1b was prepared from **9b** according to the same procedure as described for the preparation of **1a**; yield of pure **1b** 2AcOH H_2O, 95%; mp 107–112 °C, $[\alpha]_D^{20}$ −56° (H_2O).

Results and discussion. For the comparison of the synthetic peptides with I_B, TLC was performed at first using four different systems. As shown in the table, **1a** revealed identity with I_B. To compare mass spectra[10], the N-terminal arginine residues in **1a** and **1b** were converted to α-N-acetyl-pyrimidyl-ornithines and the C-terminus to methyl esters according to the procedure described in the previous paper[3]. The spectra of the derivatives from **1a** and **1b** (measured with a Hitachi RMU-6L, 70 eV, about 200 °C of sample temperature) showed quite a difference in their fragmentation patterns. The derivative from **1a** revealed ion peaks at m/e 777 (M^+), 585, 441, 344, 197 and 148, which were exactly the same as those observed for the corresponding derivative from I_B. Thus the identity of **1a** with I_B was verified unambiguously. The biological activities of **1a**, **1b** and I_B to induce agglutinability in *a* type cells (H22 strain) were assayed according to the method reported previously[4]. Contrary to our expectation, every compound tested showed the activity in almost the same degree at dosages of 2–20 ng/ml.

Ribadeau Dumas et al. elucidated the whole primary structure of bovine β-casein[11], which contained the sequence of –Arg(202)–Gly–Pro–Phe–Pro–Ile(207)–Ile–Val–OH(209) in the C-terminal portion. It is noteworthy that the part of –Arg(202)–Gly–Pro–Phe–Pro–Ile(207)– is just identical with I_B. Peptides having the structure of H–Arg–Gly–Pro–Pro–Phe–Ile–Val–OH (**10**)[12] and H–Gly–Pro–Phe–Pro–Val–Ile–OH (**11**)[13] were isolated as bitter principles from the enzymatic hydrolyzate of casein. We observed that the synthetic **1a** and **1b** had a strong bitter taste in almost the same threshold value at 0.13–0.25 mM (0.1–0.2 mg/ml). It is interesting that the peptide isolated as a sexual factor in the yeast has a very similar structure[14] to that of bitter principles.

1 Laboratory of Biochemistry, Faculty of Science, Kyushu University, Higashi-ku, Fukuoka 812, Japan.
2 The Institute of Physical and Chemical Research, Wako-shi, Saitama-ken 351, Japan.

3 A. Sakurai, K. Sakata, S. Tamura, K. Aizawa, N. Yanagishima and C. Shimoda, Agr. biol. Chem. in press.
4 A. Sakurai, S. Tamura, N. Yanagishima and C. Shimoda, Proc. Japan Acad. *51*, 291 (1975).
5 Amino acid symbols except Gly denote the L-configuration. The abbreviations recommended by the IUPAC-IUB Commission of Biochemical Nomenclature (J. biol. Chem. *247*, 977 (1972)) have been used throughout.
6 T. R. Vaughan, Jr, and R. L. Osato, J. Am. chem. Soc. *74*, 676 (1952).
7 J. C. Anderson, M. A. Barton, P. M. Hardy, G. W. Kenner and R. C. Sheppard, J. chem. Soc. *1967*, 108.
8 Satisfactory elemental analyses, ratios of amino acid analysis on acid hydrolysates, and chromatographic data were obtained for crystalline compounds. For many compounds, the description of data (yield, mp, $[\alpha]_D$ etc.) was omitted.
9 G. W. Anderson, J. E. Zimmerman and F. M. Callahan, J. Am. chem. Soc. *86*, 1839 (1964).
10 We wish to express our thanks to Dr A. Isogai and Dr K. Aizawa, Department of Agricultural Chemistry, the University of Tokyo, for measurements of mass spectra.
11 B. Ribadeau Dumas, G. Bringnon, F. Grosclaude and J. C. Mercier, Eur. J. Biochem. *25*, 505 (1972).
12 N. Minamiura, Y. Matsuura, F. Fukumoto and T. Yamamoto, Agr. biol. Chem. *36*, 588 (1972).
13 T. Matoba, R. Hayashi and T. Hata, Agr. biol. Chem. *34*, 1235 (1970).
14 We assume that the partial structure of **10** isolated will be H–Arg–Gly–Pro–Phe–Pro–Ile–, and **11** isolated will be H–Gly–Pro–Phe–Pro–Ile–.

Collagen glycopeptides from the sea cucumber Stichopus japonicus

M. Isemura and T. Ikenaka[1]

Department of Biochemistry, Niigata University School of Medicine, Niigata, 951 (Japan), 6 December 1976

Summary. Structures of several hydroxylysine-containing glycopeptides derived from the body wall of the sea cucumber have been determined. The glycosylated hydroxylysines were all the Y positions of Gly-X-Y triplets. They are thought to be distributed throughout the peptide chain, in contrast to vertebrate skin collagens. One glycopeptide demonstrated the existence of the sequence Gly-Gly-Hyp in invertebrates for the first time.

