

cDNA deduced procionin

Structure and expression in protochordates resemble that of procholecystokinin in mammals

Hans-Jürg Monstein^{*,**}, Jan Ulrik Thorup, Ronnie Folkesson, Anders H. Johnsen, Jens F. Rehfeld^{*}

Department of Clinical Biochemistry, Rigshospitalet, University of Copenhagen, DK-2100 Copenhagen, Denmark

Received 21 July 1993

Using an improved 3' RACE (PCR) amplification system containing oligonucleotide primer with an inosine at ambiguous codon positions and inverse PCR to amplify the 5' ends, we have isolated and characterized cDNA clones which encode cionin, a protochordate homologue of the mammalian hormones, cholecystokinin (CCK) and gastrin. The full-length cloned cDNA of 510 bp encoded a 128 amino acid preprocionin. Reverse transcription-PCR and subsequent cDNA cloning revealed that cionin mRNA is expressed in both the neuronal ganglion and the gut of the protochordate *Ciona intestinalis*. The primary structure of procionin resembles that of proCCK more than that of progastrin. Sequence-specific immunochemical analysis showed that the cionin gene is expressed also at peptide level in both the gut and the neural ganglion. The neuronal processing of procionin is, however, more complete both with respect to carboxyamidation and tyrosine O-sulfation. Hence, the tissue-specific expression of the cionin gene in *Ciona intestinalis* resembles that of the CCK gene in mammals.

cDNA cloning, Cholecystokinin; *Ciona intestinalis*; Cionin, Gastrin, Prohormone processing

1. INTRODUCTION

Cholecystokinin (CCK) and gastrin are important mammalian brain and gut hormones (for review, see [1]). CCK and gastrin have the same C-terminal active site, Gly-Trp-Met-Asp-Phe-NH₂. The binding to either CCK or gastrin receptors is determined by the exact position of an adjacent tyrosyl residue. Thus, tyrosine in position six (as counted from the C-terminus) renders the peptide gastrin-like, whereas tyrosine in position seven is characteristic for CCK. The active site homology indicates that CCK and gastrin may have a common ancestor [2].

In evolution, protochordates occupy a crucial position immediately prior to vertebrates. Consequently, in a search for a common ancestor we recently isolated from the neural ganglion of the protochordate, *Ciona intestinalis*, a disulfated octapeptide amide named cionin [3]. Cionin has sulfated tyrosyl residues in both positions six and seven in addition to a pentapeptide amide identical to the common C-terminus of CCK and gastrin. Thus, cionin is structurally a hybrid of CCK and gastrin, and therefore related to the putative common ancestor.

The unique structure of cionin raises a number of

questions. Does procionin resemble proCCK more than progastrin? Is procionin a polypeptide like the precursors of the related frog skin peptide, caerulein [4]? Is the structure of procionin N-terminal to the cionin octapeptide sequence sufficiently acidic to be in accordance with the suggested consensus site for tyrosyl sulfation [5]? Is cionin – like CCK and gastrin – also expressed in the gastrointestinal tract? In order to answer these questions, we have now cloned the cionin cDNA and examined the cionin gene expression in *Ciona intestinalis*. The obvious question of the effect of cionin in mammals was recently answered in studies which showed an activity of cionin similar to that of CCK [6].

2. MATERIALS AND METHODS

2.1. Amplification and cloning of the cionin 3' end

3' RACE (rapid amplification of cDNA ends) was accomplished according to the procedure of Frohman et al. [7]. Using a Gene Amp RNA-PCR kit (Perkin Elmer Cetus, Norwalk, CT, US), oligo-d(T)₁₆-primed cDNA synthesis (single stranded) was performed with 1 µg of poly(A)⁺ RNA from the *Ciona* gut and neuronal ganglion, respectively. Single stranded cDNA was amplified with AmpliTaq DNA polymerase (Perkin Elmer Cetus) and a deoxymyosine-containing sense oligonucleotide (#1, Table I) and oligo-d(T)₁₅ NotI primer/adaptor, using a DNA Thermal Reactor (Hybaid Inc., Middlesex, UK). 33 cycles of amplification were carried out as follows: 95°C, 3 min; 31°C, 3 min; 70°C, 3 min; (3 cycles); 93°C, 3 min; 42°C, 3 min, 70°C, 3 min; (30 cycles), followed by a 15-min final extension at 70°C. The PCR product was cleaved with NotI and EcoRI and cloned into the transcription vector pGEM-11 zf(-) (Promega Inc., Madison, WI, US). This construct was used to transform competent DH5α cells. Plasmid preparations [8] and subsequent sequence analysis yielded three positive clones.

*Corresponding authors. Fax: (45) (35) 375 540.

**Present address: Department of Clinical and Medical Microbiology, University Hospital Linköping, S-58185 Linköping, Sweden.

