Characterization of a cDNA encoding a precursor of *Carassius* RFamide, structurally related to a mammalian prolactin-releasing peptide

Honoo Satake[a], Hiroyuki Minakata[a], Xiaoyan Wang[b], Masaaki Fujimoto[b,*]

[a]*Suntory Institute for Bioorganic Research, Wakayamadai 1-1-1, Shimamoto-cho, Mishimagun, Osaka 618-8503, Japan*
[b]*Department of Biological Science, Faculty of Life and Environmental Science, Shimane University, Nishikawatsu-cho 1060, Matsue 690-8504, Japan*

Received 25 January 1999

Abstract We have characterized the cDNA encoding *Carassius* RFamide (C-RFa), which is structurally related to mammalian prolactin-releasing peptides (PrRPs), from the brain of the crucian carp. The deduced C-RFa precursor has been shown to comprise 117 amino acids, encoding a single C-RFa sequence. A comparative study of amino acid sequences has revealed that several sequences conserved in preproPrRPs are also found in the C-RFa precursor. Furthermore, the abundant presence of the C-RFa mRNA in the telencephalon, optic tectum, medulla oblongata, and proximal half eye ball was demonstrated by Southern blot analysis of RT-PCR products.

Key words: RFamide; Precursor; Expression; *Carassius auratus langsdorfi*

1. Introduction

Recently, many bioactive peptides containing the consensus sequence -Arg-Phe-NH_2 (RFamides) at their C-termini have been characterized from various animals. The first example of such peptides is FMRF-amide isolated from the ganglia of the bivalve mollusc, *Macrocallista nimbosa*, as a molluscan neurotransmitter or neuromodulator [1]. Several cardioactive RFamides have also been identified from other molluscs Table 1. *Achatina* cardioexcitatory peptide (ACEP-1) from the atria of the giant african snail *Achatina fulica* [2], *Lymnaea* ACEP-1-related peptide (LyCEP) from the CNS [3], and *Helix* cardioinhibitory peptide (HCIP) from the brain of the snail *Helix aspersa* [4]. Localizations of FMRF-amide, ACEP-1, LyCEP, HCIP and/or their related peptides in the central nervous system have been demonstrated by immunocytochemical analyses [3–6]. In addition, production of a structurally related peptide in the *Aplysia* neuron L5-67 has been examined [7,8]. These findings suggest that as neuropeptides, RFamides are involved in the regulation or modulation of multiple physiological events including cardiac activity in molluscs.

RFamides and/or their related peptides have been characterized from vertebrates as well as invertebrates, for instance, LPLRFamide from the chicken brain [9], FLFFQPQRFamide (F8F) and AGEGLSSPFWSLAAPQRFamide (A18F) from the bovine brain [10]. Although purification of RFamides from fish species has not been reported, the presence of RFamides in fish tissues was indicated by the detection of RFamide-like immunoreactivities in the nervus terminalis and the retinal amacrine cells of the goldfish [11–13].

A novel 20 amino acid peptide, *Carassius* RFamide (C-RFa), comprising the amino acid sequence SPEIDPFWYVGRGVRPIGRF-NH_2, similar to ACEP-1 and HCIP (Table 1), has been isolated from the brain of a Japanese crucian carp *Carassius auratus langsdorfi* and shown to stimulate contraction of several freshwater fish visceral muscle tissues [14]. Furthermore, the primary sequence of C-RFa is quite homologous to those of mammalian prolactin-releasing peptides (PrRPs) ([15], Table 1), implying that C-RFa and PrRPs may share physiological functions in fish and mammals, respectively, or that C-RFa and PrRPs may constitute a novel peptide family. However, the physiological significance of C-RFa as an endogenous ligand remains to be elucidated.

