Cloning and sequence analysis of a cDNA encoding a precursor for rat C-type natriuretic peptide (CNP)

Masayasu Kojima[1], Naoto Minamino[2], Kenji Kangawa[1] and Hisayuki Matsuo[2]

[1]*Department of Biochemistry, Miyazaki Medical College, Kihara, Kiyotake, Miyazaki 889-16, Japan* and [2]*National Cardiovascular Center Research Institute, Fujishirodai, Suita, Osaka 565, Japan*

Received 26 September 1990

Recent identification of a C-type natriuretic peptide (CNP) in porcine brain strongly suggested that a third member of the natriuretic peptide family still remains to be identified in other species of mammals. A cDNA encoding a precursor for rat CNP was cloned from a rat brain cDNA library and sequenced. The precursor was a 126-residue peptide, carrying a 23-residue signal sequence at the N-terminus and the known porcine CNP-53 sequence at the C-terminus. By RNA blot analysis, rat CNP mRNA was found to be expressed exclusively in the brain, implying that CNP may function in the central nervous system as a neuropeptide.

C-type natriuretic peptide; cDNA cloning; Precursor structure; A-type natriuretic peptide; B-type natriuretic peptide; Neuropeptide

1. INTRODUCTION

Identification of atrial natriuretic peptide (A-type natriuretic peptide: ANP) and brain natriuretic peptide (B-type natriuretic peptide: BNP) disclosed the possibility that a natriuretic peptide (NP) family comprising similar hormones participates in regulating the homeostatic balance of the body fluid volume and blood pressure [1-3]. ANP and BNP, though they are derived from distinct genes, share a highly homologous 17-residue ring structure, which is known to be essential for exerting natriuretic/diuretic and hypotensive activities. We have recently identified in porcine brain a third member of the NP family and designated it C-type natriuretic peptide (CNP) [4,5]. Porcine CNP consisting of 22 amino acid residues (CNP-22) and its N-terminally elongated form with 53 residues (CNP-53) show remarkable sequence homology to ANP and BNP within the 17-residue ring portion formed by a pair of cysteine residues. However, porcine CNP terminates at the second cysteine residue, which participates in the ring formation, and completely lacks the further C-terminal extension, which occurs in both ANP and BNP. Moreover, porcine CNP-22 stimulates guanylate cyclase activity in cultured vascular smooth muscle cells more potently than ANP and BNP, implying that CNP may function in a manner distinct from ANP and BNP [6]. In order to characterize the specified physiological role of CNP and to define its functional difference from ANP and BNP, it is essential to know the structure of rat CNP, since extensive studies on ANP and BNP have been carried out mainly with peptides of rat and human origin. Here we report cloning and sequence analysis of a cDNA encoding a precursor for rat CNP. In addition, regional difference in the expression of the CNP gene was also examined by RNA blot analysis.

Correspondence address: H. Matsuo, National Cardiovascular Center Research Institute, Fujishirodai, Suita, Osaka 565, Japan

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number no. D90219

2. MATERIALS AND METHODS

2.1 *cDNA library construction*

Total RNA was extracted from rat brain by the guanidine thiocyanate method. Poly(A)$^+$ RNA was isolated on an oligo(dT)-cellulose column (Pharmacia). Double-stranded cDNA was synthesized from 4 μg of rat brain poly(A)$^+$ RNA by the method of Gubler and Hoffman [7]. cDNA was ligated to *Eco*RI adaptors and size-fractionated by 1% agarose gel electrophoresis. After electroelution, 300-1500 bp-fraction was ligated to phage λgt10 arms (Bethesda Research Laboratory) and packaged in vitro by using Gigapack Gold (Stratagene).