Table I

Positions and DNA sequences of the primers used in this study to amplify *Cionin* DNA by PCR techniques as described in section 2

Name	Primer no.	Sequence, 5'-to-3' orientation	Start position in cDNA	Restriction enzyme site
Cionin-26mer sense	#1	GAATTCAAITAITAIGGITGGATGGAITTIGG	374 →	<i>EcoRI</i>
Cionin-26mer antisense	#2	CCIAAITCCATCCAICCCITAITAITT	399 ←	-
NTR-3' antisense	#4	GGATCCAACGTTACATTTTACGATGAATATG	463 ←	<i>BamHI</i>
NTR-3' sense	#7	GAATTCGTTTTCTTCGCTCAA	461 →	<i>EcoRI</i>
Cionin-5' sense	#9	GTCGACCAATGAGCATGGGAAGCAATATCGTTA	42 →	<i>SalI</i>
Cionin-5' sense oligo-d(T) Not	#10	GTCGACTATCATTGTTATAGTAACGTTAAATGT	79 →	<i>SalI</i>
		Commercially available	poly(A) tail	<i>NotI</i>

2.2. Amplification and cloning of the cionin 5' end by inverse PCR

1 µg of poly(A)⁺ RNA was used for oligo-d(T)/*Xba*-primed cDNA synthesis using Promega's cDNA kit. Inverse PCR was carried out as described [9], using cionin specific primers #4 and #7 (Table I). 32 cycles of amplification were performed: 94°C, 1 min; 37°C, 1 min; 72°C, 2 min (2 cycles); 92°C, 1 min; 42°C, 1 min, 72°C, 2 min: (30 cycles), followed by a final 15 min extension at 72°C. The PCR-amplified product of ~650 bp was digested with *EcoRI* and *BamHI* restriction enzymes and cloned into plasmid vector pGEM-3 (Promega Inc.). Recombinant plasmids containing appropriate inserts were sequenced using SP-6, T7, and cionin-26 oligomer sequencing primers and a commercially available DNA sequencing Kit (Sequenase, Version 2.0, US Biochemical Corp., Cleveland, OH, US).

2.3. Northern blot analysis

Total RNA from the gastrointestinal tract and neuronal ganglion was extracted as earlier described [10]. Electrophoretic separation of total RNA on a 1.2% agarose-0.7% formaldehyde gel was followed by capillary transfer onto Hybond-N membrane (Amersham, Birkerød, Denmark) in 25 mM sodium phosphate buffer, pH 6.4. Filters were fixed by UV-illumination (Stratagene, UV box). Prior to transfer, RNA preparations were routinely checked for degradation by ethidium bromide staining of the agarose gel. Hybridization analysis was performed using the cionin pC-III_m cDNA template, obtained by 3' RACE cloning of poly(A)⁺ RNA from the gastrointestinal tract under standard riboprobe conditions, as described [11].

2.4. Immunochemical analysis

The glycine-extended intermediate precursors of cionin were measured with antiserum #3208 using ¹²⁵I-labeled glycine-extended gastrin-13 as tracer and sulfated glycine-extended CCK-8 as standard [12]. Carboxyamidated cionin peptides were measured with antiserum #2609, which cross-reacts fully with sulphated and non-sulphated CCK peptides, as well as with carboxyamidated gastrin peptides [13]. Tyrosine-sulfated and carboxyamidated CCK peptides were also

measured using antiserum G160, which binds only sulfated cionin peptides [14].

2.5. Enzymatical cleavage

In order to measure peptides extended beyond glycine, i.e. pro-cionin, extracts and chromatographic fractions were cleaved with trypsin (1 mg trypsin/ml sodiumphosphate solution, pH 8.2) and carboxypeptidase B before measurements using antiserum #3208 as earlier described ([12], see also Fig. 4).

2.6 Chromatography

Extracts of 1 ml were applied to calibrated Sephadex G50 superfine columns (10 × 1,000 mm) and eluted at 4°C with 0.125 M NH₄HCO₃ at a flow rate of 4 ml/h. Fractions of 1.0 ml were collected. The void volume of the columns were indicated by ¹²⁵I-albumin and the total volume by ²²NaCl. The columns were further calibrated with synthetic cionin, glycine-extended CCK-8 (sulfated), and carboxyamidated CCK-33. Moreover, the columns were calibrated with neutral and acid extracts of porcine small intestine and cerebral cortex.

