The amino acid sequences of homologous hydroxyproline-containing myotoxins from the marine snail *Conus geographus* venom

Showbu Sato, Hideshi Nakamura, Yasushi Ohizumi, Jun'ichi Kobayashi and Yoshimasa Hirata+

Mitsubishi-Kasei Institute of Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194 and *+Faculty of Pharmacy, Meijo University, Tempaku, Nagoya 468, Japan*

Received 9 March 1983

Two homologous toxic peptides containing hydroxyproline from the venom of the marine snail *Conus geographus* have been sequenced.

Geographutoxin I:

Arg–Asp–Cys–Cys–Thr–Hyp–Hyp–Lys–Lys–Cys–Lys–Asp–Arg–Gln–Cys–Lys–Hyp–Gln–Arg–Cys–Cys–Ala-NH_2

Geographutoxin II:

Arg–Asp–Cys–Cys–Thr–Hyp–Hyp–Arg–Lys–Cys–Lys–Asp–Arg–Arg–Cys–Lys–Hyp–Met–Lys–Cys–Cys–Ala-NH_2

These peptides inhibit the contractile response of directly stimulated mouse diaphragm.

Conus geographus *Toxic peptide* *Geographutoxin I and II* *Amino acid sequence* *Hydroxyproline*

1. INTRODUCTION

Most of the marine snails belonging to *Conus* genus have some toxic substances [1–9]. Neurotoxic peptides named conotoxin GI, GIA and GII were isolated from the venoms of *Conus geographus*, and their chemical structures elucidated. They are composed of 13–15 amino acid residues including two cystine bridges [4,4a].

We have isolated also from *C. geographus* two toxic peptides different from conotoxin GI, GIA and GII in pharmacological properties and in amino acid compositions, and referred to them as geographutoxin (GTX) I and II [10]. They are composed of 21–22 amino acid residues including 3 hydroxyprolines and have 3 cystine bridges in a molecule. In pharmacological studies, they both showed an inhibitory effect on the contractile response of the mouse diaphragm to direct stimulation. The amino acid composition of a toxic peptide corresponding to GTX I was described in [11]. We present here the primary structure of GTX I and II.

2. MATERIALS AND METHODS

Snails of *C. geographus* were collected around the islands of Okinawa. GTX I and II were isolated from 30–40 pairs of the venom glands by a procedure involving gel-filtration on Sephadex G-50, ion-exchange chromatography on CM-Sephadex C-25 and high-performance liquid chromatogra-

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

phy (HPLC) as in [10]. Peptides were oxidized with performic acid as in [12]. The BrCN cleavage of GTX II was performed in 70% (v/v) formic acid with a BrCN:peptide molar ratio of 20:1, for 20 h at room temperature in the dark. After the amino groups were modified with phenylisothiocyanate, oxidized GTX I was digested in 0.1 M NH_4HCO_3 with trypsin (1-chloro-4-phenyl-3-tosylamido-butan-2-one treated trypsin, Worthington Co.) with enzyme:substrate at 1:30 (w/w), for 4 h at 37°C. The resulting fragments were fractionated by chromatography on DEAE-cellulose column (0.6 × 16 cm). The peptides were eluted with NH_4HCO_3 of a linear gradient (0.02–0.5 M). The eluate was determined by absorbance at 230 nm and by dansyl method. The purity of the peptides was checked by electrophoresis.

Electrophoresis was carried out on a cellulose thin-layer plate (Avicel, Asahi Chemical Industry, 5 or 10 × 20 cm) placed on a plate circulated with coolant at 0°C. The voltage of 1200 V was applied over a cellulose plate 20 cm long, in pyridine acetate buffer (10% pyridine/0.5% acetic acid (pH 6.5) or 0.5% pyridine/5% acetic acid (pH 3.5)). The peptides separated on a cellulose plate were eluted with 0.1 M acetic acid.

Amino acid analysis was performed on a Durrum D-500 amino acid analyzer with a program for collagen analysis. The samples were hydrolyzed in 6 M HCl at 110°C for 24 h in an evacuated tube. Edman dansyl method was applied as in [13]. Dns-amino acids were identified by two-dimensional TLC on a polyamide layer sheet and by reversed phase HPLC. 5-Dimethylamino-naphthalene sulfonyl (Dns)-alaninamide was prepared by allowing L-alaninamide (Sigma Co.) to react with dansylchloride in 0.2 M $NaHCO_3$/acetone and used as a standard without purification. The standards of other Dns-amino acids including *N*-Dns-hydroxyproline were obtained from Sigma Co.