Since investigation of the C-RFa transcript localizations and an amino acid sequence comparison analysis of the C-RFa precursor polypeptide with preproPrRPs were expected to provide important clues to clarify the physiological roles of C-RFa, we attempted to characterize a cDNA encoding a C-RFa precursor polypeptide. In the present paper, we show the C-RFa cDNA sequence, the amino acid sequence similarity of the C-RFa precursor with a preproPrRP, and the distribution of C-RFa transcript expression. To the best of our knowledge, this is the first identification of a fish RFamide cDNA sequence and also of the structurally PrRP-related peptide from non-mammalian vertebrates.

2. Materials and methods

2.1. Total RNA preparation and first strand cDNA synthesis

Carp were collected in Matsue city and used for experiments within 7 days of collection. The excised tissues were immediately frozen in liquid nitrogen. Total RNA was extracted from each 500 mg *Carassius* frozen tissue by the TRIzol reagent procedure (Gibco, USA).

2.2. Oligonucleotides

All PCR and sequencing primers were purchased from Sawady Technology (Japan). The oligo-dT anchor primer and PCR anchor primer were supplied in the 5′/3′-RACE kit (Boehringer Mannheim, Germany).

2.3. Polymerase chain reaction (PCR)

2.3.1. Amplification of the partial C-RFa cDNA fragment. All PCRs were performed in a 50 μl mixture containing 1.5 mM $MgCl_2$, 0.2 mM dNTP, 12.5 μM each primer, 2 μl cDNA solution, and 0.5 unit of rTAQ polymerase (Toyobo, Japan) or Taq^{EX} polymerase (Takara Shuzo, Japan) on a thermal cycler (Perkin Elmer, model GeneAmp PCR System 2400, UK). First strand cDNA was synthesized with the oligo-dT anchor primer and AMV reverse transcriptase sup-

*Corresponding author. Fax: (81) (852) 32 6432.
E-mail: fujimoto@life.shimane-u.ac.jp

Abbreviations: ACEP-1, *Achatina* cardioexcitatory peptide; C-RFa, *Carassius* RFamide; DIG, digoxigenin; HCIP, *Helix* cardioinhibitory peptide; LyCEP, *Lymnaea* cardioexcitatory peptide; PrRP, prolactin-releasing peptide; RACE, rapid amplification of cDNA end; RT-PCR, reverse transcription-polymerase chain reaction

PII: S0014-5793(99)00215-X

plied in the 5′/3′-RACE kit or Superscript II (Gibco, USA) and amplified using the anchor primer and the first degenerate primers 5′-CCITT(T/C)TGGTA(A/T)GTIGG-I(A/C)GIGG-3′ (I represents inosine) corresponding to the C-RFa sequence Pro^{6}–Gly^{13}. First round PCR products were reamplified using the anchor primer and the second degenerate primers 5′-GGI(A/C)GIGGIGTI(A/C)GICCIAT(A/T/C)GG-3′, corresponding to the sequence Gly^{11}–Gly^{18}. Amplifications were performed for 30 cycles at 94°C for 30 s, at 55°C for 30 s, and 1 min at 72°C (5 min for the last cycle).

2.3.2. Rapid amplification of the cDNA 5′-end (5′-RACE). The template cDNA was synthesized using a primer complementary to nucleotides 560–582 (5′-CACTTGCCAGTCAAGTTTATTTG-3′), followed by dA-tailing of the cDNA using dATP and terminal transferase (Boehringer Mannheim, Germany). The tailed cDNA was amplified using the oligo-dT anchor primer and the gene-specific primer 1 (5′-ATTGAAGGACACACAGCCAG-3′, complementary to nucleotides 532–551) followed by reamplification of the first round PCR products using the anchor primer (Boehringer Mannheim, Germany) and the gene-specific primer 2 (5′-GGACACACAGCCAGATAAAG-3′, complementary to nucleotides 526–545). Both first and second rounds of PCR were performed for 30 cycles consisting of 30 s at 94°C, 30 s at 55°C, and 1 min at 72°C.

