THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 246, No. 2, Issue of January 25, pp. 305-308, 1971
Printed in U.S.A.

Primary Structure of Freezing Point-depressing Glycoproteins*

(Received for publication, August 31, 1970)

ARTHUR L. DEVRIES,‡ JACKIE VANDENHEEDE, AND ROBERT E. FEENEY
From the Department of Food Science and Technology, University of California, Davis, California 95616

SUMMARY

The freezing point-depressing glycoproteins, which are present in the blood of the Antarctic fish, *Trematomus borchgrevinki*, were sequenced. The first 6 residues of the native glycoprotein and the β-eliminated protein were identified by phenyl isothiocyanate degradation. The native glycoproteins were hydrolyzed into several small glycopeptides by the proteolytic enzymes subtilisin or elastase. The resulting hepta, hexa, penta, tetra, tri, and diglycopeptides were separated on Dowex 50 and were sequenced by phenyl isothiocyanate degradation. The sequences of these glycopeptides showed that the molecule contained the same repeating unit found in the first 6 residues of the unhydrolyzed protein. The repeating structural unit is a tripeptide of alanine-alanine-threonine in which the last residue is glycosidically linked to the disaccharide galactosyl-*N*-acetylgalactosamine.

The fish *Trematomus borchgrevinki* is a common inhabitant of McMurdo Sound, Antarctica, and experiences water temperatures near the freezing point of the seawater there (−1.9°). However, the fish does not freeze even when the water is seeded with ice crystals. The freezing point of its blood is approximately −2.0° (1), which is 1.3° lower than that of the blood of most temperate marine fishes. Thirty percent of the serum freezing point depression is due to nondialyzable solutes, most of which are three glycoproteins of relatively high molecular weights (1). These glycoproteins have been isolated and their physical and chemical properties studied (1–3). They have been termed freezing point-depressing glycoproteins and at low concentrations depress the freezing point of water to the same extent as NaCl on a weight basis (1, 2). They have been isolated from the blood serum and separated into three distinct groups based on results from ion exchange chromatography and analytical acrylamide gel electrophoresis. These three glycoproteins, which are referred to as glycoproteins 3, 4, and 5, differ in molecular weights (10,500, 17,000, and 21,500), but have identical compositions on a weight basis (2). They are composed primarily of threonine (16%), alanine (23%), *N*-acetylgalactosamine (29%), and galactose (28%), and the different forms occur in equal amounts in the blood of this fish. Studies of the structure of the glycoproteins have shown that all threonines are linked glycosidically to C-1 of the *N*-acetylgalactosamine of a galactosyl-*N*-acetylgalactosamine disaccharide (3). Periodate oxidations indicate that the internal linkage of the disaccharide is a 1,4 rather than a 1,3 or 1,6.

The fact that the composition of the glycoproteins is so simple and that the ratio of alanine to threonine is 2:1, leads one to suggest that the glycoproteins are composed of a basic repeating triglycopeptide whose structure is Ala–Ala–Thr–*O*-disaccharide. This paper presents data which supports this structure.

EXPERIMENTAL PROCEDURE

Materials—Specimens of *T. borchgrevinki* weighing between 50 and 100 g were caught through holes in the ice of McMurdo Sound, Antarctica, and their blood was collected with a hypodermic syringe from the heart as described previously (2). All blood samples were allowed to clot for 4 hours at 0° and then centrifuged for 10 min at 2000 × *g*, the sera collected, and stored at −20° until used. The FPD glycoproteins[1] were isolated from the blood sera using DEAE-cellulose ion exchange chromatography, and the various glycoproteins were then separated using QAE-Sephadex A-25 ion exchange chromatography according to the procedures recently described (2).