Solid-phase sequencing was performed on an LKB 4030 sequencer with manual conversion in 1 N HCl at 80°C for 10 min. Phenylthiohydantoin (PTH)-amino acids were analyzed by reversed-phase HPLC using two-gradient elution systems, 5% tetrahydrofuran (THF) in acetate-Na^+ (pH 5.3)–acetonitrile/THF (9:1, v/v) and phosphate-triethylamine (pH 3.2)–acetonitrile/THF (9:1, v/v). PTH-homoserine (Hse) was synthesized by the method for PTH-serine or PTH-threonine. The peptide mixture of GTX II treated with BrCN and followed by performic acid oxidation was reacted with aminoethylaminopropyl (AEAP) glass after trifluoroacetic acid treatment to prepare the GTX II derivative attached with glass via the homoserine residue. Oxidizd GTX I was coupled with phenylenediisothiocyanate (DITC)-treated AEAP glass. The two tryptic peptides derived from the oxidized phenylthiocarbamyl GTX I were coupled with AEAP glass by *N*-ethyl-*N*-dimethylaminopropyl carbodiimide (EDC) method.

3. RESULTS

The oxidized GTXs were electrophoresed at pH 6.5. The relative mobility to aspartic acid was −0.02 for oxidized GTX I and −0.14 for oxidized GTX II. The electrophoretic mobility plotted in the diagram of [14] suggests that the net charge of oxidized GTX I and II were 0 and +1, respectively, assuming the M_r of the oxidized GTXs to be 3300.

The oxidized GTXs were analyzed by Edman dansyl method. The sequence of Arg–Asx–X–X–Thr–Hyp–Hyp– (X stands for a non-determined residue) was determined for both oxidized GTXs. The oxidized GTX I was reacted with phenylisothiocyanate, digested with trypsin and subjected to chromatography on DEAE-cellulose

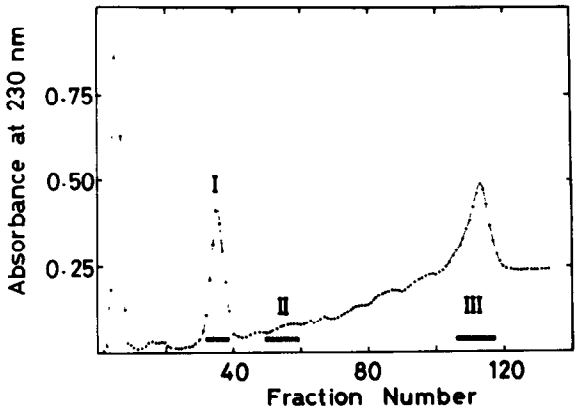

Fig.1. Separation of the tryptic peptides of oxidized phenylthiocarbamyl GTX I by chromatography on DEAE-cellulose column. The eluate was determined by absorbance at 230 nm and Dns-method. Peptide I, II and III were obtained from the fractions indicated with bars.

column (elution profile in fig.1). The eluate was analyzed by absorbance at 230 nm and dansyl method, and peptide I, II and III were obtained from the fractions indicated with the bars in the figure. The electrophoretic mobility and the results of amino acid analysis were as follows: peptide I, $m_{Asp} = -0.01$, cysteic acid (1.00), hydroxyproline (1.21), glutamic acid (2.23), lysine (0.71) and arginine (1.00); peptide II, $m_{Asp} = 0.50$, cysteic acid (1.70) and alanine (1.00); and peptide III, $m_{Asp} = 0.61$, cysteic acid (2.53), hydroxyproline (1.90), aspartic acid (1.96), threonine (0.99), lysine (2.19) and arginine (1.00). Although the value of lysine was rather low, the amino acid composition of the 3 peptides explains that of the whole oxidized GTX I except one arginine residue. Peptide I and III were coupled with AEAP glass by EDC method and sequenced. The results were: peptide I, Gln–Cya–Lys–Hyp–Gln–Arg (Cya stands for cysteic acid); peptide III, X–Cya–Cya–Thr–Hyp–Hyp–Lys–Lys–Cya–Lys–X–Arg. Peptide II was determined to be Cya–Cya–Ala-NH_2 by Edman dansyl method. Dns-alaninamide was detected without hydrolysis after 2 cycles of Edman degradation. Peptide II is the C-terminal peptide. The solid phase sequencing of the whole oxidized GTX I showed that the 'X' residues in the sequence of the peptide III were both aspartic acid and that the peptide I was located after the peptide III. The sequence of GTX I is shown in fig.2.