2.2 *cDNA library screening and sequence analysis*

High- and low-stringencies in hybridization were controlled by the two different concentrations of formamide (50% and 20%). Hybridization was performed at 37°C in a solution of 5 × SSPE, 5 × Denhardt's solution, 0.1% SDS, 100 μg/ml denatured salmon sperm DNA, containing 50% or 20% formamide. Rat CNP genomic clones were obtained from Charon4A rat genomic library (Clonetech). Recombinants from the library were screened under low-stringency conditions, by using a 150-bp genomic DNA fragment encoding porcine CNP-53 as a probe, which was kindly donated by Y. Tawaragi (Suntory Institute for Biomedical Research). As reported [8], the probe was prepared from porcine liver DNA by the PCR-amplified method, using sense and antisense primers, corresponding to the N-terminal and C-terminal sequences of porcine CNP-53. One of the clones thus isolated was partially sequenced by the dideoxy chain termination method. Then, the *Sma*I/*Mva*I fragment of the clone, corresponding to nucleotides 245-394 in Fig. 1, was used as a

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

```
                                                    -116 CC GCA CAG CAG TAG GAC CCG  -97

-96  TGC TCG CTT GGC AAT CCT GCT CTG CAA CCG CTT GTC GGA CTG CTC ACC  -49

-48  GGC CGT CCC GGC TGC AGA CTG TCT GCA CCC CTC GGT CCC ATC GGC ACC   -1
     -23          -20                                      -10
     Met His Leu Ser Gln Leu Ile Ala Cys Ala Leu Leu Leu Ala Leu Leu
  1  ATG CAC CTC TCC CAG CTG ATC GCC TGT GCC CTG CTG CTC GCG CTA CTC   48
                                  -1  1
     Ser Leu Arg Pro Ser Glu Ala Lys Pro Gly Thr Pro Pro Lys Val Pro
 49  TCA CTC CGG CCC TCC GAA GCC AAG CCC GGG ACA CCA CCG AAG GTC CCG   96
     10                                          20
     Arg Thr Pro Pro Gly Glu Glu Leu Ala Glu Pro Gln Ala Ala Gly Gly
 97  AGA ACC CCG CCA GGG GAG GAG CTG GCA GAG CCC CAG GCA GCT GGT GGC  144
                 30                                          40
     Asn Gln Lys Lys Gly Asp Lys Thr Pro Gly Gly Gly Gly Ala Asn Leu
145  AAT CAG AAA AAG GGT GAC AAG ACT CCA GGC GGC GGG GGA GCC AAT CTC  192
                                    50
     Lys Gly Asp Arg Ser Arg Leu Leu Arg Asp Leu Arg Val Asp Thr Lys
193  AAG GGA GAC CGA TCG CGA CTG CTT CGG GAC CTG CGT GTG GAC ACC AAG  240
         60                                          70
     Ser Arg Ala Ala Trp Ala Arg Leu Leu His Glu His Pro Asn Ala Arg
241  TCC CGG GCG GCG TGG GCT CGC CTT CTG CAC GAG CAC CCC AAC GCG CGC  288
                       80
     Lys Tyr Lys Gly Gly Asn Lys Lys Gly Leu Ser Lys Gly Cys Phe Gly
289  AAA TAC AAA GGC GGC AAC AAG AAG GGC TTG TCC AAA GGC TGC TTT GGC  336
     90                                  100         103
     Leu Lys Leu Asp Arg Ile Gly Ser Met Ser Gly Leu Gly Cys ***
337  CTC AAG CTG GAC CGG ATC GGC TCC ATG AGC GGT CTG GGA TGT TAG TGC  384

385  AGC GAC CCC TGG CGG CGG ATT GGG AAC TGC ACT GTG CAC TGA GGT CAT  432

433  CCT TGG TCA TCA GCC TCC AGC ATC TGG AAA CAC CTC CAA CGC AAT GTG  480

481  GCT TTT ACA TTT CTT TTT ATT TTT TCC TCC TGG TAC TGG CAA TAC ACA  528

529  ACA CCA GCT GTT TTA TTA TTA TTT GGG GAG GGG AGG GGA TGA TTT TAT  576

577  TGT TTG GGG TTT TTT TTT TGA AAA TGA AAA ATA AAA AAT TAT ATA TTA  624

625  TAT ATA TAT TAT ATA CAT GAG ACA CAC ACT CCC ACA CCG ACT TGA TGA  672

673  CAA GGG ACG GTT TTT AAA GTG ACT GAC AAA ACC AGC TAG CTG TAA AAA  720

721  CAT TGC TGT TTG TAA ATT CAC ATC ATG CAT AAA TGT ATT TAT GTT GTA  768

769  AAG CTA TTT ATA TTG TTT ATA AAG AGA TAT TTA TAA AAA TTT TAT TTA  816

817  TGT AAC TAA ATG AAA GAA GCC AAC CAT TGT AAT GTT TTT GTC CTA ACT  864

865  AGT TGA AAA AAA ATG TTA AAA AAA AAA AAG CCA TTC CAT G(Poly A)    904
```