Subtilisin (Nagarse) and elastase were purchased from Miles Labs, Elkhart, Indiana, and Worthington Biochemical Company, respectively. The anhydrous trifluoroacetic acid, phenyl isothiocyanate, and 1-fluoro-2,4-dinitrobenzene were purchased from Eastman Organic Chemicals. Anhydrous hydrazine (97%+), ethyl ether, and pyridine were obtained from Mallinckrodt Chemical Works. Dinitrophenyl L-alanine and dinitrophenyl L-threonine were purchased from Nutritional Biochemicals. All other reagents were of analytical grade. The cellulose thin layer chromatography sheets with fluorescent indicator were purchased from Eastman Kodak Company, Distillation Products Industries, New York.

Enzymatic Digestion with Subtilisin—Hydrolysis of the FPD glycoproteins by subtilisin was done at pH 8.2 in 0.02 M $NaHCO_3$ containing 0.01 M $CaCl_2$. The concentration of FPD glycoprotein was 30 mg per ml while that of subtilisin was 1% by weight of the glycoprotein. Incubations for 12 and 24 hours at 37° resulted in completely soluble digests. In order to avoid any losses of material, the entire digest was fractionated using

* This work was supported in part by Grant GA-12607 from the Environmental Sciences Division of the National Science Foundation. Presented in part at the Pacific Slope Biochemical Conference, San Diego, California, June 24 to 26, 1970.

‡ Present address, Scripps Institution of Oceanography, University of California, San Diego.

[1] The abbreviation used is: FPD glycoprotein, freezing point-depressing glycoprotein.

a Technicon amino acid analyzer (Technicon, Ardsley, New York). Usually about 30 mg of digested glycoprotein was applied directly on top of the resin of a column (0.6 × 140 cm) of Dowex 50 (Technicon Chromasorb, A type) in a volume of 1 ml and eluted. Elution was with a sodium citrate buffer (pH 2.3) which contained 27.4 g of sodium citrate, 61.4 ml of 6 N HCl, and 63.0 ml of 2 N sodium hydroxide in 2 liters of water. This was also the starting buffer for the subsequent gradient elution going to a pH 5 buffer containing 29.4 g of sodium citrate, 70.1 g of NaCl, 50 ml of 2 N NaOH, and 26.6 ml of 6 N HCl in 2 liters of water. The nonlinear gradient was obtained by filling the Autograd of the Technicon amino acid analyzer as used for a standard 21-hour determination. The effluent (95%) was diverted to a fraction collector. The fractions containing the glycopeptides were consolidated, lyophilized, and desalted on a Sephadex G-10 column. The purified glycopeptides were analyzed for amino acids and then sequenced. The presence of carbohydrate in the peptides was determined by the methods recently employed (2).

Enzymatic Digestion with Elastase—Digestion of the FPD glycoproteins with elastase was done in 0.15 M $NaHCO_3$ buffer containing 0.1 M NaCl at pH 8.8. The concentration of FPD glycoprotein was 10 mg per ml, and the ratio of FPD glycoprotein to elastase was 50:1 on a weight basis. Incubations were at 37° for 12 hours. The reaction was stopped by lowering the pH to 2 with 10% perchloric acid. At this low pH, the enzyme is very rapidly denatured and precipitates, and therefore must be removed before applying to the analyzer. After removal of the precipitate by centrifugation, the supernatant fluid was desalted on Sephadex G-10 and lyophilized. The resulting peptides were separated on Dowex 50 and desalted again on Sephadex G-10. Determination of the yield of peptides was done on the basis of the weight of desalted peptides. Alanine was determined colorimetrically because of the difficulty of separating it from the buffer salt. The structure of the purified glycopeptides was identified by the same methods used in the subtilisin digestion experiment.

