amino acid analysis was performed by the technique of Spackman et al. (13), using a Beckman UniChrom Analyzer with a 10-mm flow cell which increased the sensitivity of the analysis to 0.01 μmol.

The analysis of whole mucin and glycopeptides was repeated with a special procedure recommended by Laskowski, which separates leucine from the carbohydrate peak. Peptides containing large amounts of serine and threonine were run at two different dilutions of the same hydrolysate.

**Determination of NH₂-terminal Amino Acid**

The NH₂-terminal amino acids of the peptides were identified from their dansyl derivatives (14) using thin layer chromatography on micropolyamide layers. The NH₂-terminal alanine from BCM (4) was determined by treatment (15) with 1-fluoro-2,4-dinitro-3,6H benzeno from the Radiochemical Centre, U.K. Standards were prepared by treating the free amino acid with [3H]PDBN in the same way as the mucin. For labeling 20 mg of BCM or 10 mg of alanine, respectively, 163 μCi (0.6 μCi) and 40.8 μCi (180 μCi) of [3H]PDBN were used. After the reaction, excess PDBN was extracted with peroxide-free ether. As BCM did not precipitate on acidification, it was lyophilized after drying against desiccated water. After hydrolysis with 5.7 M HCl for 8 h, [3H]PDBN-alanine was extracted with ether three times and redissolved in 1 ml of ether.

For the standard curve, fractions of 2, 4, 6, and 8 μl of the DNP-alanine preparation were spotted on micropolyamide layers with reference DNP-amino acids, and resolved by two-dimensional chromatography using the system of Wang and Wang (16). The spots were scraped off, the [3H]PDBN-alanine eluted with 1 ml of dioxane, and the radioactivity was counted. For the determination of the DNP-alanine released from BCM, 15 and 25 μl of the extract were treated in the same way as the standards.

**Bovine Cervical Mucin**

7789

Analytical Ultracentrifugation

BCM and a tryptic glycopeptide of BCM were centrifuged at 20°C in a Spinco model E ultracentrifuge using a 12 mm double sector
of BCM preparations was measured in a Perkin-Elmer fluorescence spectrophotometer using a cuvette of 1-cm path length. The incident wavelength was 350 nm and the analyzer wavelength 360 nm.

**RESULTS**

Tryptic Digestion of BCM — In pilot experiments, we have tried to digest reduced and alkylated BCM with L-L-tosylamido-2-phenylethyl chloromethyl ketone-treated trypsin (Worthington) at 37°, in 0.126 M ammonium bicarbonate buffer, pH 8.2, at an enzyme:substrate ratio of 1:50. The maximal yield of peptides obtained after 24 h was 5 to 10% of the initial material. Incubation was limited to 24 h because of the risk of trypsin autodigestion (25). In a second series of experiments, insolubilized trypsin (Enzite-agarose-trypsin from Miles-Seravac) was used. BCM, which had been reduced and alkylated with 14C-labeled iodoacetic acid, was incubated at 37° with 500 units of enzyme/g of BCM, in ammonium bicarbonate buffer, pH 8.2, for various times, i.e., 24, 48, 72, 90, 125, and 150 h. The suspension was saturated with toluene to prevent the growth of bacteria. The undigested material was then redigested with trypsin.

The yield of peptides for each cycle was calculated by measuring the area of the elution profile monitored at 280 nm. The maximal yield reached about 60% after 125 h incubation. Incubation of BCM under the same conditions but without trypsin failed to provide any peptides detectable in the Sepharose 6B chromatography. The high molecular weight material was then redigested with trypsin. The yield of peptides for each cycle was calculated by measuring the area of the elution profile monitored at 280 nm. The maximal yield reached about 60% after 125 h incubation. Incubation of BCM under the same conditions but without trypsin failed to provide any peptides detectable in the Sepharose 6B chromatography.