Glycoprotein nature of collagens has been established and many works have been carried out concerning the structure and biosynthesis of the carbohydrate units of the proteins[2–5]. In our previous works on invertebrate collagens[6,7], we observed the differences from vertebrates in the characteristics of the amino acid residues around the glycosylated sites. The present work was initiated to confirm whether such differences represent the reflection of the phylogenetic positions of the animals studied.

In this paper, we determined the structures of collagen glycopeptides from the body wall of the sea cucumber and compared them with those of vertebrate and invertebrate collagens.

Materials and methods. The body walls of Stichopus japonicus were defatted with acetone and minced in a blendor. The pronase P digestion was performed as described previously[7]. The digest was treated with cetylpyridinium chloride[8,9] and fractionated by gel-filtration through

1 Acknowledgments. The authors wish to express their gratitude to Dr R. Hayashi of Kyoto University for a generous gift of carboxypeptidase Y. This study was supported in part by a scientific research grant from the Ministry of Education, Science and Culture of Japan.

2 R. G. Spiro, in: Glycoproteins, vol. 5, p. 964. Ed. A. Gottschalk. BBA Library, Elsevier, Amsterdam 1972.
3 R. D. Marshall, A. Rev. Biochem. *41*, 673 (1972).
4 R. G. Spiro, Adv. Protein Chem. *27*, 350 (1973).
5 P. M. Gallop and M. A. Paz, Physiol. Rev. *55*, 418 (1975).
6 M. Isemura, R. K. Zahn and K. Schmid, Biochem. J. *131*, 509 (1973).
7 M. Isemura, T. Ikenaka and Y. Matsushima, J. Biochem., Tokyo *74*, 11 (1973).
8 M. Isemura and T. Ikenaka, Biochim. biophys. Acta *404*, 11 (1975).
9 M. Isemura, T. Ikenaka, T. Mega and Y. Matsushima, Biochem. biophys. Res. Comm. *57*, 751 (1974).

Sephadex G-50[7]. The main hexose-containing fraction was further fractionated by conventional chromatographic techniques[7].

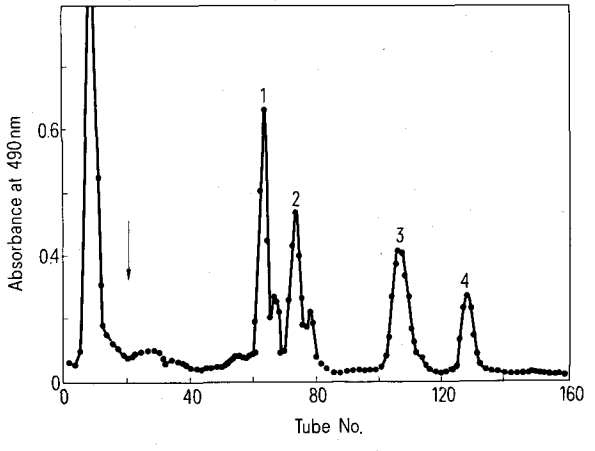
Amino acid analysis and sequence determination were performed as described previously[7,8]. To the serine-containing glycopeptide, the partial acid hydrolysis[10] with 12 N HCl at 23 °C for 24 h was successfully applied. The resultant peptides were purified by high voltage paper electrophoresis[7] at pH 3.6. A glycopeptide was digested with Clostridium collagenase (Seikagaku Kogyo Co, Tokyo, Japan) in 0.1 M pyridine-acetate, pH 6.5 for 20 h. The resultant peptides were fractionated on a column of Bio-Gel P-2 (1.2 × 108 cm) using 0.1 M acetic acid, and further purified by preparative paper electrophoresis at pH 3.6. In some cases, carboxypeptidase B or carboxypeptidase Y[11] was utilized to confirm the carboxyl terminal residues[7]. Glycosylated hydroxylysines were characterized by paper electrophoresis and by the determination of the hexose content[6,7,9].

Results and discussion. As shown in the figure, 4 major glycopeptide fractions were obtained by Dowex 50 chromatography. From peaks 1, 2, 3 and 4,5 homogeneous glycopeptides **I**, **II**, **III A** plus **III B** and **IV** were isolated, respectively.