2.4. Subcloning and sequencing

The second round PCR products were subcloned into a TA cloning vector according to the manufacturer's instructions (Invitrogen, USA). The DNA inserts of the positive clones were amplified by PCR using universal M13 primers. All nucleotide sequences were determined using Big-Dye sequencing kits (Perkin Elmer, UK) and an automated DNA sequencer (Perkin Elmer, model 373A, UK), and analyzed on GENETYX-MAC software (Software Development, Japan). Universal M13 primers or gene-specific primers were used to sequence both strands. The nucleotide sequences reported in this paper have been submitted to the DDBJ/EMBL/GenBank nucleotide sequence databases with the accession number for the C-RFa cDNA and for the partial *Carassius* β-actin cDNA.

2.5. Northern blot hybridization

A full length digoxigenin C-RFa precursor cDNA was synthesized using a DNA labeling kit (Boehringer Mannheim, Germany) and used as a probe for Northern and Southern blot analysis. Total RNA was separated on a denaturing 1% agarose-formaldehyde gel and fixed onto abound N^{+} membrane (Amersham Life Science, UK) by UV irradiation. Hybridization and detection was carried out according to the manufacturer's standard procedure (Boehringer Mannheim, Germany). RNA size was estimated using DIG-labeled RNA molecular markers (Boehringer Mannheim, Germany).

2.6. Southern blot hybridization of RT-PCR products

First strand cDNA was synthesized from total RNA (1 μg) prepared from each tissue using Superscript II reverse transcriptase and an oligo-dT primer. The oligonucleotide primer set used for amplification of C-RFa cDNA fragments was 5′-GTTACTGAAGAGCATACGTCC-3′ (identical to nucleotides 40–60) and 5′-ATTGAAGGACACACAGCCAG-3′ (complementary to nucleotides 532–551), and primers for amplification of β-actin cDNA fragments were 5′-ATGTGCAAAGCCGGATTCGC-3′ (identical to nucleotides 52–71 in the zebrafish β-actin cDNA) and 5′-GAACCTCTCATTGCCAATGG-3′ (complementary to nucleotides 752–771 in the zebrafish β-actin cDNA). PCR products were resolved on a 1.5% agarose gel followed by transfer to Hybond N^{+} membrane. Hybridization and detection were carried out according to the procedure of the DIG System protocol (Boehringer Mannheim, Germany).

3. Results and discussion

3.1. Characterization of a cDNA encoding a precursor polypeptide of C-RFa

In an attempt to obtain C-RFa precursor polypeptide cDNA fragments, we first performed an RT-PCR experiment using degenerate primers corresponding to the partial C-RFa sequence Pro^{6}–Gly^{13} and the anchor primer (Section 2) followed by further amplification of the PCR products using degenerate primers corresponding to the sequence Gly^{11}–Gly^{18}. Sequencing of the subcloned PCR products showed that all clones included essentially identical nucleotide sequences except for minor differences in the 3′-terminal sequence attributable to various lengths of the poly(A) tract. The predicted amino acid sequence comprised the sequence Arg-Phe-Gly-Lys-Arg located immediately after a partial C-RFa sequence derived from the second round PCR primers. To determine the 5′-end sequence, 5′-RACE using specific primers for the clone (Section 2) was performed, revealing that the 5′-RACE products contained the only putative ATG initiation codon. The RT-PCR and 5′-RACE products amplified using different reverse transcriptases and/or polymerases contained identical nucleotide sequences, confirming that these cDNA clones were not generated by artifacts.