Tryptic Peptide Map of BCM — After ninhydrin staining of the trypsin digest of BCM (Fig. 1), one major and some minor spots were revealed. The major spot contained carbohydrate, as it reacted strongly with the phenol/sulfuric acid reagent (26). The other spots corresponded to free neutral and basic amino acids as shown by HV electrophoresis, amino acid analysis and dansylation with or without acid hydrolysis of the eluted spots.

On radioautography, two spots were visualized in the zone of the acidic markers. They were just visible on the ninhydrin-stained paper (Fig. 1). The radioactive spot migrating with Red Pentel was identified as free carboxymethylcysteine. The second spot represented the major cysteinyl containing peptide, which will be described hereafter.

Fractionation of BCM Peptides (Fig. 2) — The tryptic peptides of BCM were passed through a column of cation exchange resin (Aminex Q15S) and eluted with a pH and ionic strength gradient of pyridine/acetate buffer. Estimated by amino acid analysis, 24% of the starting material was retained by resin. This material consisted mainly of free amino acids, as basic hydrolysis did not affect their staining with ninhydrin. By HV electrophoresis at pH 1.9 and 3.6, we have recognized leucine, serine, threonine, glycine, valine, phenylalanine, tyrosine, and proline. A few di- or tripeptides were identified by amino acid analysis and dansylation. No arginine or lysine was found in these oligopeptides.

The major fraction from the Aminex column resolved into two peaks on Bio-Gel P-2 (Fig. 3). The first one eluted in the void volume contained most of the carbohydrate and was slightly radioactive. The second one was highly radioactive. These two fractions will be referred hereafter to as glycopeptides and cysteinyl peptides, respectively.

Fractionation and Characterization of Glycopeptides — The glycopeptides were further purified by chromatography on Dowex 1-X2 in a picoline/acetate acid system which removed small amounts of contaminating cysteinyl peptides and resolved the glycopeptides themselves into two fractions called glycopeptide I and glycopeptide II.

After desalting on Sephadex G-15, these two fractions were passed through a column of Sephadex G-25 in 2% dodecanoic acid adjusted to pH 10.0 with ammonium hydroxide (27). This resulted in the elimination of some residual free amino acids.

After this treatment, each glycopeptide gave a single diffuse band in HV electrophoresis at pH 9.2, with slightly different mobilities.

The amino acid compositions were found to be similar (Table I). Trace amounts of methionine were detected in glycopeptide II. However, after treatment of the glycopeptides with CNBr, we failed to obtain any fragment detectable by...
we calculated a minimal molecular weight of 4,000 for the peptide core of glycopeptide I and 4,400 for that of glycopeptide II, taking the amounts of arginine and aspartic acid as the basal units for estimating the number of residues.

The COOH-terminal residues of the two glycopeptides were identified by digestion with soluble carboxypeptidase A and B at an enzyme:substrate molar ratio of 6:100 and 20:100 for glycopeptide I and II, respectively. Carboxypeptidase A was added after a 1-h incubation with carboxypeptidase B. A rapid release of arginine was observed from glycopeptide I, reaching a maximum (16.5 nmol/mg) after a 4-h incubation. Incubation with carboxypeptidase B did not cause any release of amino acids from glycopeptide II, but the addition of carboxypeptidase A caused the release of significant amounts of threonine (21.1 nmol/mg), arginine (6.4 nmol/mg), alanine (6.0 nmol/mg), and valine (6.0 nmol/mg). Autodigestion of carboxypeptidase A due to the high enzyme/substrate ratio could have been responsible for this.

The amino acid sequences of glycopeptide I and II were determined by manual dansyl monitored Edman degradation up to the 22nd residue and found to be identical (Fig. 4).

Fractionation and Characterization of Cysteinyl Peptides – When the retained fraction from Bio-Gel P-2 chromatography of the tryptic peptides of BCM was applied to a Sephadex G-15 column in 0.1 M ammonium hydroxide, two radioactive peaks were obtained. One corresponded to the excluded fraction and contained the cysteinyl peptides, while the second, which was eluted in a similar position to that of an ε-dinitrophenyl lysine used as reference, appeared to be free carboxymethyllysine. From the radioactivity as well as from the amino acid composition of the two fractions, it was calculated that 50% of the carboxymethyllysine of the material applied on the Sephadex G-15 column was in a free form.

