Volume 250, number 1, 85–90 FEB 07237 June 1989

Discussion Letter

A new family of growth factor-like peptides

'Trefoil' disulphide loop structures as a common feature in breast cancer associated peptide (pS2), pancreatic spasmolytic polypeptide (PSP), and frog skin peptides (spasmolysins)

Lars Thim

Novo Research Institute, DK-2880 Bagsvaerd, Denmark

Received 25 March 1989; revised version received 20 April 1989

Four peptides present in completely different biological sources have been shown to exhibit a large degree of structural similarity. The peptides include: (i) a 60 amino acid residue breast cancer associated pS2 peptide isolated from human gastric juice and the culture media of the human breast cancer cell line MCF-7; (ii) a 106 amino acid residue pancreatic spasmolytic polypeptide (PSP) isolated from porcine pancreas and pancreatic juice; and (iii) a 49 and 50 amino acid residue peptide predicted from a cDNA isolated from the skin of the frog, *Xenopus laevis*. These peptides are characterized by having one (pS2 and the frog peptides) or two (PSP) domains of a highly conserved 38–39 amino acid residue consensus sequence not found in any other known peptides or proteins. The domain sequences contain 6 cysteine residues in nearly the same positions and it is suggested that these 6 residues are linked by 3 disulphide bonds to form a characteristic 'trefoil' disulphide loop structure common in all four peptides. From the sources of which the peptides have been isolated and from experiments showing that PSP has a growth factor stimulatory effect on MCF-7 cells, it is further suggested that these peptides may represent members of a new family of growth factors.

Amino acid sequence comparison; Domain structure; Disulfide bond; Receptor binding

1. INTRODUCTION

A significant number of peptides with growth stimulatory effects on mammalian cells have been isolated and characterized over the last 15 years. These peptides include epidermal growth factor (EGF) [1], transforming growth factor alpha (TGFα) [2], transforming growth factor beta (TGFβ) [3], platelet-derived growth factor (PDGF) [4], insulin-like growth factors (IGF-I and IGF-II) [5–7], nerve growth factor [8], and heparin-binding growth factor (acidic fibroblast growth factor) [9]. These peptides are all characterized by having a certain number of cysteine residues linked by inter or intra chain disulphide bonds. The detailed disulphide bond configuration which is important for the tertiary structure and biological activity of the molecules has been elucidated for EGF, TGFα and IGF-I. These peptides all contain three disulphide bonds in a region comprising 36 (TGFα), 37 (EGF) or 56 (IGF-I) amino acid residues [1,2,5–7]. If the cysteine residues are numbered from the N-terminal end, EGF and TGFα have a disulphide bonding of 1-3, 2-4 and 5-6 [10,11] whereas IGF-I is believed to have a disulphide bonding of 1-4, 2-6, and 3-5 [12]. In the present paper I wish to suggest the existence of a

Correspondence address: L. Thim, Novo Research Institute, DK-2880 Bagsvaerd, Denmark

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

new family of growth factors characterized by having one or two domains, in which each domain consists of 38 or 39 amino acid residues and three disulphide bonds in the configuration of either 1-5, 2-4, 3-6 or 1-4, 2-5, 3-6. In either case, the disulphide bond configuration is different from the one in EGF/TGFα and IGF-I. The members of the new family constitute human breast cancer associated peptide (pS2), pancreatic spasmolytic polypeptide (PSP), and two frog skin peptides (spasmolysin I and II).

2. THE pS2 PEPTIDE

The gene encoding the human pS2 peptide has been discovered by differential screening of a cDNA library obtained from the human breast cancer cell line MCF-7 induced by estrogen [13,14]. The pS2 gene has been cloned by two groups [15,16]. The nucleotide sequence of a pS2 full-length cDNA encoded a peptide of 84 amino acid residues out of which 26 or 21 residues represented a putative signal peptide corresponding to a mature pS2 peptide of 58 or 63 residues [15]. By the use of an antibody raised against a synthetic peptide corresponding to a 31 amino acid residue sequence from the C-terminal end of the predicted pS2 peptide, it has been possible to immuno-precipitate the pS2 peptide from the culture media of MCF-7 cells [17]. Amino acid labelling experiments indicated that the pS2 peptide is secreted as a 58 amino acid residue peptide [17]. The pS2 peptide is present in the culture media of MCF-7 cells after estrogen induction at a concentration of 30–100 μg/l [18,19]. Recently, it has been possible to purify a sufficient amount of the peptide from this source for gas-phase protein sequencing. These analyses showed that the pS2 peptide is secreted from the MCF-7 cells as a 60 amino acid residue peptide [18,19]. By use of the above mentioned pS2 antibodies, a large variety of human tissues (including normal breast tissue and pancreas) has been screened for pS2-specific staining [20]. The pS2 peptide was found to be specifically expressed and secreted by the mucosa cells of the human stomach, and the concentration in the gastric juice was 30–100 μg/l [20]. Analysis of the purified gastric pS2 peptide indicated that the first 48 amino acid residues are identical to those of the secreted MCF-7 pS2 peptide, although the N-terminal amino acid residue of the gastric peptide may be present as a pyroglutamic acid [19]. However, the MCF-7 pS2 peptide and the gastric pS2 peptide do not show a completely identical tryptic mapping [19], and it cannot be excluded that the gastric peptide has some C-terminal post-translational modification (e.g. at Cys-58) different from that of the MCF-7 pS2 peptide.