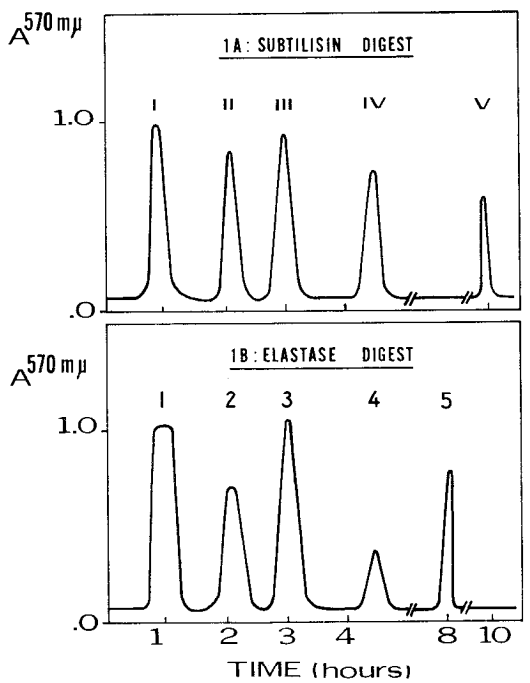

FIG. 1. *A*, elution profile obtained when 2 mg of FPD glycoprotein was treated for 12 hours at 37° with subtilisin and chromatographed on a column (0.6 × 140 cm) of Dowex 50 attached to a Technicon amino acid analyzer. Even when 30 mg of subtilisin digest was applied to the column no components other than those in the five peaks were noted. *B*, elution profile obtained when FPD glycoproteins were treated for 12 hours at 37° with elastase and chromatographed on a column (0.6 × 140 cm) of Dowex 50 attached to Technicon amino acid analyzer.

Amino Acid Analyses—The glycopeptides were hydrolyzed at 110° for 22 hours with 6 N HCl under reduced pressure according to the method of Moore and Stein (4). Amino acid analyses were made with a Technicon Autoanalyzer (Ardsley, New York) with a standard 21-hour run. Loss of threonine in a 22-hour hydrolysis was 4%.

Amino- and Carboxy-terminal Amino Acid Determinations—The NH_2-terminal amino acids of the native glycoproteins (glycoproteins 3, 4, and 5 (2)), the β-eliminated proteins, and the glycopeptides which were obtained from subtilisin digestion were identified as their dinitrophenol derivatives (5). The dinitrophenylamino acids were separated and identified by thin layer chromatography with a benzene-pyridine-acetic acid (80:20:2, v/v) solvent system. Since the FPD glycoproteins contain only alanine and threonine and their dinitrophenyl derivatives separate widely, there was never any question as to the identification of the products.

Hydrazinolysis was employed for identification of the COOH-terminal amino acids of the glycoproteins and glycopeptides. The native glycoproteins and glycopeptides were reacted with anhydrous hydrazine for 24 hours at 60° according to the method of Fraenkel-Conrat and Tsung (6). The COOH-terminal amino acids liberated in this reaction were separated from the majority of the hydrazides by chromatography on Amberlite IRC-50 (H^+ form) and were identified on the amino acid analyzer.

Determination of Sequences of Amino Acids—Initial attempts to sequence from the NH_2-terminal end with the Edman degradation as described by Sjoquist, Blomback, and Wallen (7) were unsuccessful with the native FPD glycoprotein (mixtures of glycoproteins 3, 4, and 5). This difficulty was probably due to interference by the large amounts of carbohydrate.

In order to avoid the apparent interference, sugar-free protein was prepared by base elimination. β-Elimination with 0.2 M NaOH (3) yielded practically sugar-free protein with no detectable splitting of peptide bonds. The solution of FPD glycoproteins (glycoproteins 3, 4, 5) in 0.2 M NaOH was incubated at 45° for 6 hours and the reaction stopped by neutralization with HCl. Part of the protein precipitated, and the precipitate was separated by centrifugation and washed several times with distilled water. The supernatant fluid was desalted by passing it through a Sephadex G-10 column. Both the precipitated and soluble fractions were subjected to Edman degradation using the sequence analysis of Sjoquist *et al.* (7). The phenylthiohydantoin amino acids were identified by thin layer chromatography on silica gel plates, which were developed in the solvent system pyridine-heptane (30:70, v/v), and the spots were made visible by spraying with a 1:1 mixture of iodine and sodium azide solutions.