The cysteinyl peptides were further analyzed on a column of Chromobeads. Two fractions, representing 30% and 70% of the starting material, were eluted with 20% acetic acid and pyridine/acetic acid buffer, pH 3.1, respectively. The first fraction, called Fraction A, had aspartic acid as the NH₂-terminal residue, and resolved into three bands on HV electrophoresis at pH 3.6. These bands were only detected by radioautography and did not stain with cadmium/ninhydrin, suggesting that the NH₂-terminal residue was asparagine rather than aspartic acid. The amino acid composition of Fraction A is reported in Table II.

Fraction B from the Chromobeads resolved in HV electrophoresis at pH 3.6 into six ninhydrin-positive and radioactive bands. These bands were eluted from the paper and their amino acid compositions determined (Table II). Significant differences were observed in the numbers of cysteine, threonine, and glycine residues. Aspartic acid was identified as the NH₂-terminal residue of all these peptides. Bands II and III, which had very similar amino acid compositions, were pooled and analyzed for COOH-terminal amino acids using insolubilized carboxypeptidase A, as described under "Materials and Methods" for whole BCM, with incubation times of 5, 60, 120, and 180 min. After 5 min, 70% of the leucine had been released, with the maximum level (580 nmol/mg) being reached in 60 min.

When the amino acid composition of Fractions A and B from the Chromobeads were compared, significant differences appeared only for the carboxylic amino acids. Fraction A contained 2 aspartic acid residues and 3 glutamic acid residues; Fraction B, 3 aspartic acid and 2 glutamic acid residues.

The major electrophoretic bands from Fraction B, i.e. Bands

---

**Table I**

| Amino acid | Glycopeptide I residues | Glycopeptide II residues |
|------------|-------------------------|-------------------------|
| Aspartic   | 1                       | 0                       |
| Threonine  | 14                      | 15                      |
| Serine     | 5                       | 6                       |
| Glutamic   | 2-3                     | 2-3                     |
| Proline    | 5                       | 5                       |
| Glycine    | 2                       | 2                       |
| Alanine    | 5                       | 5                       |
| Cysteine   | 0                       | 0                       |
| Valine     | 3                       | 4                       |
| Methionine | 0                       | 0-1                     |
| Isoleucine | 1                       | 1                       |
| Leucine    | 1                       | 1                       |
| Tyrosine   | 0                       | 0                       |
| Phenylalanine | 0                   | 0                       |
| Lysine     | 0                       | 0-1                     |
| Histidine  | 0                       | 0                       |
| Arginine   | 1                       | 1                       |
| Total number of residues | 40-41 | 42-45 |
| NH₂-terminal residue | Alanine | Alanine |
| COOH-terminal residue | Arginine | Arginine |

---

gel acrylamide electrophoresis in SDS, and no homoserine or homoserine lactone were detected in the amino acid composition of the cyanogen bromide-treated samples.

The partial specific volume for glycopeptide I was found to be 0.56 and this value was used in calculating its average molecular weight. However, owing to polydispersity, ultracentrifugation failed to provide a precise molecular weight. The values lay between 6,400 and 15,800, corresponding to 1,320 and 4,720 for the peptide core, allowing a carbohydrate content of 70% for glycopeptide I. From the amino acid composition,

---

**FIG. 3.** Gel filtration of tryptic peptides of BCM on a column of Bio-Gel P-2 in 0.1 M acetic acid. Fractions, 5.5 ml. Absorbance at 280 nm was less than at 220 nm, but was parallel.

The amino acid compositions of tryptic glycopeptides of BCM are reported in Table II. Amino acid Glycopeptide I residues Glycopeptide II residues

---

...
Bovine Cervical Mucin

Glycopeptide

FIG. 4. Amino acid sequence of glycopeptides I or II determined by manual dansyl-monitored Edman degradation. The amino acid composition of the peptide segment of which the sequence has not been determined is reported in parenthesis.