3. RESULTS AND DISCUSSION

3.1. Amplification and cloning of the cionin 3' end by the inosine primer-PCR cloning (IPPC) technique

Synthetic deoxyribo-oligonucleotides, the sequences of which were deduced from those of known amino acid sequences, have been used as powerful tools in cloning cDNAs. Most often mixed oligonucleotide pools, consisting of oligonucleotides with sequences degenerate at the ambiguous codon positions, have been used. However, Ohtsuka et al. [15] described an alternative approach using oligonucleotides with deoxyinosine in-

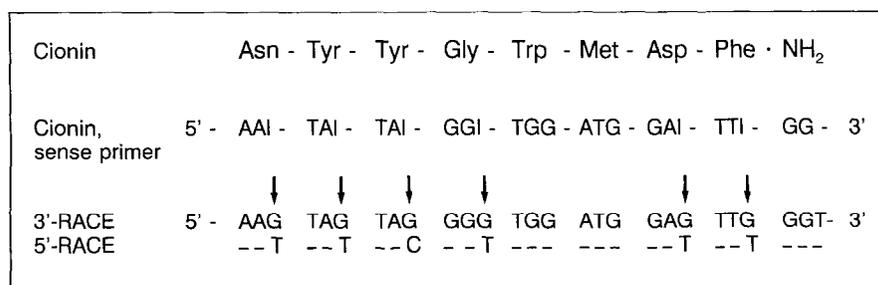


Fig. 1. The nucleotide sequence of Cionin-26 mer, containing deoxyinosine residues at ambiguous codon positions. The cionin amino acid sequence was as previously described [3]. Arrows indicate insertion of G residues at inosine positions (3' RACE). T or C indicate the correct nucleotides at ambiguous codon positions, obtained by inverse PCR cloning.

	5' GTCGACACGAAGCGTCGGT	GATTGGTTTTTCTGGAAAATTTCAATGAGC	49
1	MetGlySerAsnIleValIleTyrPheSer ATGGGAAGCAATATCGTTATTTATTTTCT	IleIleValIleValThrLeuAsnValAsn ATCATTTGTTATAGTAACGTTAAATGTGAAT	109
21	GlyValProAlaSerAspLeuPheLysSer GGCGTTCCAGCATCTGATTTGTTCAAGTCC	ValSerGlnTyrHisIleProArgSerLys GTGTCGCAATACCACATTCCTCCGGAAGCAAA	169
41	ValIleAsnLysGluThrValThrLysPro GTCATCAATAAAGAGACCGTAACAAAACCG	LeuGlnPheGlnArgAlaIleCysArgLeu CTTCAGTTTCAAAGAGCTATTTGCCGGCTC	229
61	LeuGlnLysLeuGlyGluGluThrPheAla TTGCAAAAACCTGGGGAGGAAACATTTGCT	ArgLeuSerGlnSerGluLeuGluAlaLys CGGTTGTCTCAATCGGAACCTGAAGCAAAA	289
81	GlnLeuAspLeuIleLysThrCysTyrGln CAGCTGGATTTAATCAAAAACCTGCTACCAG	AlaAsnSerPheGlyAspAsnGluAsnGln GCAAACAGCTTTGGGGATAATGAGAATCAG	349
101	GlyHisMetGlnArgMetAspArgAsnTyr GGTCATATGCAAGAATGGATCGAAACTAT	TyrGlyTrpMetAspPheGlyLysArgAla TACGGCTGGATGCATTTTGGTAAAGAGCA	409
121	IleGluAspValAspTyrGluTyrEnd ATCGAAGATGTTGATTATGAATATTAAGAA	CATATTCATCGTAAAATGAACGTTTCTCTC	469
	GCTCAATGCTTATATTGTTTAAATAAAATTT	CGTGCACGCAAAAAAAAAAAAAA 3'	529

Fig. 2. Nucleotide and predicted amino acid sequence of the protochordan pre-procionin peptide. The consensus poly(A)⁺ addition site, AATAAA, is underlined.

serted at ambiguous codon positions. The development of methods for the rapid amplification of 3' and 5' cDNA ends, the RACE techniques [7,16], using single gene-specific oligonucleotide primers, prompted the use of a deoxyinosine-containing cionin-specific oligoprimers in an effort to clone the cionin cDNA. Based on the cionin amino acid sequence, a 26-oligomer was synthesized, containing deoxyinosines at six ambiguous codon positions (Table I, Fig. 1). This deoxyinosine oligomer primer was used in amplification of the cionin cDNA 3' end and the PCR product was subsequently cloned and sequenced.