By combining nucleotide sequences determined by these experiments, the entire cDNA sequence encoding a C-RFa precursor was identified. Fig. 1 shows the complete sequence of the longest cDNA. The C-RFa precursor cDNA is composed of 722 nucleotides, consisting of a 96 bp 5′-untranslated region (UTR), a 354 bp open reading frame (ORF), and a 272 bp 3′-UTR followed by various lengths of poly(A) tail. The ORF region begins with the putative start codon present at position 97 and terminates with a TGA stop codon at position 448. Four polyadenylation signals (AATAAA) were found in the 3′-UTR at positions 561, 619, 623, and 699. Nucleotide sequence analysis of all clones indicated that the last polyadenylation signal was used much more frequently than other signals, although the biological significance of this remains unclear. Northern blot analysis of total RNA using a DIG-labeled C-RFa precursor cDNA as a probe detected a single band of ca. 0.8 kb (Fig. 1B) even after longer exposure (not shown), suggesting that the C-RFa gene produces a single transcript. The apparent migration of the 0.8 kb was well in accordance with the estimated length of the cDNA, confirming that the longest cDNA includes a full-length nucleotide sequence encoding the C-RFa precursor polypeptide.

The ORF region encoded a 117 residue polypeptide. The precursor consists of a long signal peptide, a spacer sequence, a single C-RFa sequence, and a functionally unknown C-terminal domain. The cleavage site of the signal peptide is predicted to be the Ala^{37}–His^{38} bond according to the −3, −1 rule [16]. A single copy of the C-RFa sequence, flanked by

Table 1
Amino acid sequences of C-Rfa and its homologous peptides

Peptide	Sequence	Animal	Ref.
C-Rfa	**SPEIDPFWYVGRGVRPIGRFamide**	**Crucian carp**	[12]
PrRP20 (bovine)	TPDINPAWYAGRGIRPVGRFamide	Bovine	[13]
PrRP31 (bovine)	SRAHQHSMEIRTPDINPAWYAGRGIRPVGRFamide	Bovine	[13]
ACEP-1	SGQSWRPQGRFamide	*Achatina fulica*	[2]
HCIP	VFQNQFKGIQGRFamide	*Helix aspersa*	[4]
LyCEP	TPHWRPQGRFamide	*Lymnaea stagnalis*	[3]