 Ala-Thr-Pro-Gly-Pro-Ser-Val-Thr-Gly-Thr-Ser-Pro-Ala-Thr-Pro-Thr-Pro-Thr-Val

20 22
Pro-Ala-Ser-(Asp₂₋₁; Thr₉; Ser₄; Glu₂₋₃; Ala₂; Val₂; Met₀₋₁; Leu₁; Lys₀₋₁)-Arg

Table II

Amino acid composition of tryptic cysteinyl peptides from BCM, after fractionation on a Chromobeads column

| Amino acid            | Fraction A | Eluted bands of HVE, pH 3.8 of Fraction B |
|-----------------------|------------|------------------------------------------|
|                       | I         | II          | III         | IV          | V          | VI          |
| Carboxymethyl cysteine| 1         | 1-2         | 2           | 1           | 1          | 2           |
| Aspartic acid         | 2         | 3-4         | 3-4         | 3           | 3          | 3           |
| Threonine             | 1-2       | 1           | 1           | 1           | 1          | 2           |
| Serine                | 1         | 1           | 1           | 1           | 1          | 1           |
| Glutamic acid         | 3         | 2           | 2           | 2           | 2          | 2           |
| Proline               | 1         | 1           | 1           | 1           | 1          | 1           |
| Glycine               | 1         | 2           | 1           | 1           | 1-2        | 2           |
| Alanine               | 1         | 1           | 1           | 1           | 1          | 1           |
| Valine                | 1         | 1           | 1           | 1           | 1          | 1           |
| Methionine            | 0         | 0           | 0           | 0           | 0          | 0           |
| Isoleucine            | 1         | 1           | 1           | 1           | 1          | 1           |
| Leucine               | 1         | 1           | 1           | 1           | 1-2        | 1           |
| Tyrosine              | 1         | 1           | 1           | 1           | 1          | 1           |
| Phenylalanine         | 1         | 1           | 1           | 1           | 1          | 1           |
| Lysine                | 0         | 0           | 0           | 0           | 0          | 0           |
| Histidine             | 0         | 0           | 0           | 0           | 0          | 0           |
| Arginine              | 0         | 0           | 0           | 0           | 0          | 0           |
| Number of residues    | 16–17     | 17–19       | 17–19       | 16          | 16–18      | 19          |
| NH₂-terminal residue  | Asx       | Asx         | Aspartic acid| Aspartic acid| Asx        | Asx         |
| COOH-terminal residue | Leucine   | Leucine     | Leucine     | Leucine     | Leucine    | Leucine     |

Cysteinyl peptide

Asp-Gln-Asn-Cmcys-Asn-Cmcys-Glu-Thr-Gly-Ser

THR Pro-Ile-Gly-Val-Ala-Tyr-Phe-Leu

Fig. 5. Amino acid sequence of a tryptic cysteinyl peptide from BCM. The cysteinyl peptide corresponded to the material eluted from the electrophoretic bands V and VI from the Chromobeads Fraction B (see Fig. 2). The manual Edman degradation was difficult to pursue after the 12th residue because of the loss of the residual peptide in the benzene washing and ethyl acetate extractions. Dns-isoleucine and Dns-glycine of positions 13 and 14 were detected in trace amounts, whereas Dns-valine of position 15 was clearly visualized on the thin layer plate. The results obtained with carboxypeptidase A were evident up to the 14th residue (glycine). The yield was 169 nmol/mg for leucine, 76 nmol/mg for phenylalanine, 89 nmol/mg for tyrosine, 42 nmol/mg for alanine, 20 nmol/mg for valine, and 15 nmol/mg for glycine. The tyrosine residue was placed in position 17 before phenylalanine because the release of the latter was faster than that of tyrosine.

V and VI, which had identical amino acid compositions, were pooled. The amino acid sequence of this material was determined up to the 15th residue. The sequence of the COOH-terminal part from the 14th residue was obtained by carboxypeptidase A digestion (Fig. 5).