GTX II was treated with BrCN and oxidized with performic acid. The peptide with homoserine at the C-terminus was selectively coupled with AEAP glass via the homoserine residue. The solid-phase Edman degradation of the peptide gives a sequence of

Arg–Asp–Cya–Cya–Thr–Hyp–Hyp–Arg–Lys–Cya–Lys–Asp–Arg–Arg–Cya–Lys–Hyp–Hse

The peptide not coupled with the AEAP glass was purified from the reaction mixture by electrophoresis at pH 6.5. The peptide was neutral and the amino acid analysis showed that it had two cysteic acids, one alanine and one lysine. Edman dansyl analysis gave the sequence Lys–Cya–Cya–Ala-NH_2. The C-terminal alaninamide was determined by dansylation of non-hydrolyzed sample after 3 cycles of the Edman degradation. The whole sequence of GTX II is shown in fig.2. Three arginine residues were reported present in GTX II [10]. The underestimated value might be due to a small amount of the sample analyzed or due to contaminants in the sample. GTX I and II have 7 and 8 basic amino acids, respectively, and they both have amide group at the carboxyl terminus and two aspartic acid residues. On an increase of 6 minus charges on performic oxidation, the net charge of oxidized GTXs (zero for GTX I and +1 for GTX II) is consistent with the results of electrophoresis at pH 6.5.

```
                                  5                  10
Geographutoxin I   Arg-Asp-Cys-Cys-Thr-Hyp-Hyp-Lys-Lys-Cys-Lys-Asp-
Geographutoxin II  Arg-Asp-Cys-Cys-Thr-Hyp-Hyp-Arg-Lys-Cys-Lys-Asp-

                              15                  20
                   Arg-Gln-Cys-Lys-Hyp-Gln-Arg-Cys-Cys-Ala-NH2
                   Arg-Arg-Cys-Lys-Hyp-Met-Lys-Cys-Cys-Ala-NH2
```

Fig.2. The amino acid sequences of geographutoxin I and II. The amino acid residues which are not common in both toxins are underlined.

4. DISCUSSION

We have sequenced two muscle toxins, geographutoxin I and II, which have three hydroxyprolyl residues. They both consist of 22 amino acids and their primary structures are very similar to each other. There are 4 amino acid substitutions between them, at the 8th, 14th, 18th and 19th residue. Except for the substitution at the 18th residue, Gln ↔ Met, the substitution of the amino acids was explained by single base substitution in the triplet codons. Since both have two aspartic acid residues, they have 5 or 6 positive net charges in a molecule. Some of basic amino acids are expected to be involved in their functions. The guanidino-group at the 14th in GTX II seems to enhance its biological activity since the lethal dose of GTX II against mouse is less than that of GTX I (LD_{50} of GTX I and GTX II were 340 μg/kg body wt and 110 μg/kg body wt, respectively). GTX II is superior to GTX I also in the effect on the contractile response of directly stimulated diaphragm. The pharmacological actions of GTX I and II are now studied in detail.