Fig. 1. Nucleotide sequence of the cDNA insert in λrCNP21 with predicted amino acid residues. Nucleotide residues are numbered beginning with the first residue of ATG triplet encoding a putative initiating methionine, and those on the 5′-side of nucleotide 1 are indicated by negative numbers. The termination codon is marked with three consecutive asterisks. The AATAAA sequence is underlined. Dashed lines under the 3′-untranslated region indicate ATTTA motifs.

probe to screen the rat brain cDNA library. Approximately 6×10^6 recombinant phages were transferred to nitrocellulose membranes and screened under high-stringency conditions. A positive clone λrCNP21 harboring the longest cDNA insert was subcloned into M13 vector and sequenced. Both strands of the cDNA clone were sequenced.

2.3 *RNA blot analysis*

Poly(A)$^+$ RNA (20 μg) was denatured using glyoxal and dimethylsulfoxide, and was fractionated on a 1.4% agarose gel. After electrophoresis, RNA was transferred to a nylon membrane (Zeta Probe, Bio-Rad) and fixed by utraviolet irradiation (Stratalinker, Stratagene). The membrane was prehybridized and hybridized at 37°C in 50% formamide, 6 × SSPE, 5 × Denhardt's solution, 0.5% SDS, and 100 μg/ml denatured salmon sperm DNA. A cDNA insert of λrCNP21 was labeled by the random-primed method and used for the hybridization. The blot was washed at 37°C once in 3 × SSC, 0.1% SDS, once in 0.5 × SSC, 0.1% SDS, and finally twice in 0.1 × SSC, 0.1% SDS. After autoradiography, the radiolabeled probe was removed by boiling the membrane in 0.1 × SSC, 0.5% SDS for 20 min. Then, the membrane was used for the hybridization by the ANP or BNP probe.

3. RESULTS AND DISCUSSION

Recombinants obtained from a rat genomic library were first screened under low-stringency conditions by using a 150-bp porcine CNP probe corresponding to porcine CNP-53 [5,8]. Five positive clones thus obtained exhibited the identical restriction inserts of about 19 kbp long. One of the clones obtained above was partially sequenced to verify that the clone encoded the rat CNP gene. Then, the *SmaI*/*MvaI* fragment of the clone corresponding to the CNP-53 region was used as a pro-

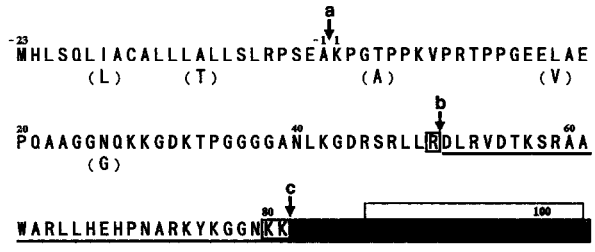

Fig. 2. Amino acid sequence of rat prepro-CNP. Amino acid replacements observed between rat and porcine CNP precursors are shown in the parentheses. One-letter amino acid notation is used. Arrows indicate the putative processing sites for the signal peptidase (a), and for generating CNP-53 (b) and CNP-22 (c). Processing signals for CNP-53 and CNP-22 are boxed. CNP-53 portion is underlined; CNP-22 portion is shaded.

be for screening the rat brain cDNA library under high-stringency conditions.

Among 22 positive clones obtained from 6×10^6 recombinant phages, a clone, λrCNP21, harboring the longest cDNA insert, was sequenced. Fig. 1 shows the complete nucleotide sequence of the cDNA, which is 1020 bp long (excluding poly(A) tail). A putative initiation codon ATG is located at nucleotides 1–3, preceded by the consensus sequence for the initiation, while a termination codon TAG is found 126 codons later at nucleotides 379–381. A typical polyadenylation signal, AATAAA, is found only at nucleotides 606–611, but no poly(A) tail follows the signal. A GTTAAA sequence, the closest match to the signal, which is located 26 nucleotides upstream of the polyadenylation site, may serve as a signal for the polyadenylation.