NH₂- and COOH-terminal Amino Acids of BCM—To quantify the NH₂-terminal amino acid of BCM, i.e. alanine (4), a standard curve with pure alanine was prepared (Fig. 6). It was calculated that two 1-mg samples of DNP-mucin con-
found to be polydisperse. An average molecular weight of
chloride; Doe, deoxycholate.

A for 24 h released 21 nmol of leucine, 12 nmol of phenylala-
reagents on BCM. DTT, dithiothreitol; Guan HC1, guanidinium

happened in saline, the addition of dithiothreitol to BCM in
SDS or guanidinium chloride decreased $I_{\alpha}$ by 6 to 8%.

The addition of a reducing agent such as 6 mM dithiothreitol
to BCM caused a slight increase of $I_{\alpha}$, whereas 3
molecular size is lower than the wavelength of the incident light. Therefore, it is possible with a relatively small
instrument such as a fluorometer used as a nephelometer to
monitor the changes in the size of macromolecules. This
system is commonly used in immunochemistry (28).

The addition of a reducing agent such as 6 mM dithiothreitol
to bcm caused a slight increase of $I_{\alpha}$, whereas 3
molecular size is lower than the wavelength of the incident light. Therefore, it is possible with a relatively small
instrument such as a fluorometer used as a nephelometer to
monitor the changes in the size of macromolecules. This
system is commonly used in immunochemistry (28).

By ultracentrifugation in 10 mM dithiothreitol, BCM was
found to be polydisperse. An average molecular weight of
850,000 was estimated. When 10 mM SDS was added to this
preparation, the weight average molecular weight dropped by
41% to 510,000.

**DISCUSSION**

The present work indicates that BCM consists of similar
subunits which yield two populations of tryptic peptides, *i.e.*
one group of peptides carrying all the oligosaccharide side
chains, and a second group of peptides characterized by their
high content of carboxylic amino acids and by the presence of
cysteinyl residues (Fig. 8). Gelman and Vered (29) have
obtained cyanogen bromide fragments of BCM differing in
the same way as our tryptic peptides. However, this effect of
CNBr is questionable because these authors did not report
the presence of homoserine or homoserine lactone in their
CNBr fragments. Our sequence data showed that the cysteinyl
peptides are devoid of methionine. In addition, we have been
unable to split the glycopeptides by CNBr, and no homoserine
or homoserine lactone was detected in the CNBr-treated
glycopeptides. The fragmentation observed by Gelman and
Vered (29) could be due to proteases from leukocytes or
bacteria which abound in the cervical secretion (30). In other
glycoproteins, such as blood group substances from human
ovarian cyst mucin, two regions differing in their content of
carbohydrate and charged amino acids have also been de-
scribed (31, 32).

Special conditions were necessary for the trypsin digestion
of BCM. Even after removal of sialic acid, which is known to
protect glycoproteins against proteolysis (33), a satisfactory
yield of peptides was obtained only after 6 days incubation
with insolubilized trypsin. Digestion by contaminating bacte-
rial or leukocytic proteases was unlikely since no peptides
were recovered from BCM incubated under the same condi-
tions in the absence of trypsin. For the present study, we
have prepared, with consistent results, 12 tryptic digests from
different samples of BCM.

From the analysis of the amino acid composition of the
whole mucin, BCM would contain 2 arginine and 1 lysine
residues. Hence, we should have obtained four tryptic peptides
in place of two. This discrepancy could be due to the inaccess-
sibility of some basic residues of BCM, as suggested by the
particular conditions required for the tryptic digestion. An-
other explanation, which appears more likely, is that some of
the basic amino acids detected in the whole mucin were
contaminants. The cysteinyl peptides are devoid of basic
amino acids (Table II), the glycopeptides contain only 1
arginine residue (Table I), and the presence of lysine in
glycopeptide II was questionable. Therefore, we believe that
the BCM subunit contains only 1 arginine residue.

