EVIDENCE FOR OCCURRENCE OF A CHARACTERISTIC AMINO ACID SEQUENCE OF GLYCOPEPTIDES IN THE LINKAGE REGION BETWEEN PEPTIDE AND CARBOHYDRATE

Mamoru Isemura*, Tokuji Ikenaka* and Yoshio Matsushima

Department of Chemistry, Osaka University College of Science, Toyonaka, Osaka, Japan

Received November 1, 1971

SUMMARY: The amino acid sequences of the several glycopeptides derived from the cuttlefish skin collagen were established and compared with those derived from vertebrate collagens. The result indicates that there exists the common amino acid sequence which seems to act as a recognition sign for the glycosyltransferase to attach the first sugar to the polypeptide chain.

In a course of the study of elucidating a chemical nature of Cuverian tubules from a sea cucumber, Holothuria folskali (1), we obtained evidence which might be inconsistent with some of suggestions given by Morgan et al. (2), relating to the glycosylation mechanism in the hydroxylysine-containing polypeptide chain. This led us to further investigation on the glycopeptides derived from invertebrate collagens and indeed, we could confirm the previous findings obtained from the tubules.**On the other hand, this study provided evidence for occurrence of the possible recognition sign made of specific amino acid sequence for a glycosyltransferase in invertebrate as well as in vertebrate as pointed out by above authors (2).

---

* Present address: Department of Biochemistry, Niigata University School of Medicine, Niigata, Japan

** The results will be discussed elsewhere.

MATERIALS AND METHODS

Skins of cuttlefish, Sepia subaculeata, were defatted with acetone, cut into small pieces, allowed to swell in water and minced with the aid of a blendor. The preparation was successively washed with 10 % NaCl, 0.5 M Na_2HPO_4 and acetone, then finally air-dried. The insoluble skin collagen thus obtained was digested with Pronase P at pH 8, care being taken to maintain this pH by adding $Ba(OH)_2$ during a course of 24 hours. The lyophilyzed digestion products were subjected to gel filtration through Sephadex G-50 yielding two hexose-containing fractions as evidenced by the phenol-H_2SO_4 method (Fig.1). The main hexose-containing fraction (see Fig. 1) was subsequently fractionated on a column of Dowex 50-X2 (3) yielding four major glycopeptide fractions. These major fractions were further fractionated by a combination of Bio-Gel P-4, Dowex 1-X2 and Bio-Gel P-2, yielding eight glycopeptides homogeneous as judged by high voltage electrophoresis at pH 3.6 and by amino acid analysis. Six glycopeptides out of these were obtained in an amount enough for structural studies. All glycopeptides thus isolated were shown gas-chromatographically to have

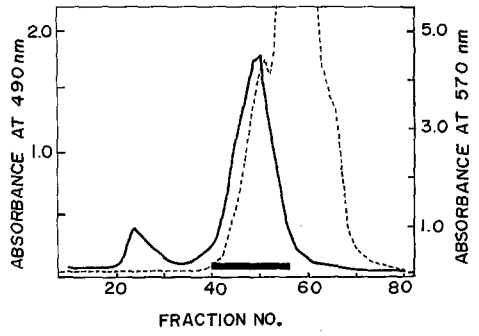

Fig. 1. Gel filtration of the Pronase digest (3 g.) of the cuttlefish skin collagen through Sephadex G-50 (2.8 x 95 cm). The column, previously equilibrated with 0.1 M pyridine-acetate, pH 6.0, was eluted with the same buffer. Each fraction contained 10 ml. A portion (0.2 ml) of each fraction was analyzed by the phenol-H_2SO_4 method measured at 490 nm (——) and 0.05 ml by the ninhydrin procedure at 570 nm (----). Fractions 40-56 as indicated were collected to prepare the glycopeptides.

one mole each of glucose and galactose per mole and gave the hydroxylysine-disaccharide upon alkaline hydrolysis as has been obtained previously (1,4).

The amino acid sequence of each glycopeptide was determined by the direct Edman procedure (5), thin layer chromatography being used to identify the PTH-amino acid including PTH-hydroxyproline. When serine was expected from the above procedure, subtractive Edman method (6) was also utilized for the confirmation. When Edman cycle was repeated up to the step just before the last as expected from the amino acid composition, the reaction residue was subjected to the analysis by an automatic amino acid analyzer to identify a carboxy-terminal amino acid.