Three hydroxyprolines in GTXs are all 4-*trans*-L-hydroxyproline, judged from the elution position of the free amino acid in the amino acid analysis and from the R_f value on two-dimensional polyamide TLC and the elution time on the reverse phase HPLC of the Dns-derivative. 4-*trans*-

Hydroxyproline is best known as a constituent of collagen where the sequence –Gly–X–Hyp–Gly– (X is commonly Pro) repeats frequently. In addition, it is found in the sequence Arg–Pro–Hyp–Gly– of bradykinin-like peptide which is reminiscent of the sequence in collagen [15]. In higher plants, the sequence –Ser–$(Hyp)_4$–(Ser, Lys or Val)– was found in cell-wall proteins where the hydroxyl groups of hydroxyproline residues are glycosylated [16]. Protocollagen prolylhydroxylases hydroxylate a prolyl residue in the sequence –Gly–X–Pro–Gly–. Plant prolylhydroxylases hydroxylate almost all prolines in the proline cluster –Ser–$(Pro)_4$–, recognizing a helix of polyproline type II [17]. Hydroxyprolines in GTXs are also expected to be formed by hydroxylation of proline residues in peptide linkage as those in collagen and cell-wall proteins. However, it is difficult to find out similarities among the sequences around the 3 hydroxyproline residues in GTXs. There might be a novel and specific proline hydroxylase in the venom glands which acts on 'Pro'-geographutoxins.

Other noted features in the primary structures are a basic amino acid cluster in the middle of the sequence, and Cys–Cys sequence in both terminal regions. The primary structures of homologous neurotoxins from *C. geographus*, conotoxin GI, GIA and GII, which consist of 13–15 amino acids including 4 half-cystine residues have been elucidated in [4]. Some similarities between geographutoxins and conotoxins are observed in the distribution of half-cystine and in the blocked carboxyl-terminus.

In [11], a muscle toxin from *C. geographus* (conotoxin GIII) was purified. Its amino acid composition suggests that conotoxin GIII seems to be identical with GTX I. But they did not refer to a peptide corresponding to GTX II at all. GTX II is always found in our venom preparation and richer than GTX I at times.

ACKNOWLEDGEMENTS

We would like to thank Mr Z. Nagahama of Okinawa island for supplying specimens of *Conus geographus* and Mrs K. Tomizawa of this institute for amino acid analysis. We are also grateful to Dr T. Uchida for encouraging us during this work.

REFERENCES

[1] Endean, R., Parish, G. and Gyr, P. (1974) Toxicon 12, 131–138.
[2] Spence, I., Gillessen, D., Gregson, R.P. and Quinm, R.J. (1977) Life Sci. 21, 1759–1770.
[3] Cruz, L.J., Gray, W.R. and Olivera, B.M. (1978) Arch. Biochem. Biophys. 190, 539–548.
[4] Gray, W.R., Luque, A., Olivera, B.M., Barrett, J. and Cruz, L.J. (1981) J. Biol. Chem. 256, 4734–4740.
[4a] Nishiuchi, Y. and Sakakibara, S. (1982) FEBS Lett. 148, 260–262.
[5] Kobayashi, J., Ohizumi, Y., Nakamura, H. and Hirata, Y. (1981) Toxicon 19, 757–762.
[6] Kobayashi, J., Nakamura, H., Hirata, Y. and Ohizumi, Y. (1982) Toxicon 20, 823–830.
[7] Kobayashi, J., Nakamura, H., Hirata, Y. and Ohizumi, Y. (1982) Life Sci. 31, 1085–1091.
[8] Kobayashi, J., Nakamura, H., Hirata, Y. and Ohizumi, Y. (1982) Biochem. Biophys. Res. Commun. 105, 1389–1395.
[9] McIntosh, M., Cruz, L.J., Hunkapiller, M.W., Gray, W.R. and Olivera, B.M. (1982) Arch. Biochem. Biophys. 218, 329–334.
[10] Nakamura, H., Kobayashi, J., Ohizumi, Y. and Hirata, Y. (1983) Experientia (accepted).
[11] Stone, B.L. and Gray, W.R. (1982) Arch. Biochem. Biophys. 216, 765–767.
[12] Hirs, C.H.W. (1967) Methods Enzymol. 11, 197–199.
[13] Hartley, B.S. (1970) Biochem. J. 119, 805–822.
[14] Offord, R.E. (1966) Nature 211, 591–593.
[15] Nakajima, T., Yasuhara, T., Falconere-Erspamer, G. and Visser, J. (1979) Experientia 35, 1133.
[16] Brown, R.G. and Kimmins, W.C. (1977) Int. Rev. Biochem. Plant Biochem. II 13, 183–209.
[17] Tanaka, M., Sato, K. and Uchida, T. (1981) J. Biol. Chem. 256, 11397–11400.