The fractionation of the tryptic peptides was complicated
by contamination with free amino acids and di- or tripeptides.
These were not part of the BCM molecule, since the sum of
the amino acid compositions of the glycopeptides and the
cysteinyl peptide accounted for the composition of the whole
molecule (Table III). The abundance of free amino acids and
oligopeptides contaminating BCM is probably related to partial
digestion of the soluble proteins and the mucin itself before fractionation of the cervical secretion. The numerous
leukocytes infiltrating the cervical submucosa and present in
the mucus itself may account for this proteolysis (34). Studies
on human cervical secretions (30) indicated that the carbohy-
drate-free segment would be particularly sensitive to such
endogenous proteolysis. This probably explains why the yield
of cysteinyl peptides was lower than that expected from the
recovery of glycopeptides. A ratio of cysteinyl peptides to
glycopeptides of about 0.43 (Table III) was expected, whereas
a ratio of 0.23 was obtained (Fig. 2). The free carboxymethyl-
cysteine which contaminated the cysteinyl peptides was pre-

**FIG. 7.** Nephelometric study of the effect of various dissociating
reagents on BCM. DTT, dithiothreitol; Guan HC1, guanidinium
chloride; Doc, deoxycholate.

**FIG. 8.** Structure of the BCM subunit. The glycopeptides and the
cysteinyl peptide represent the NH$_2$-terminal and COOH-terminal
parts, respectively, of the BCM subunit.
From the amount of the COOH-terminal amino acid, leucine, a yield of 19% for the carboxypeptidase digestion. This yield the molecular weight of the subunit would be 27,000 assuming contains 2 leucine residues (Table III).

The NH₂-terminal amino acid, alanine, the molecular weight, residues found in the whole mucin. From the quantity of the depending on the experiment, would be 16,300 or 30,000. Small subunits of 21,700. The existence of such small subunits is also suggests that BCM is composed of similar, if not identical, side chains located near the COOH-terminal residue.

Table III

| Amino acid   | Whole mucin | Glycopeptides | Cysteinyl peptides |
|--------------|-------------|---------------|--------------------|
| Aspartic acid| 4           | 0-1           | 3-4                |
| Threonine    | 16          | 14-15         | 1-2                |
| Serine       | 7           | 5-6           | 1                  |
| Glutamic acid| 6           | 2-3           | 2                  |
| Proline      | 5-6         | 5             | 1                  |
| Glycine      | 4-5         | 2             | 1-2                |
| Alanine      | 5           | 5             | 1                  |
| Cysteine     | 2           | 0             | 1-2                |
| Valine       | 6           | 3-4           | 1                  |
| Methionine   | 0-1         | 0-1           | 0                  |
| Isolaurine   | 2           | 1             | 1                  |
| Leucine      | 3           | 1             | 1-2                |
| Tyrosine     | 1           | 0             | 1                  |
| Phenylalanine| 1           | 0             | 1                  |
| Lysine       | 1           | 0-1           | 0                  |
| Histidine    | 0-1         | 0             | 0                  |
| Arginine     | 2           | 1             | 1                  |
| 66-70        | 39-45       | 16-21         |                    |

* From Bhushana Rao et al. (4).

sumably released from BCM before its isolation by the proteases occurring in the cervical secretion.

The glycopeptides recovered from BCM were resolved into two fractions, which displayed a diffuse pattern on HV electrophoresis at pH 9.2. Without excluding slight differences in the amino acid composition, we believe that the separation of the two glycopeptides on Dowex 50-X2 was mainly due to differences in their carbohydrate contents. This would also account for the continuous variation observed in electrophoresis. The COOH-terminal amino acid of glycopeptide I, as expected, was a basic amino acid, i.e. arginine. That of glycopeptide II, although less clearly identified, was also presumably arginine. The different behavior towards carboxypeptidase A and B of glycopeptides I and II, despite their similar amino acid compositions, might be related to differences in their carbohydrate contents. The activity of the enzymes on glycopeptide II could be impeded by polyanionic side chains located near the COOH-terminal residue.