The reason for the presence of the pS2 peptide in breast cancer cells and in the human gastric juice is unknown. Neither the pS2 peptide nor the pS2 mRNA have been found in significant amounts in normal breast tissue of women [21]. The pS2 peptide is also absent from the breast tissue late in pregnancy, during lactation, and in human milk [20]. The possibility of the pS2 peptide being an autocrine growth factor for MCF-7 cells has been examined by Davidson et al. [22]. From the study of the pS2 gene expression in two variants of MCF-7 cells they concluded that the pS2 peptide did not seem to play a role as a major autocrine growth factor [22]. However, the effect of the pS2 peptide on the growth of MCF-7 cells or other mammalian cells has not been investigated, and the possibility of the pS2 peptide being a paracrine growth factor still exists.

3. PANCREATIC SPASMOLYTIC POLYPEPTIDE

Pancreatic spasmolytic polypeptide (PSP) was originally discovered in a side fraction from the purification of porcine insulin [23]. The concentration of PSP in porcine pancreas is approx. 100 mg/kg [24], which is about half the content of insulin. The initial characterization of the peptide showed that it contains 106 amino acid residues in a single chain; the peptide is N-terminally blocked and heavily crosslinked by seven disulphide bridges [23]. In 1985, the complete amino acid sequence of PSP was published [25]. These studies showed that PSP is composed of two homologous domains probably derived from a common gene [25]. Recently, the amino acid sequence of PSP has been revised at four positions using a combination of mass spectrometry and Edman degradations [26]. The presence of a N-terminal pyroglutamic acid residue was also confirmed in this study [26]. The amino sequence of PSP is given in fig.1, which also shows the large degree of similarity between the

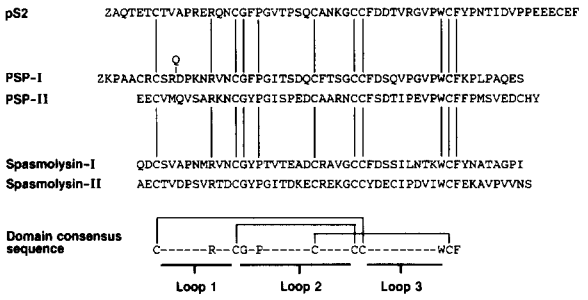

Fig.1. Alignment of the amino acid sequences for the human breast cancer associated pS2 peptide [15,16,18,19], porcine pancreatic spasmolytic polypeptide domain I and II [25,26], and the frog skin peptides spasmolysin I and II [35]. It should be noted that Gln-11 of PSP-domain I has been deleted in order to obtain exact alignment of cysteine residues. The suggested disulphide bond configuration is shown for the domain consensus sequence.

two domains of the molecule. In the two PSP domains (PSP(8–47) and PSP(58–96)), 19 out of 39 residues are identical.