Carboxypeptidase B was also used to confirm the C-terminal arginine of Glycopeptide II to VI and quantitative release of this amino acid was observed in every case. When neither absorption at 270 nm expected for PTH-amino acid, nor any spot on a thin layer chromatogram was observed, hydroxylysine carrying carbohydrate was tentatively assigned and confirmed finally after all other amino acid residues were positioned properly.

Table I. The amino acid composition of the cuttlefish skin preparation

Amino acid	Residues per 100,000 g.	Amino acid	Residues per 100,000 g.
Hydroxyproline	71.4	Methionine	17.9
Aspartic acid	61.6	Isoleucine	21.4
Threonine*	25.4	Leucine	30.6
Serine*	42.1	Tyrosine	11.1
Glutamic acid	76.6	Phenylalanine	14.7
Proline	90.4		
Glycine	277.3	Hydoxylysine	10.2
Alanine	73.0	Lysine	15.1
Cystine	0	Histidine	8.7
Valine	23.0	Arginine	60.7

* not corrected for loss occurring during hydrolysis

Table II. The structures of hydroxylysine-containing glycopeptides derived from vertebrate and invertebrate collagens

Occurrence	Glycopeptide	Amino acid sequence
Cuttlefish skin	I	Gly-Ala-Hyl*-Gly-Asp
	II	Hyl*-Gly-Asp-Arg
	III	Gly-Ala-Hyl*-Gly-Asp-Arg
	IV	Asp-Gly-Ser-Hyp-Gly-Glu-Hyl*-Gly-Ala-Arg
	V	Gly-Phe-Hyp-Gly-Ile-Hyp-Gly-Gln-Hyl*-Gly-Ala-Arg
	VI	Ser-Gly-Pro-Hyl*-Gly-Ala-Arg
Human skin(2)	Peak II	Gly-Phe-Hyl*-Gly-Ile-Arg
Human skin, guinea pig skin, and carp swim bladder(2)	Peak III and V	Gly-Met-Hyl*-Gly-His-Arg
Carp swim bladder(2)	Peak IV	Gly-Ile-Hyl*-Gly-His-Arg

* A disaccharide unit, glucosylgalactosyl residue is attached to hydroxylysine

RESULTS AND DISCUSSION

The amino acid composition shown in Table I indicates that this preparation is mainly composed of the collagenous protein. The structures of the glycopeptides derived from the cuttlefish collagen are summarized in Table II together with those of the glycopeptides from selected vertebrate collagens (2). As seen clearly from this table, the common sequence of Gly-X-Hyl-Gly-Y-Arg found in vertebrate collagens is held in the cuttlefish collagen. Thus this study would extend the rule found in vertebrate (2) to invertebrate collagens, hence to a collagen in general.

There has been known evidence for occurrence of the particular amino acid sequence of Asn-X-Thr or Ser in the glycopeptides where

Carbohydrate

N-acetylglucosamine links to aspargine (7). Therefore, it may be speculated that in any kind of carbohydrate-protein complex, there would exist such a characteristic sequence that is the site recognized by glycosyltransferases. Namely, there seems to be an information directed by the amino acid sequence. This hypothesis can answer the question why only the limited combinations of sugar and amino acid are found in glycoproteins.

Search for a characteristic amino acid sequence in mucoproteins such as chondromucoprotein (8), where xylose links to serine, and in glycoproteins (9) including submaxially mucin and blood group substances where N-acetylgalactosamine is believed to link to threonine or serine would provide an answer for this problem.

REFERENCES

1. Isemura, M., and Schmid, K., in preparation.
2. Morgan, P. H., Jacobs, H. G., Segrest, J. P., and Cunningham, L. W., J. Biol. Chem., 245, 5042 (1970).

3. Schroeder, W. A., (1967), "Methods in Enzymology", Vol. XI, ed. by Hirs, C. H. W., Academic Press, New York, p.351.
4. Spiro, R. G., J. Biol. Chem., 242, 4813 (1967).
5. Edman, P., and Begg, G., Europ. J. Biochem., 1, 80 (1967).
6. Elzinga, M., Lai, C. Y., and Hirs, C. H. W., Arch. Biochem. Biophys., 123, 353 (1968).
7. Spiro, R. G., Ann. Rev. Biochem., (1970), Vol. 39, p.599.
8. Katsura, N., and Davidson, E. A., (1967), 7th Internat. Congr. Biochem., Tokyo, Abstracts Vol. IV, p.701.
9. Eylar, E. H., J. Biol. Chem., 246, 2519 (1971).