The cysteinyl peptides resolved into nine bands on HV electrophoresis. Slight differences in amino acid composition, especially in the proportions of amides, could explain these differences of mobility.

The fact that only one tryptic cleavage was necessary to break the molecule into two peptides of 45 and 19 residues suggests that BCM is composed of similar, if not identical, subunits of 64 amino acid residues. The sum of the amino acid residues and the carbohydrates gives a molecular weight of 21,700. The existence of such small subunits is also supported by the large numbers of NH₂- and COOH-terminal residues found in the whole mucin. From the quantity of the NH₂-terminal amino acid, alanine, the molecular weight, depending on the experiment, would be 16,300 or 30,000. From the amount of the COOH-terminal amino acid, leucine, the molecular weight of the subunit would be 27,000 assuming a yield of 19% for the carboxypeptidase digestion. This yield was estimated from the assumption that the BCM subunit contains 2 leucine residues (Table III).

The sharp decrease in viscosity of cervical mucus caused by reducing agents suggests that the cysteinyl residues play a role in the binding of the subunits. However, the reduction with dithiothreitol did not decrease the nephelometric effect of BCM. Hence, linkages other than disulfide bridges are presumably involved. The nephelometric experiments showed that detergents are the most efficient dissociating agents. Hydrophobic bonds, therefore, appear important in the association of the subunits. The finding of a cluster of hydrophobic amino acid residues in the COOH-terminal part of the cysteinyl peptide agrees with this suggestion. However, the minimal molecular weight of 510,000 obtained for BCM in the presence of SDS, greatly exceeds the value of 220,000 proposed for the subunit. Gibbons et al. (35) reported a minimal molecular weight for BCM of 550,000. We have no clear explanation for the discrepancy between the biochemical and physical estimations of the BCM minimal molecular size, but we suspect that, as in some membrane proteins, the hydrophobic linkages are so strong that their disruption requires special dissociating agents (36).

A study of the mucin from a human ovarian cyst (37) has recently confirmed the subunit structure described here for BCM. Some similarities of the BCM subunit with the stamaglycoprotein of human erythrocyte membrane, the so-called glycoporin are worth mentioning (38). The latter contains a NH₂-terminal segment rich in threonyl and seryl residues carrying the carbohydrate side chains. Beyond the 50th residues, the molecule is devoid of carbohydrate. This "naked" segment is 4 to 5 times longer in glycoporin than in BCM, but it contains as does the naked segment of BCM a cluster of hydrophobic amino acid residues.

Acknowledgments—Our thanks are due to Dr. C. Meeus for providing bovine secretions and to Mr. E. Van Roost who performed the ultracentrifugations. The competent editorial assistance of Mr. M. Delory is gratefully acknowledged.

REFERENCES
1. Gibbons, R. A., and Roberts, G. P. (1963) Ann. N. Y. Acad. Sci. 196, 218–232
2. Gibbons, R. A. (1969) in Protides of the Biological Fluids (Peters, H., ed.) Vol. 16, pp. 299–305, Pergamon Press, Inc., Elmsford, N. Y.
3. André, F., André, C., Bhushana Rao, K. S. P., Masson, P. L., and Heremans, J. F. (1972) Carbohydr. Res. 25, 395–401
4. Bhushana Rao, K. S. P., Van Roost, B., Masson, P. L., Heremans, J. F., and André, F. (1973) Biochim. Biophys. Acta 317, 296–302
5. Pugman, W., Fayza, N., Moschera, J., and Weiss, M. (1970) Fed. Proc. 29, 599
6. Gibbons, R. A., and Glover, F. A. (1959) Biochem. J. 73, 217
7. Gibbons, K. A. (1966) in Glycoproteins: Their Composition, Structure and Function (Gottschalk, A., ed) p. 29, Elsevier, Amsterdam
8. Bellingter, J. F., and Huist, N. K. M. (1973) J. Chromatog. 77, 513–522
9. Schroeder, W. A. (1967) Methods Enzymol. 11, 351–361
10. Schroeder, W. A. (1967) Methods Enzymol. 11, 361–369
11. Ambler, R. P. (1963) Biochem. J. 90, 341–349; 349–378
12. Waley, S. G., and Watson, J. (1953) Biochem. J. 55, 328–373
13. Spackman, D. H., Stein, W. H., and Moore, S. (1958) Anal. Chem. 30, 1190–1206
14. Hartley, B. S. (1970) Biochem. J. 119, 805–822
15. Levy, A. L., and Li, C. H. (1955) J. Biol. Chem. 213, 487–497
16. Wang, K.-T., and Wang, I. S. Y. (1967) J. Chromatog. 27, 318–320
17. McKean, D., Potter, M., and Hood, L. (1973) Biochemistry 12, 749–759
18. Gray, W. R. (1972) Methods Enzymol. 25B, 121–138
19. Woods, K. R., and Wang, K.-T. (1967) Biochim. Biophys. Acta
Bovine Cervical Mucin