The amino acid sequence encoded in the open reading frame is deduced as shown in Figs 1 and 2. The amino acid sequence flanked by Cys^{87} and Cys^{103} (nucleotides 328–378) corresponds to the 17-residue ring structure characteristic of the NP family. In the cDNA sequence, a TGT codon for the C-terminal Cys^{103} is directly followed by a termination codon. Thus, the open reading frame encodes a putative precursor for rat CNP (rat prepro-CNP), which consists of 126 amino acid residues and carries a bioactive unit at its C-terminus.

The first 23-residue peptide starting from the initial methionine is thought to be a signal peptide, based on its characteristic hydrophobic features. Consequently, it is most likely that the first processing of the precursor takes place in between Ala^{-1} and Lys^{1} to generate a 103-residue pro-CNP (Fig. 2). In the rat prepro-CNP, a typical processing signal, Lys^{80}–Lys^{81}, is followed by a 22-residue C-terminal peptide, which is identical to porcine CNP-22, one of the endogenous CNPs in porcine brain [4]. This fact strongly suggests that the C-terminal 22-residue peptide (rat CNP-22) can also be regarded as an endogenous CNP in the rat. Moreover, the C-

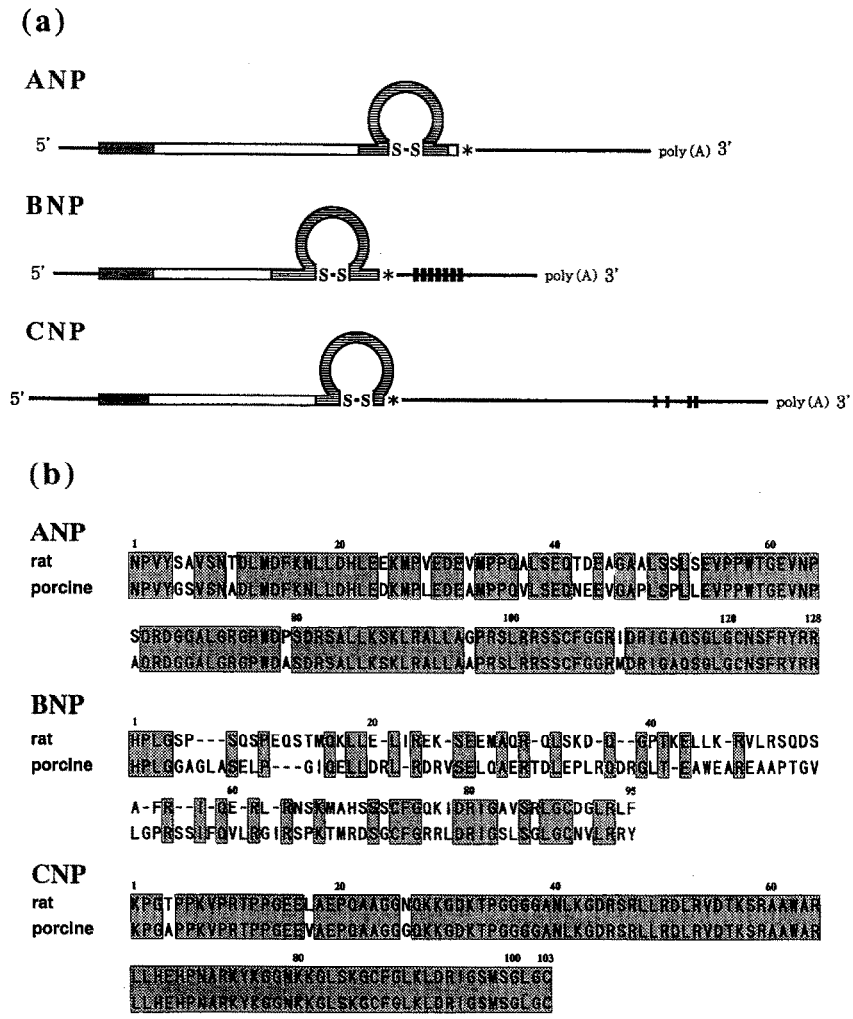

Fig. 3. (a) Schematic diagram of cDNAs for rat natriuretic peptide family. ▨ signal peptide; ▭ bioactive unit; S–S disulfide bond; * stop codon; ■ ATTTA motif. (b) Comparison in amino acid sequence of rat and porcine natriuretic peptide families. Sets of identical residues are shaded. References of the sequences; rat ANP [9], porcine ANP [13], rat BNP [10], porcine BNP [14], porcine CNP [8].