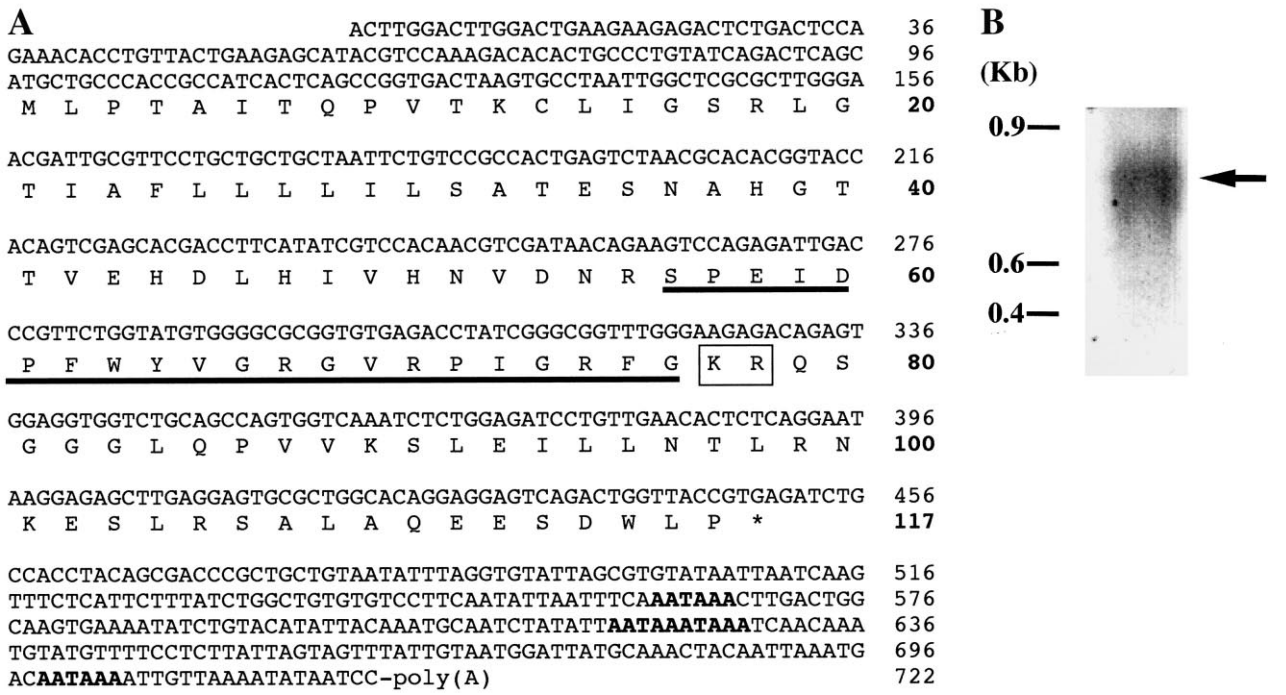

Fig. 1. A: Nucleotide sequence and deduced amino acid sequence of the C-RFa cDNA. The C-RFa sequence is underlined. The dibasic cleavage site is boxed. The potential poly(A) adenylation signals AATAAA are indicated in bold face type. B: Northern blot analysis of total RNA. Total RNA was extracted from the brain of the carp and about 25 μg of RNA was subjected to Northern blot hybridization using a DIG-labeled C-RFa cDNA probe. RNA molecular weight markers are shown on the left.

Gly as a C-terminus amidation signal and the typical dibasic processing sequence Lys-Arg, was located at position 56–75 (Fig. 1A). Interestingly, instead of the typical dibasic amino acid sites, a single Arg residue was found to precede the C-RFa sequence, and the amino acid sequence adjacent to this Arg residue is different from the furin-like endopeptidase cleavage sites Arg-Xaa-(Lys/Arg)-Arg [17–19]. Thus, it is likely that an unusual endoproteolytic processing mechanism is associated with the maturation of C-RFa.

3.2. Amino acid sequence comparison of the C-RFa precursor with a bovine preproPrRP

The amino acid sequence of the C-RFa precursor was aligned with a bovine preproPrRP polypeptide (Fig. 2). The C-RFa precursor displayed relatively low (29.9%) amino acid similarity with the preproPrRP, although the C-RFa sequence is 65% homologous to PrRP20. This low similarity is due to the difference in the amino acid sequence of the C-terminal domain between the C-RFa precursor and PrRP precursor, and the extraordinarily long signal peptide region of the C-RFa precursor. In addition, the C-terminal sequences of preproPrRPs show relatively high diversity even among the mammalian species [15]. The C-RFa precursor was found to share characteristic sequences at the N-terminal region with the preproPrRP, such as the Leu-rich hydrophobic core and putative mono-Arg processing site located at the N-termini of C-RFa and PrRP20 (Fig. 2). The latter findings indicate that the specific cleavage between the Arg residue and the N-terminus of C-RFa and PrRP20 occurs through a common processing mechanism. Taken together, C-RFa and PrRPs are likely to belong to the same peptide class.

3.3. Expression of C-RFa mRNA in carp tissues

The presence of the C-RFa transcript was investigated by Southern blot analysis of the RT-PCR products that were prepared from the telencephalon, optic tectum, cerebellum, medulla oblongata, olfactory bulbs, and proximal half eye ball. As an internal control, the presence of the β-actin cDNA fragment with the size of ca. 0.75 kb amplified using the primer set based on the zebrafish β-actin cDNA sequence

```
C-RFa precursor          MLPTAITQPVTKCLIGSRLGTIAFLLLLILSATESNAHGTTVEHDLHIVHNVDNRSPEID   60
Bovine PrRP precurosor   MK--AV---------G------AWLLLLGLAALQGAASRA------HQ-HSMEIRTPDIN   36
                         *   *          *      * **** * *    *         *  *    * * *

(continued)
C-RFa precursor          PFWYVGRGVRPIGRFGKRQS---GGGLQPVVKSLEILLNTLRNKESLRS-ALAQEESDWLP   117
Bovine PrRP precurosor   PAWYAGRGIRPVGRFGRRRAATPGDGPRPGPRRVPACFRLEGGAEP--SRALPGRLTAQLVQE   90
                         * ** *** ** **** *     * *  *               *   *  *
```

Fig. 2. Comparison of the amino acid sequence of the C-RFa precursor with the bovine preproPrRP. Asterisks denote identical amino acid residues among all precursors. Gaps marked by hyphens have been inserted to optimize homology.