Studies have been initiated to determine the disulphide bond configuration of PSP. These studies are rendered difficult by the remarkable resistance of the native PSP to proteolytic digestion [27]. By a combination of CNBr-treatment followed by thermolysin digestion it has been possible to divide the native PSP into two separate moieties; one containing residues 1–55 disulphide linked to residues 100–106 and another comprising residues 56–58 disulphide linked to residues 61–95. By further trypsin digestion of the latter fragment, a three chained and a two chained molecule was obtained; the three chained molecule comprises the residues 56–58, 62–72, and 79–84, and the two chained molecule comprises the residues 73–78 and 85–95 (Christensen, M., personal communications). From these results the disulphide bond configuration of PSP-domain II is either A: 78-95, 68-83, 58-84 or B: 78-95, 68-84, 58-83. Preliminary data from Edman degradation of the three chained molecule suggest possibility A. By homology to domain II the disulphide bond configuration in PSP-domain I (see fig.1) is either A: 29-46, 19-34, 8-35, or B: 29-46, 19-35, 8-34. Assuming possibility A, the entire structure of PSP including the seven disulphide bond can be deduced (fig.2). It should be emphasized that from the preliminary experimental data presently available

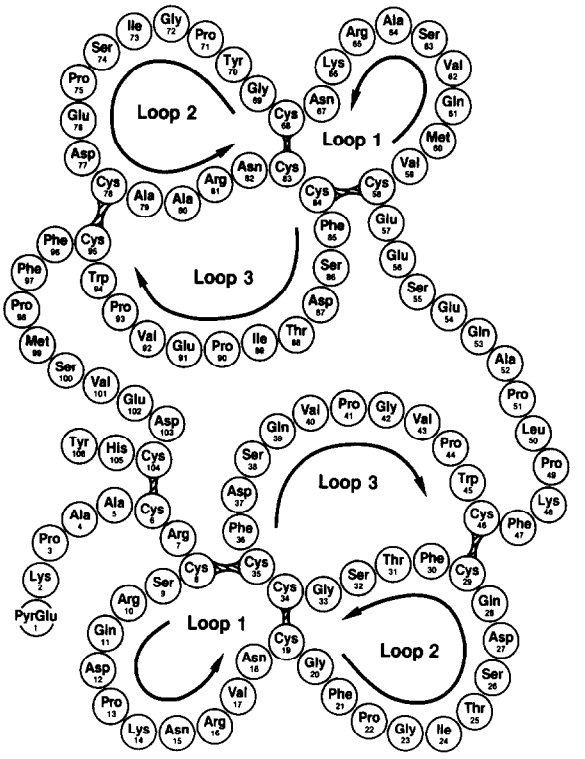

Fig.2. Proposed structure of porcine pancreatic spasmolytic polypeptide (PSP).

possibility B ('cross-over' of two disulphide bonds in each domain) cannot be excluded.

Several studies have been carried out to elucidate the possible physiological function of PSP. In extracts from 22 different porcine tissues the pancreas was found to be the only organ containing substantial amounts of PSP [24]. After stimulation with pancreozymin or secretin, large amounts of PSP (50–100 mg/l) were found to be secreted into the pancreatic juice, indicating that PSP is related to the exocrine function of the pancreas [24]. A pharmacological screening has indicated that PSP has spasmolytic, and gastric acid secretion inhibitory effects [28]. Although we were not completely certain that these effects reflected the physiological functions of the peptide, we decided to name it pancreatic spasmolytic polypeptide. Specific binding of PSP to rat intestinal mucosa cells and membrane preparations from these cells have been demonstrated [29,30]. These binding studies showed the presence of high-affinity, low-capacity sites (dissociation constant: 130 nM), and

low-affinity, high-capacity sites (dissociation constant 4.2 μM) [29]. The presence of PSP recognition sites (receptors) on intestinal mucosa cells in combination with the high concentration of PSP in pancreatic juice [24] and the remarkable resistance towards proteolytic digestion [27] suggests a unique intraluminal function of the peptide. Recently, it has been shown that PSP has growth stimulatory effect on the colon carcinoma cell, HCT 116, and the human breast cancer cell, MCF-7 [31]. The growth stimulatory effect of porcine PSP on human MCF-7 cells could very likely be due to cross-recognition with a pS2 site [31], thus indicating a very closely related tertiary structure of the porcine PSP and the human pS2 peptide.

4. SPASMOLYSIN I AND II

The amphibian skin has proven to be a valuable source for the isolation of new biological active peptides [32]. Thus peptides isolated from amphibian skin often have been found to have their counterpart in the brain and gut of mammalians. This has been the case with the cholecystokinin/gastrin-like peptide caerulein [33] and the neurotensin-like peptide xenopsin [34].

Recently, a new protein from the skin of the frog *Xenopus laevis*, highly homologous to PSP and pS2 has been cloned [35]. Processing at two dibasic sequences in N- and C-terminal parts of the molecule would lead to the generation of two peptides consisting of 49 and 50 amino acid residues, respectively [35]. Both of these peptides (spasmolysin I and II) contain 6 half-cystine residues, and the sequences are highly homologous to the pS2 peptide and the two domains of PSP [35] (fig.1). The biological function of these peptides in the frog skin is unknown. Due to their structural similarities with PSP/pS2 and their coexistence with biological active peptides in the frog skin it is tempting to suggest that spasmolysin I and II have a growth stimulatory effect on at least some of the cells in the frog skin.