199, 260-270
20. Pisano, J. J., and Bronzert, T. J. (1969) J. Biol. Chem. 244, 5697-5697
21. Summers, M. R., Smythers, G. W., and Oroszlan, S. (1973) Anal. Biochem. 53, 624
22. Van Holde, K. E. (1967) Fractions, Vol. 1, pp. 1-10, Spinco Division of Beckman Instruments, Inc., Palo Alto, Calif.
23. Chervenka, C. H. (1970) Anal. Biochem. 34, 24-29
24. Kratky, O., Leopold, H., and Stabinger, H. (1969) Z. Angew. Phys. 27, 273-277
25. Densuelle, P. (1966) in The Enzymes (Boyer, P. D., Lardy, H., and Myrback, K., eds) Vol. 4, pp. 119-132, Academic Press, N. Y.
26. Adachi, S. (1965) J. Chromatogr. 17, 295-299
27. Strid, L. (1973) FEBS Lett. 33, 192-196
28. Marrack, J. R., and Grant, R. A. (1953) Brit. J. Exp. Pathol. 34, 263
29. Gelman, R. A., and Vered, J. (1976) Biochim. Biophys. Acta 427, 627-633
30. Bhushana Rao, K. G. P., Barbier, B., Masson, P. L., Heremans, J. F., and Ferin, J. (1976) The Human Semen and Fertility Regulation in the Male (Hafez, E. S. E., ed) pp. 237-242, The C. V. Mosby Co., Saint Louis, Mo.
31. Goodwin, S. D., and Watkins, W. M. (1974) Eur. J. Biochem. 47, 371-382
32. Kristiansen, T. (1974) Biochim. Biophys. Acta 338, 374-381
33. Winzler, R. J., and Bocci, V. (1972) in Glycoproteins: Their Composition, Structure and Function (Gotschalk, A., ed), Vol. 5, Part B, pp. 1228-1245, Elsevier, Amsterdam
34. Bhushana Rao, K. S. P., Roberts, T. K., Masson, P. L., and Heremans, J. F. (1973) J. Reprod. Fertil. 32, 89-92
35. Gibbons, R. A., Boyd, L. J., Dixon, S. N., Parker, J., Sellwood, R., and Tasker, J. B. (1973) J. Reprod. Fertil. 35, 469-476
36. Marchesi, V. T., and Andrews, E. P. (1971) Science 174, 1247-1248
37. Huberty-Barbier, B. (1977) Ph.D. thesis, University of Louvain
38. Tomita, M., and Marchesi, V. T. (1975) Proc. Natl. Acad. Sci. U. S. A. 72, 2964-2968
Study of the primary structures of the peptide core of bovine estrus cervical mucin. Possible existence of small similar subunits.

K S Rao and P L Masson

J. Biol. Chem. 1977, 252:7788-7795.

Access the most updated version of this article at http://www.jbc.org/content/252/21/7788

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/252/21/7788.full.html#ref-list-1