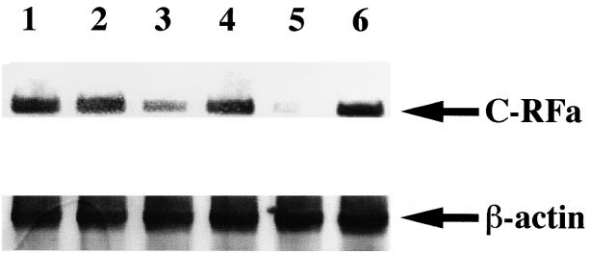

Fig. 3. Southern blot analysis of RT-PCR products for C-RFa (upper) and β-actin (lower) transcripts isolated from the telencephalon (lane 1), optic tectum (lane 2), cerebellum (lane 3), medulla oblongata (lane 4), olfactory bulbs (lane 5), and proximal half eye ball (lane 6). PCR products were resolved on a 1.5% agarose gel followed by transfer to a nylon membrane and hybridization with a DIG-labeled C-RFa or β-actin cDNA probe, respectively.

(Section 2) was observed. Sequence analyses showed that this RT-PCR products included an amino acid sequence that is approximately 95% homologous to the zebrafish β-actin, confirming the specific amplification of the carp β-actin cDNA fragment (data not shown). The β-actin transcript was demonstrated to be present in all tissues to a similar degree (Fig. 3, lower).

Expression of the C-RFa mRNA was detected predominantly in the telencephalon, optic tectum, medulla oblongata, and proximal half eye as shown in Fig. 3 (upper, lanes 1, 2, 4, and 6). Furthermore, the C-RFa mRNA was detected at a lower level in the cerebellum (lane 3). On the other hand, very little C-RFa transcript was found to be expressed in the olfactory bulbs (lane 5). These results obtained by Southern blot analysis of RT-PCR products for the C-RFa and β-actin mRNA lead to a conclusion that C-RFa is synthesized mainly in the telencephalon, optic tectum, medulla oblongata, and the proximal half eye ball.

The widely distributed expression of C-RFa gene in several cortex regions including the proximal half eye ball suggests that C-RFa is responsible for the regulation of multiple physiological events. One of the proposed functions of C-RFa is the control of the activity and/or sensitivity of the optic organ, at least in fish, because the abundant presence of C-RFa mRNA was observed in the proximal half eye ball including the retina, which is the central tissue regulating optic information processing. Moreover, in goldfish, RFamide-like immunoreactivities were shown to be localized in a sub-population of retinal amacrine cells as well as in nerve fibers of the nervus terminalis that project to the retina [10,11]. However, quite little C-RFa transcript was detected in the olfactory bulbs including the ganglial cells of the nervus terminalis, where RFamide immunolabelling was observed in goldfish [10,11]. In combination, these findings indicate the possibility that C-RFa produced in the retina, rather than in the nervus terminalis, participates in retinal communications as a signalling molecule through peptidergic pathways. In addition, RFamide-like immunoreactivities in the nervus terminalis of the goldfish may demonstrate the presence of RFamide-related peptides other than C-RFa. To investigate the peptidergic regulation or modulation mechanism of the retinal component cells by C-RFa, more detailed studies by both immunocytochemical and in situ hybridization experiments are necessary.