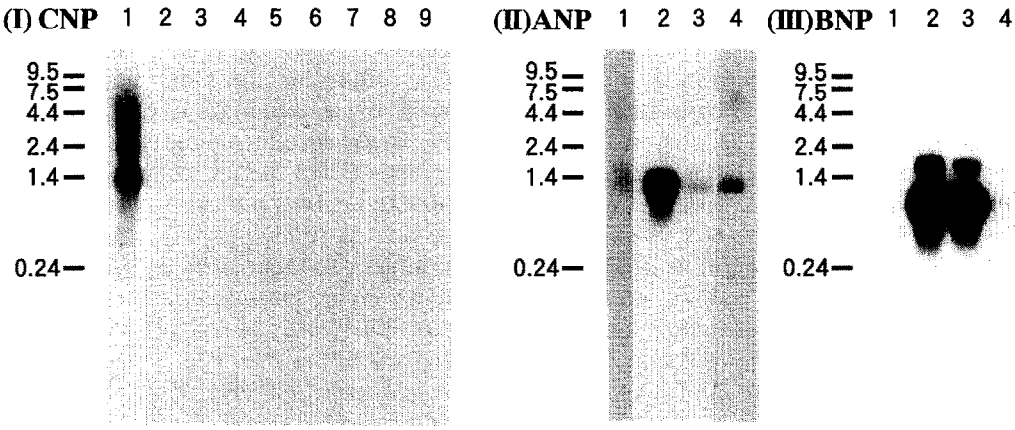

Fig. 4. RNA blot analysis of rat natriuretic peptide transcripts in rat tissues. Each lane contained 20 μg of poly(A)$^+$ RNA. Numbers on the left indicate kilobase as determined from RNA size markers. The probes used for hybridization are (I) CNP probe, (II) ANP probe and (III) BNP probe. Lanes: (1) brain; (2) atrium; (3) ventricle; (4) lung; (5) liver; (6) kidney; (7) stomach; (8) small intestine; (9) large intestine. Autography was at -80°C for 5 days except for (II)-2 and -3 (-80°C for 2 h).

terminal 53-residue sequence (rat CNP-53) is preceded by a single Arg^{50}, as in the case of the porcine CNP precursor (Fig. 2). Thus, it its likely that the processing may take place after Arg^{50} to give another endogenous form (CNP-53).

As reported in our previous papers, porcine CNP-22 as well as CNP-53, unlike ANP and BNP, have the unique structural feature that the C-terminal extension from the ring structure is completely absent [4,5]. The present analysis shows that the codon for the C-terminal Cys^{103} is directly followed by a termination codon. This fact clearly indicates that rat CNP also lacks the C-terminal extension and that this form is generated without any additional post-translational modification.

We have recently accomplished the structural analysis of the porcine CNP gene and deduced the amino acid sequence of a porcine CNP precursor (porcine prepro-CNP) [8]. Surprisingly, rat prepro-CNP is identical to porcine prepro-CNP, except that only 5 amino acid residues are replaced (Fig. 2).

As schematically indicated in Fig. 3a, the molecular construction of prepro-CNP is very similar to those of ANP and BNP. All members of the rat NP family carry the signal sequences at their N-terminal regions and the bioactive units at their C-terminal regions, although they are derived from distinct genes [9,10]. However, the 3′-untranslated region of the rat CNP cDNA is about two times longer than those of rat ANP and BNP cDNAs. Furthermore, the rat CNP cDNA contains in the 3′-untranslated region four copies of an ATTTA motif (nucleotides 757–761, 774–778, 797–801 and 812–816), which is known to destabilize mRNA in the cell [11]. Since the motif is very much clustered in BNP, but is not found at all in ANP, the expression of the three genes for the natriuretic peptide family may be regulated through the different mechanisms [10]. It is well known that nucleotide and amino acid sequences of ANP precursors are highly conserved among various mammals [9,12] (Fig. 3b). In contrast with this, sequence homology among BNP precursors is remarkably low and homologous sequences are limited in the signal peptide and the C-terminal bioactive unit [10]. However, homology between rat and porcine CNP precursors is remarkably high (97.1% in the amino acid sequence of the pro-form), much higher than that between ANP precursors (86.7%). Thus, mammalian CNP is thought to be the most conserved NP in the family. In this respect, it should be mentioned that the C-terminal 22-residue peptide of a human CNP precursor has very recently been established to be identical to those of rat and porcine CNP-22 (to be published).