The method of Gray and Smith (8) was used for the determination of the sequences of both native FPD glycoproteins and glycopeptides resulting from subtilisin or elastase digestion. The NH_2-terminal amino acids were removed by phenylisothiocyanate degradation and the new NH_2-terminal residue determined by dansylation or by making the dinitrophenol derivative with

TABLE I

Sequences of glycopeptides from subtilisin hydrolysis of freezing point-depressing glycoprotein

Glycopeptides corresponding to peaks[a]	Alanine to threonine ratio	NH_2-terminal as dinitrophenol residue	COOH-terminal (hydrazinolysis)	Sequences
I	1.9	Ala, Thr	Ala	Ala–Thr–Ala–Ala–Thr–Ala–Ala Ala–Ala–Thr–Ala–Ala–Thr–Ala Ala–Thr–Ala–Ala–Thr–Ala Thr–Ala–Ala–Thr–Ala–Ala Thr–Ala–Ala–Thr–Ala
II	3.0	Ala	Ala	Ala–Thr–Ala–Ala
III	2.0	Ala, Thr	Ala	Ala–Thr–Ala Thr–Ala–Ala
IV	1.1	Thr	Ala	Thr–Ala
V		Ala	Ala	Ala–Ala

[a] Peak fractions from Fig. 1*A*.

TABLE II

Identification of glycopeptides of Peak I of subtilisin hydrolysate

Number of Edman cycles	Amino acid removed	Resulting glycopeptide peaks[a]	Identified glycopeptide sequences	Designated structures of original glycopeptides
0		I		
1	(Ala) (Thr)	I		
	(Thr)	I′	Ala–Ala–Thr–Ala–Ala	Thr–Ala–Ala–Thr–Ala–Ala
	(Thr)	II	Ala–Ala–Thr–Ala	Thr–Ala–Ala–Thr–Ala
2	(Ala, Ala) + ···	I	Thr–Ala–Ala–Thr–Ala + ···[b]	Ala–Ala–Thr–Ala–Ala–Thr–Ala
	(Ala, Thr)	I′	Ala–Ala–Thr–Ala–Ala	Ala–Thr–Ala–Ala–Thr–Ala–Ala
	(Thr, Ala)	II	Ala–Thr–Ala–Ala	Thr–Ala–Ala–Thr–Ala–Ala

[a] Peaks produced on chromatography after removal of NH_2-terminal amino acids. Numerical designations correspond to those from patterns of Fig. 1 and sample of Table I.

[b] Small amounts of longer peptides still present.

Sanger's reagent. The dansyl amino acids were identified on polyamide sheets, with the use of a solvent of benzene-acetic acid (9:1, v/v), while the dinitrophenylamino acids were identified by thin layer chromatography on silica gel plates in the solvent system benzene-pyridine-acetic acid (80:20:2, v/v).

RESULTS

Sequence of Native and β-Eliminated Freezing Point-depressing Glycoproteins—Both NH_2-terminal and COOH-terminal residues of all three FPD glycoproteins (3, 4, and 5) were found to be alanine. The one free amino group previously reported (3) was thus shown to be from NH_2-terminal alanine rather than from unacetylated galactosamine.

With the sequence analysis of Gray and Smith (8), the sequence of the first six amino acids of the native FPD glycoproteins was found to be Ala–Ala–Thr–Ala–Ala–Thr–. The identical sequence was also obtained with the β-eliminated peptide by the method of Sjoquist *et al.* (7). By this latter method, threonine was identified as the Δ-threonine phenylthiohydantoin, also named the anhydrothreonine derivative. Both the precipitate and soluble fractions gave identical sequences.

Sequences of Glycopeptides from Subtilisin Treatment of Freezing Point-depressing Glycoproteins—After 12 hours of digestion with subtilisin, the glycoproteins were hydrolyzed into a number of glycopeptides which were separated into five distinct groups with the use of a Dowex 50 column of a Technicon amino acid analyzer.