Structures of glycopeptides derived from the body wall of the sea cucumber

Glycopeptide	Structure
I	Gly-Ala-Hyp-Gly-Pro-Hyl*-Gly-Glu
II	Thr-Gly-Pro-Hyl*-Gly-Pro-Hyl*-Gly-Glu-Pro-Gly-Gln-Thr-Pro
III A	Gly-Gly-Hyp-Gly-Ser-Hyp-Gly-Ser-Hyl*-Gly-Ser-Arg
III B	Gly-Ala-Hyl*-Gly-His
IV	Gly-Gln-Hyl*-Gly-Glu-Arg

*The disaccharide, glucosylgalactose is attached to this residue.



Separation of glycopeptides derived from the body wall of the sea cucumber by Dowex 50X2 chromatography. A column (1.9 × 38 cm) equilibrated with 0.2 M pyridine-acetate, pH 3.1, was eluted with the same buffer. At the position indicated by the arrow, the linear gradient elution was started with pH 3.1 buffer (500 ml) and 2 M pyridine-acetate, pH 5.0 (500 ml). Each fraction contained 5.5 ml and 0.3 ml-portions were analyzed for carbohydrate[7]. The first peak was not studied in the present work.

The direct Edman degradation of glycopeptide **I** and its amino acid composition established the sequence Gly-Ala-Hyp-Gly-Pro-Hyl-Gly-Glx. After 7 cycles of the degradation, glutamic acid was identified on the amino acid analyzer[7,8].

The 6 cycles of the Edman degradation for **II** showed the sequence Thr-Gly-Pro-Hyl-Gly-Pro-. Carboxypeptidase Y released proline (0.5 and 0.86 moles per mole after 1 and 3 h, respectively) and threonine (0.24 and 0.7 moles per mole after corresponding time). From the collagenase digest of **II**, 5 peptides were isolated. These were shown to be Thr-Gly-Pro-Hyl (Co-1), Gly-Pro-Hyl-Gly-Glu-Pro-Gly (Co-2), Gly-(Thr, Glx, Pro), Gln-Thr-Pro and a peptide with the composition corresponding to that of parent peptide **II** minus Co-1. These results suggested the sequence of **II** as shown in the table. The structure of **III A** was first assigned as shown in the table by the direct Edman method. In order to confirm this sequence, the partial acid hydrolysis[10] was utilized. 3 peptide fractions were obtained by paper electrophoresis. The amino acid compositions and subtractive Edman method showed that they were Ser-Arg, Ser-Hyl-Gly and the fraction with a composition, Hyp_2, Ser_1, Gly_4. Quantitative release of arginine from **III A** was accomplished by carboxypeptidase B digestion for 20 h[7]. The structures of **III B** and **IV** were similarly deduced as shown in the table.

Upon the alkaline hydrolysis[7] each parent glycopeptide, Co-1 and Co-2 gave glucosylgalactosylhydroxylysine, indicating that the carbohydrate units linked to hydroxylysines are all disaccharides.

The sequences of glycopeptides supported the view for the existence of glycine triplets in invertebrate collagens[6,7]. The disaccharide-linked hydroxylysines were all at the Y positions of the helical triplet. In 2 cases, this residue was followed by the sequence Gly-X-Arg, which has been found most frequently in collagen glycopeptides from both vertebrate and invertebrate[6,7,9,12,13]. In this sense, glycopeptide **II** was exceptional. The high affinity of glycosyltransferase(s) for this general sequence has been suggested[6,7,9,12]. In the previous studies of invertebrate collagens[6,7], the amino acids abundant in collagen such as proline, glutamic acid and alanine, were found in the glycopeptide linkage region, which is in contrast to vertebrates[12]. The present results are generally in agreement with this observation, but glycopeptide **III B** contained histidine in this region like vertebrate collagen glycopeptides[12].

The isolation of glycopeptide **II**, which is conceivably derived from the carboxyl terminal region, suggests that in invertebrate collagen the glycosylated hydroxylysines are distributed throughout the peptide chain, a feature similar to vertebrate basement membrane collagen[14], but in contrast to vertebrate Type I collagen[13].

The present study also showed the existence of the sequence Gly-Gly-Hyp (glycopeptide **III A**) in invertebrate as well as in vertebrate collagens[15,16].

10 F. Sanger and H. Tuppy, Biochem. J. *49*, 463 (1951).
11 R. Hayashi, S. Moore and W. H. Stein, J. biol. Chem. *248*, 2296 (1973).
12 P. H. Morgan, H. G. Jacobs, J. P. Segrest and L. W. Cunningham, J. biol. Chem. *245*, 5042 (1970).
13 J. H. Aguilar, H. G. Jacobs, W. T. Butler and L. W. Cunningham, J. biol. Chem. *248*, 5106 (1973).
14 N. A. Kefalides, Biochem. biophys. Res. Comm. *47*, 1151 (1972).
15 P. P. Fietzek and J. Rauterberg, FEBS Lett. *49*, 365 (1975).
16 P. P. Fietzek and K. Kühn, Mol. cell Biochem. *8*, 141 (1975).