Fig. 4 shows the RNA blot analyses of the rat NP family. When a cDNA insert of λrCNP21 was used as a CNP probe, mRNA of approximately 1200 bp, which is longer than those of rat ANP (950 bp) and rat BNP (850 bp), was detected only in brain, and not in other examined tissues including cardiac atrium and ventricle. Furthermore, no positive plaque was obtained, when a rat atrial cDNA library was screened with the CNP probe. Thus, rat CNP is concluded to be expressed exclusively in brain but not in atrium and ventricle, while rat ANP and BNP are mainly localized in heart. ANP and BNP, although their expression is regulated in their own manner, are thought to be cardiac hormones mainly secreted from heart, while CNP is expressed only in brain and presumably functions as a neuropeptide in the central nervous system. Thus, ANP, BNP and CNP are most likely to function in concert with each other for maintaining the homeostatic balance of the body fluid volume and blood pressure.

The present identification of CNP in the rat gives a clue to solve the question as to whether or not a receptor specific for CNP is present. In this context, it should be

noted that natriuretic/diuretic and hypotensive activities of CNP are about 100 times less potent than those induced by ANP and BNP, while the rectum-relaxant effect of CNP is 3–4 times more potent than ANP [4]. Furthermore, CNP potently increases cGMP levels in cultured vascular smooth muscle cells, and the extent of the maximum elevation induced by CNP is 3.2 times higher than that by ANP [6]. Such a pharmacological dissociation of CNP from ANP and BNP, combined with regional differences in their expression, suggests a possible existence of a receptor specific for CNP.

Acknowledgements: This work was supported in part by research grants from the Ministry of Education, Science and Culture, and from the Ministry of Health and Welfare of Japan.

REFERENCES

[1] de Bold, A.J. (1985) Science 230, 767–770.
[2] Kangawa, K. and Matsuo, H. (1984) Biochem. Biophys. Res. Commun. 118, 131–139.
[3] Sudoh, T., Kangawa, K., Minamino, N. and Matsuo, H. (1988) Nature 332, 78–81.
[4] Sudoh, T., Minamino, N., Kangawa, K. and Matsuo, H. (1990) Biochem. Biophys. Res. Commun. 168, 863–870.
[5] Minamino, N., Kangawa, K. and Matsuo, H. (1990) Biochem. Biophys. Res. Commun. 170, 973–979.
[6] Furuya, M., Takehisa, M., Minamitake, Y., Kitajima, Y., Hayashi, Y., Ohnuma, N., Ishihara, T., Minamino, N., Kangawa, K. and Matsuo, H. (1990) Biochem. Biophys. Res. Commun. 170, 201–208.
[7] Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263–269.
[8] Tawaragi, Y., Fuchimura, K., Nakazato, H., Tanaka, S., Minamino, N., Kangawa, K. and Matsuo, H. (1990) Biochem. Biophys. Res. Commun. (in press).
[9] Kangawa, K., Tawaragi, Y., Oikawa, S., Mizuno, A., Sakuragawa, Y., Nakazato, H., Fukuda, A., Minamino, N. and Matsuo, H. (1984) Nature 312, 152–155.
[10] Kojima, M., Minamino, N., Kangawa, K. and Matsuo, H. (1989) Biochem. Biophys. Res. Commun. 159, 1420–1426.
[11] Shaw, G. and Kamen, R. (1986) Cell 46, 659–667.
[12] Oikawa, S., Imai, M., Inuzuka, Y., Tawaragi, Y., Nakazato, H. and Matsuo, H. (1985) Biochem. Biophys. Res. Commun. 132, 892–899.
[13] Porter, J.G., Arfsten, A., Palisi, T., Scarborough, R.M., Lewicki, J.A. and Seilhamer, J.J. (1989) J. Biol. Chem. 264, 6689–6692.
[14] Maekawa, K., Sudoh, T., Furusawa, M., Minamino, N., Kangawa, K., Ohkubo, H., Nakanishi, S. and Matsuo, H. (1988) Biochem. Biophys. Res. Commun. 157, 410–416.