TABLE III

Sequences of glycopeptides from elastase hydrolysis of freezing point-depressing glycoprotein

Peak[a]	Yield[b]	Alanine to threonine ratio	Glycopeptides
	mg		
1	40	2:1	Ala–Ala–Thr–Ala–Ala–Thr–Ala + ···[c] Ala–Thr–Ala–Ala–Thr–Ala
2	6	3:1	Ala–Ala–Thr–Ala Ala–Thr–Ala–Ala
3	10	2:1	Ala–Thr–Ala
4	2	1:1	Thr–Ala
5	(5)[d]	1:0	Ala

[a] Peak fractions from Fig. 1*B*.

[b] Protein (100 mg) was treated with elastase. After removal of elastase, desalting, and lyophilization, 71 mg of dry salt-free peptides were obtained and fractionated.

[c] Small amount (*ca.* 5 mg) of longer peptides also present.

[d] The 5 mg of alanine were in the salt fraction. This represents the lower value of an estimate because of the difficulty of accurate measurement in the presence of a high concentration of salt.

Fig. 1*A* shows the separation obtained when a 12-hour digest was applied to the column. The peaks on the chromatogram have been designated as Peaks I, II, III, IV, and V. After isolation of the glycopeptides corresponding to Peak I on the chromatogram, further digestion (24 hours) converted these glycopeptides to smaller ones which were separated on Dowex 50 into glycopeptides corresponding to Peaks II, III, IV, and V on the chromatogram.

With the use of the Dowex 50 column of the Technicon amino acid analyzer, several milligrams of the five classes of glycopeptides were separated and their amino acid composition and sequence determined.

The ratios of alanine to threonine in the glycopeptides are given in Table I. The NH_2- and COOH-terminal amino acids of the glycopeptides corresponding to the chromatogram Peaks I through V are also given in Table I. All COOH-terminals were found to be alanine while the NH_2-terminals were both alanine and threonine.

Material in Peak V was identified as the dipeptide Ala–Ala. By phenylisothiocyanate degradation coupled with identification of the new NH_2-terminal, the glycopeptide corresponding to Peak IV was found to be a diglycopeptide of galactosyl-*N*-acetylgalactosamine-threonyl-alanine (Table I). Peak III was found to contain two triglycopeptides both having alanine at the COOH-terminal end but one having alanine and the other threonine at the NH_2-terminal. After one phenylisothiocyanate degradation, rechromatography of this mixture yielded equal amounts of one glycopeptide which corresponded to Peak IV and a dipeptide Ala–Ala. Therefore, two original triglycopeptides must have been present in equal amounts in the subtilisin digest. Peak II was found to be homogenous and contained the tetraglycopeptide with the sequence Ala–Thr–Ala–Ala. After the first phenylisothiocyanate degradation of this glycopeptide, the new NH_2-terminal was threonine. After a second phenylisothiocyanate degradation, rechromatography of the remaining peptide on Dowex 50 yielded the dipeptide Ala–Ala.

When those glycopeptides corresponding to Peak I were treated with phenylisothiocyanate and the NH_2-terminal cleaved with acid, the resulting mixture of glycopeptides, when chromatographed on Dowex 50, corresponded to Peaks I and II as well as a peak which eluted intermediate to I and II (Peak I′). The sequence of the tetraglycopeptide corresponding to Peak II in this case had a sequence of Ala–Ala–Thr–Ala. This could have only come from an original pentapeptide having the amino acid sequence, Thr–Ala–Ala–Thr–Ala. A single phenylisothiocyanate degradation of the new glycopeptide (Peak I′), whose NH_2-terminal was alanine, produced a tetraglycopeptide whose sequence was Ala–Thr–Ala–Ala. Therefore the sequence of this pentaglycopeptide (Peak I) was Ala–Ala–Thr–Ala–Ala, coming originally from the hexapeptide, Thr–Ala–Ala–Thr–Ala–Ala. After a second round of phenylisothiocyanate degradation most of the glycopeptides corresponding to Peak I, when chromatographed on Dowex 50, were eluted in the position of Peak II. Sequence studies indicated the presence of a tetraglycopeptide of Ala–Ala–Thr–Ala. A few of the glycopeptides eluted in positions corresponding to Peaks I and I′ indicating that they were originally heptaglycopeptides. After the second phenylisothiocyanate degradation of the original Peak I the amounts of these glycopeptides which were recovered as glycopeptides corresponding to Peaks I and I′ were too small for sequence studies. This is illustrated in Table II.