5. SEQUENCE COMPARISON

Fig.1 shows a comparison of the sequences of the pS2 peptide, PSP and the frog skin peptides. The sequence of the 60 amino acid residue pS2 peptide is taken from [15,16,18,19]. The N-terminal amino acid residue (Z) is glutamic acid in the pS2 peptide isolated from the culture media of MCF-7 cells [18,19] and probably pyroglutamic acid in the pS2 peptide isolated from human gastric juice [19]. The sequence of PSP is taken from [26], in which the original PSP sequence [25] was revised at four positions. The N-terminal amino acid residue (Z) is pyroglutamic acid [25,26]. The 106 amino acid residue PSP has been divided into two homologous domains, PSP-I (residue 1–55) and PSP-II (residue 56–106). The sequences of the frog skin peptides (spasmolysin I and II) have been taken from [35] assuming processing of the 'spasmolysin precursor' at Ala-20, Gln-21 (signal peptidase) and at the two dibasic sequences Lys-70, Lys-71 and Arg-349, Lys-350 as suggested by Hoffmann [35]. With the exception of a single residue (e.g. Gln-11) in PSP-domain I it is possible to align all five domains without introducing gaps or insertions in a configuration which places all of the 6 half-cystine residues in exactly the same positions (fig.1). This surprising sequence homology has previously been noticed [35–37]. Besides some rather weak homology to the kringle regions of plasminogen [25,36], a search in the latest version (Ver. 19, updated February 1989) of the Dayhoff protein sequence database (National Biomedical Research Foundation, Georgetown, Washington DC, USA) using the programs SEARCH, RELATE and FASTP [38] did not pick up any protein or peptide with significant homology to the domain structures (author's observation).

From the alignments it is possible to deduce a domain consensus sequence comprising 38 or 39 amino acid residues (fig.1).

6. 'TREFOIL' DISULPHIDE LOOP STRUCTURES

Although it is possible to align the individual domains (fig.1) two by two to obtain a larger degree of homology, it is highly probable that the conserved part constitutes the six half-cystine residues and may be the Trp-Cys-Phe sequence in the C-terminal part of the domains. Thus the conserved element seems to be more related to the disulphide bond configuration and the *number* of residues in the three loop (the secondary and tertiary struc-

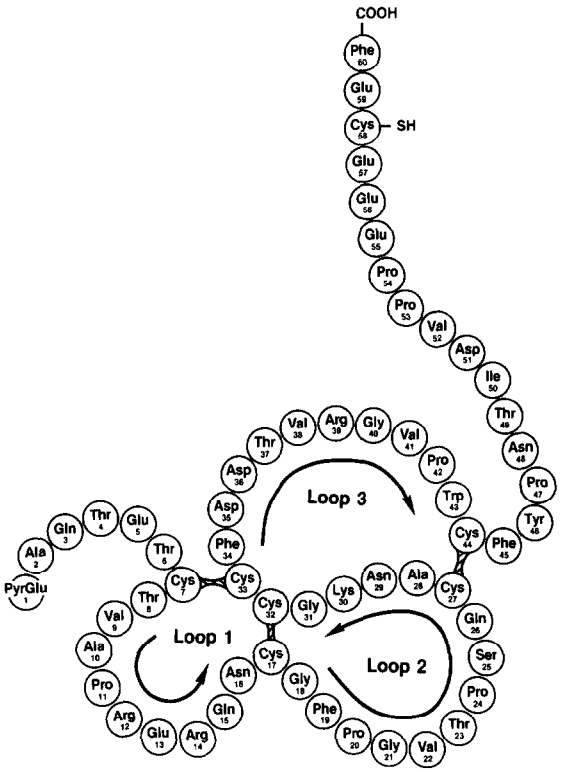

Fig.3. Proposed structure of human breast cancer associated pS2 peptide.

ture) than it is to the exact nature of the amino acid residues within the individual loops. Although the disulphide bond configuration is unknown for the pS2 peptide and frog peptides and only tentatively known for PSP as previously discussed, the sequence alignments have prompted me to suggest a trefoil disulphide loop structure common in all five domains (fig.4). From this structure, which is based on the preliminary assignment of disulphide bonds in PSP (fig.2), it is possible to suggest the structure of the pS2 peptide (fig.3) and the frog skin peptides (not shown).