Also of particular interest is whether C-RFa serves as a fish prolactin-releasing factor. The PrRP mRNA, like the C-RFa mRNA, was found to be present at the highest level in the medulla oblongata [15], and C-RFa shows a high amino acid sequence homology with PrRPs, especially PrRP20 (Table 1). Furthermore, several unique amino acid sequences have been shown to be conserved between the C-RFa and PrRPs precursors (Fig. 2) as described above. Taken together, these findings imply that both C-RFa and PrRPs play similar physiological roles such as release of prolactin in fish and mammals, respectively. The prolactin-releasing activity of C-RFa in both mammals and fish is being examined. In the publication, we show the first identification of the cDNA encoding the *Carassius* RFamides, C-RFa, and the localizations of C-RFa transcript. Our data possibly contribute to further studies of the physiological roles of RFamides.

References

[1] Price, D.A. and Greenberg, M.J. (1977) Science 197, 670–671.
[2] Fujimoto, K., Ohta, N., Yoshida, M., Kubota, I., Muneoka, Y. and Kobayashi, M. (1990) Biochem. Biophys. Res. Commun. 167, 777–783.
[3] Tensen, C.P., Cox, K.J.A., Smit, A.B., van der Schors, R.C., Meyerhof, W., Richter, D., Planta, R.J., Hermann, P.M., van Minnen, J., Geraerts, W.P.M., Knol, J.C., Bruke, J.F., Vreugdenhil, E. and van Heerikhuizen, H. (1998) J. Neurosci. 18, 9812–9821.
[4] Baud, C., Darbon, P., Li, K.W. and Marchand, C.R. (1998) Cell. Mol. Neurobiol. 18, 413–424.
[5] Greenberg, M.J. and Price, D.A. (1992) Prog. Brain Res. 92, 25–37.
[6] Fujisawa-Sakata, M. and Kobayashi, M. (1994) Cell. Tissue Res. 278, 451–460.
[7] Shyamala, M., Fisher, J.M. and Scheller, R.H. (1986) DNA 5, 203–208.
[8] Alloyz, R.S. and des Groseillers, L. (1995) Peptides 16, 331–338.
[9] Dockray, G.J., Reeve Jr., J.R., Shinvery, J., Gayton, R.J. and Barnard, C.S. (1983) Nature 305, 328–330.
[10] Yang, H.Y., Fratta, W., Majene, E.A. and Costa, E. (1985) Proc. Natl. Acad. Sci. USA 82, 7757–7761.
[11] Steel, W.K., Walker, S.E., Chohan, K.S. and Ball, A.K. (1984) Proc. Natl. Acad. Sci. USA 81, 940–944.
[12] Fisher, A.J., Reish, H.M., Kyle, A.L. and Steel, W.K. (1996) Regul. Peptides 62, 73–87.
[13] Fisher, A.J. and Stell, W.K. (1997) Neuroscience 77, 585–597.
[14] Fujimoto, M., Takeshita, K., Wang, X., Takabatake, I., Fujisawa, Y., Terranishi, H., Ohtani, M., Muneoka, Y. and Ohta, S. (1998) Biochem. Biophys. Res. Commun. 242, 436–440.
[15] Hinuma, S., Habata, Y., Fujii, R., Kawamata, Y., Hosoya, M., Fukusumi, S., Kitada, C., Masuo, Y., Asano, T., Matsumoto, H., Sekiguchi, M., Kurokawa, T., Nishimura, O., Onda, H. and Fujino, M. (1998) Nature 393, 272–276.
[16] von Heijne, G. (1986) Nucleic Acids Res. 14, 4683–4690.
[17] Hosaka, M., Nagahama, M., Kim, W.S., Watanabe, T., Hatsuzawa, K., Ikemizu, J., Murakami, K. and Nakayama, K. (1991) J. Biol. Chem. 266, 12127–12130.
[18] Nakayama, K. (1997) Biochem. J. 327, 625–635.
[19] Day, R., Lazure, C., Basak, A., Boudreault, A., Limperis, P., Dong, W. and Linberg, I. (1998) J. Biol. Chem. 273, 829–836.