Sequence of Peptides from Elastase Treatment of Freezing Point-depressing Glycoproteins—After 12 hours of incubation of FPD glycoproteins with elastase, the product was a mixture of glycopeptides which chromatographed into five distinct peaks on the Dowex 50 column (Fig. 1*B*). Peaks 1, 2, 3, and 4 eluted with the initial pH 2.3 buffer. Peak 5 eluted with the nonlinear pH and salt gradient and was identified as alanine.

The sequence of the first four peaks was determined by the method of Gray and Smith (8) (Table III). Hydrazinolysis showed that all COOH-terminals were alanine. When the glycopeptides of Peak 1 were redigested several times with elastase, they were completely hydrolyzed into peptides of Peaks 2, 3, and 4 and into alanine. Phenylisothiocyanate degradation of peptides of Peaks 2, 3, and 4 changed them into the peptides of the appropriate larger numbered peaks, in a similar way as described for subtilisin fractions.

DISCUSSION

The fact that the FPD glycoprotein appears in the blood of Antarctic fishes in a number of sizes (mol wt 10,500, 17,000, and 23,500) led DeVries *et al.* (2) to postulate that the various sizes of glycoproteins differ from each other only in the number of a basic repeating unit. This unit was thought to be composed of Ala–Ala–Thr to which the galactosyl-*N*-acetylgalactosamine disaccharide was attached. Six sequential degradations on both the native and β-eliminated glycoproteins beginning from the NH_2-terminal prove the repeating unit hypothesis for the first part of the molecules. Although the sequence studies of the glycopeptides isolated from the subtilisin and elastase digests are not direct proof of the repeating structure, they definitely lend credence to it. In the larger glycopeptides (hepta, hexa, penta, and tetra) there were never more than two alanine residues found adjacent to each other, nor were there ever threonine residues found separated by less than two alanine residues. Although the COOH-terminal residue has been identified as alanine the identity of the amino acid adjacent to it is not known because it was not possible to identify those peptides containing the COOH-terminal residue.

The pattern of peptides which are formed upon treatment with subtilisin can be explained by the fact that this enzyme cleaves between alanine residues and alanine and threonine residues but never between threonine and alanine residues of the glycoproteins. The peptides obtained after elastase digestion suggest that the enzyme cleaves the FPD glycoproteins preferentially between two alanine residues, resulting in the peptides of Peaks 1, 2, 3, and 5. Eventually the alanine-threonine peptide bond is also broken, but at a much slower rate, resulting in the small amount of the dipeptide Thr–Ala (Peak 4).

REFERENCES

1. DeVries, A. L., and Wohlschlag, D. E., *Science*, **163**, 1073 (1969).
2. DeVries, A. L., Komatsu, S. K., and Feeney, R. E., *J. Biol. Chem.*, **245**, 2901 (1970).
3. Komatsu, S. K., DeVries, A. L., and Feeney, R. E., *J. Biol. Chem.*, **245**, 2909 (1970).
4. Moore, S., and Stein, W. H., in S. P. Colowick and N. O. Kaplan (Editors), *Methods in enzymology, Vol. 6*, Academic Press, New York, 1963, p. 819.
5. Sanger, F., *Biochem. J.*, **39**, 507 (1945).
6. Fraenkel-Conrat, H., and Tsung, C. M., in S. P. Colowick and N. O. Kaplan (Editors), *Methods in enzymology, Vol. 11*, Academic Press, New York, 1967, p. 151.
7. Sjoquist, J., Blomback, B., and Wallen, P., *Ark. Kemi*, **16**, 425 (1961).
8. Gray, W. G., and Smith, J. F., *Anal. Biochem.*, **33**, 36 (1970).