Using antibodies against porcine PSP, no significant amount of immunoreactive material could be detected in human pancreatic extract or pancreatic juice (author's observation). It is interesting to note that also the human pS2-specific antibodies did not pick up any pS2-like material in the human pancreas [20]. As extracts from the human stomach, human gastric juice or MCF-7 cell media have not been investigated using the porcine PSP antibodies, the possible immunochemical cross-reactivity between the human pS2 peptide and the porcine PSP still needs to be determined.

However, the observed growth stimulatory effect of porcine PSP on human MCF-7 cells may be mediated by a recognition site (receptor) specific for the trefoil structure of PSP and pS2 [31]. This recognition site may be similar to the PSP-binding site (receptor) identified in rat mucosa cells and membrane preparations from these cells [29,30]. Although cross-recognition between the pS2 peptide and PSP may exist on the receptor level, further studies are needed to fully evaluate to which degree the suggested structural similarities of the pS2 peptide, the PSP, and the frog skin peptide imply similar biological effects such as growth stimulation.

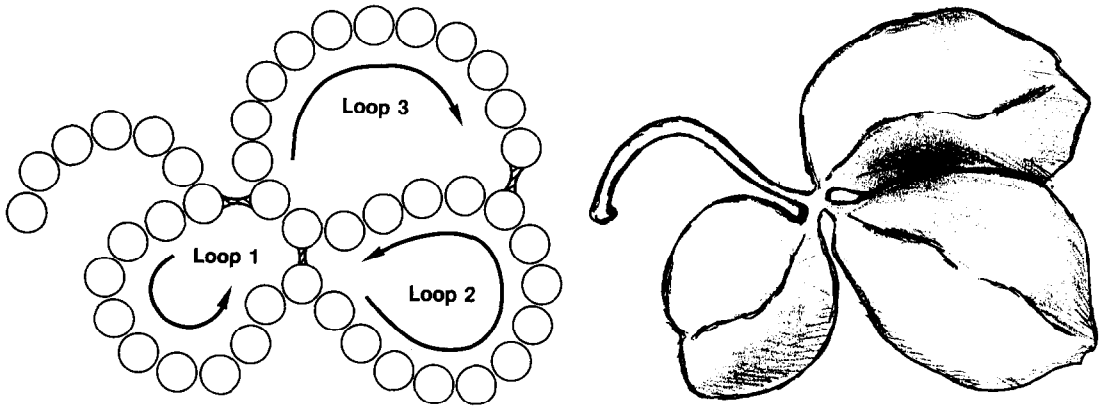

Fig.4. The 'trefoil' disulphide loop structure suggested as a common feature in the pS2 peptide, PSP and the frog skin peptides.

Acknowledgements: The author is grateful to Dr M. Christensen for making available the preliminary data on disulphide bonding in PSP and to Dr K.H. Jørgensen and Dr B.A. Nexø for valuable discussions.

REFERENCES

[1] Gregory, H. and Preston, B.M. (1977) Int. J. Pept. Protein Res. 9, 107–118.
[2] Derynck, R., Roberts, A.B., Winkler, M.E., Chen, E.Y. and Goeddel, D.V. (1984) Cell 38, 287–297.
[3] Derynck, R., Jarrett, J.A., Chen, E.Y., Eaton, D.E., Bell, J.R., Assoian, R.K., Roberts, A.B., Sporn, M.B. and Goeddel, D.V. (1985) Nature 316, 701–705.
[4] Betsholtz, C., Johansson, A., Heldin, C.H., Westermark, B., Lind, P., Urdea, M.S., Eddy, R., Shows, T.B., Philpott, K., Mellor, A.L., Knott, T.J. and Scott, J. (1986) Nature 320, 695–699.
[5] Rinderknecht, E. and Humbel, R.E. (1978) J. Biol. Chem. 253, 2769–2776.
[6] Rotwein, P., Pollock, K.M., Didier, D.K. and Krivi, G.G. (1986) J. Biol. Chem. 261, 4828–4832.
[7] Rinderknecht, E. and Humbel, R.E. (1978) FEBS Lett. 89, 283–286.
[8] Ullrich, A., Gray, A., Berman, C. and Dull, T.J. (1983) Nature 303, 821–825.
[9] Harper, J.W., Strydom, D.J. and Lobb, R.R. (1986) Biochemistry 25, 4097–4103.
[10] Savage, C.R., jr, Hash, J.H. and Cohen, S. (1973) J. Biol. Chem. 248, 7669–7672.
[11] Winkler, J. (1986) J. Biol. Chem. 261, 13838–13843.
[12] Blundell, T.L. and Humbel, R.E. (1980) Nature 287, 781–787.
[13] Masiakowski, P., Breathnach, R., Bloch, J., Gannon, F., Krust, A. and Chambon, P. (1982) Nucleic Acids Res. 10, 7895–7903.
[14] Brown, A.M.C., Jeltsch, J.M., Roberts, M. and Chambon, P. (1984) Proc. Natl. Acad. Sci. USA 81, 6344–6398.
[15] Jakowlew, S.B., Breathnach, R., Jeltsch, J.M., Masiakowski, P. and Chambon, P. (1984) Nucleic Acids Res. 12, 2861–2878.
[16] Prud'homme, J.-F., Fridlansky, F., Le Cunff, M., Atger, M., Mercier-Bodart, C., Pichon, M.-F. and Milgrom, E. (1985) DNA 4, 11–21.
[17] Nunez, A.-M., Jakowlev, S., Briand, J.-P., Gaire, M., Krust, A., Rio, M.-C. and Chambon, P. (1987) Endocrinology 121, 1759–1765.
[18] Mori, K., Fujii, R., Kida, N., Ohta, M. and Hayaski, K. (1988) Biochem. Biophys. Res. Commun. 155, 366–372.
[19] Rio, M.-C., Lepage, P., Diemunsch, P., Roitsch, C. and Chambon, P. (1988) C.R. Acad. Sci. 307, 825–832.
[20] Rio, M.C., Bellocq, J.P., Daniel, J.Y., Tomasetto, C., Lathe, R., Chenard, M.P., Batzenschlager, A. and Chambon, P. (1988) Science 241, 705–708.
[21] Rio, M.C., Bellocq, J.P., Gairard, B., Rasmussen, U.B., Krust, A., Koehl, C., Calderoli, H., Schiff, V., Renaud, R. and Chambon, P. (1987) Proc. Natl. Acad. Sci. USA 84, 9243–9247.
[22] Davidson, N.E., Bronzert, D.A., Chambon, P., Gelmann, E.P. and Lippman, M.E. (1986) Cancer Res. 46, 1904–1908.
[23] Jørgensen, K.H., Thim, L. and Jacobsen, H.E. (1982) Regul. Peptides 3, 207–219.
[24] Thim, L., Jørgensen, K.H. and Jørgensen, K.D. (1982) Regul. Peptides 3, 221–230.
[25] Thim, L., Thomsen, J., Christensen, M. and Jørgensen, K.H. (1985) Biochim. Biophys. Acta 827, 410–418.
[26] Rose, K., Savoy, L.-A., Thim, L., Christensen, M. and Jørgensen, K.H. (1989) in preparation.
[27] Jørgensen, K.H., Jørgensen, K.D. and Thim, L. (1983) U.S. Patent 4.307.317.
[28] Jørgensen, K.D., Diamant, B., Jørgensen, K.H. and Thim, L. (1982) Regul. Peptides 3, 231–243.
[29] Frandsen, E.K., Jørgensen, K.H. and Thim, L. (1986) Regul. Peptides 16, 291–297.
[30] Frandsen, E.K. (1988) Regul. Peptides 20, 45–52.
[31] Hoosein, N.M., Thim, L., Jørgensen, K.H. and Brattain, M.G. (1989) FEBS Lett. 247, 303–306.
[32] Erspamer, V. (1983) Trends Neurosci. 6, 200–201.
[33] Anastasi, A., Bertaccini, C., Cei, J.M., De Caro, G., Erspamer, V., Impiccioatore, M. and Roseghini, M. (1970) Br. J. Pharmacol. 38, 221–228.
[34] Araki, K., Tachibana, S., Uchyama, M., Nakajima, T. and Yasuhara, T. (1973) Chem. Pharmacol. Bull. 21, 2801–2804.
[35] Hoffmann, W. (1988) J. Biol. Chem. 263, 7686–7690.
[36] Baker, M.E. (1988) Biochem. J. 253, 307–309.
[37] Thim, L. (1988) Biochem. J. 253, 309.
[38] Lipman, D.J. and Pearson, W.R. (1985) Science 227, 1435–1441.