Primed cDNA synthesis was performed with poly(A)⁺ RNA from both the gastrointestinal tract and neuronal ganglion, and the single-stranded cDNAs were subjected to DNA polymerase amplification with the 26-sense oligomer (Table I) and oligo(dT)₁₅ *NotI* primer/adaptor. The amplification resulted in a small ~170 bp product. The product was cleaved with *NotI* and *EcoRI* and cloned into pGEM11 *zf(-)* vector. This construct was used to transform competent DH5 α cells. Ampicillin resistant colonies were selected at random and subsequently plasmid mini-preparations and restriction enzyme analysis were performed. Southern blot analysis with the 26-sense oligomer as probe revealed three positive clones, named pCIII_{M2}, pCIII_{M3}, and pCIII_{G7}, containing a PCR-amplified CCK-like sequence. pCIII_{M2} and pCIII_{M3} contain amplified cDNA synthesized from gastrointestinal mRNA, and pCIII_{G7} contains amplified cDNA synthesized from ganglionic mRNA.

Sequence analysis of the three partial cDNAs revealed that cionin mRNA is expressed in both the neuronal ganglion and the gastrointestinal tract. When the resulting recombinant plasmid DNAs were sequenced, all the positions for deoxyinosine were replaced by deoxyguanosine (G), indicating that the *Taq* polymerase selected deoxycytidine to pair with deoxyinosin (Fig. 1).

3.2. Amplification and cloning of the cionin cDNA 5' end by the inverse PCR technique

The inverse polymerase chain reaction (IPCR) has been used to amplify low-abundance sequences, i.e. the human deoxycytidine kinase (dC Kinase) mRNA [9]. The cloning and subsequent sequence analysis of the partial cionin cDNAs, obtained by the IPPC technique (see above) allowed us to design a set of 3' end-specific cionin PCR oligoprimers (Table I, #4 and #7). cDNA from the gastrointestinal tract was circularized using T4 DNA and T4 RNA ligase as described [9], and cDNA was used in 100 μ l reactions with oligoprimers #4 and #7 (Table I) in order to amplify the cDNA derived from the gastrointestinal tract mRNA. Amplified PCR prod-

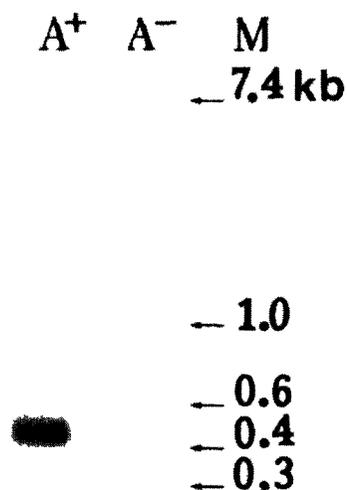


Fig. 3. Northern blot analysis of poly(A)⁺ and poly(A)⁻ RNA, isolated from the gastrointestinal tract of *Ciona intestinalis*. The Hybond-N membrane was subsequently hybridized with a [³²P]UTP-labelled cionin cRNA probe, derived from the DNA template, pCIII_m. M indicates the positions of RNA size markers.

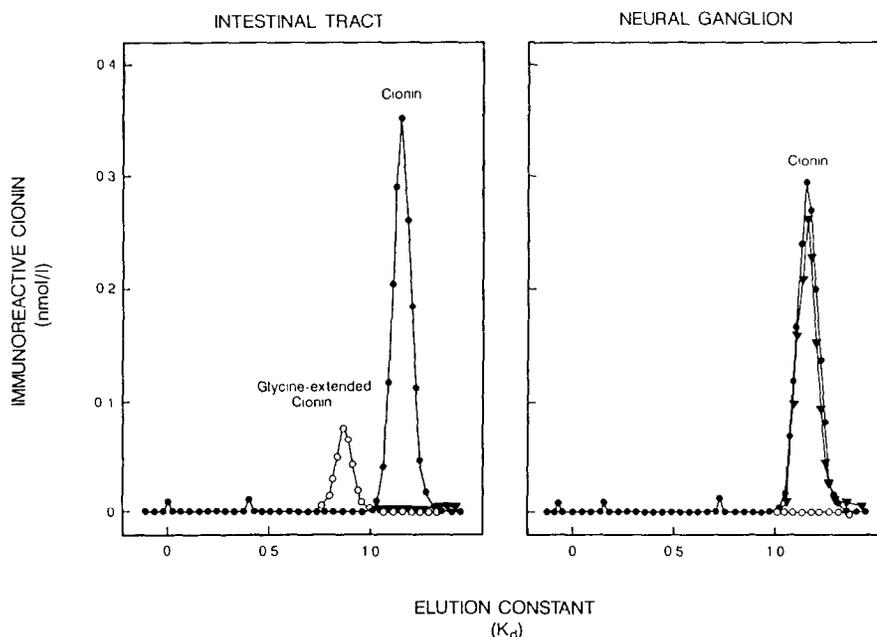


Fig. 4. Gel chromatography of extracts from the gastrointestinal tract (left) and the neural ganglion (right) of *Ciona intestinalis*. 1 ml of the extracts was applied to calibrated Sephadex G50 superfine columns (10 × 1,000 mm), eluted at 4°C with 0.125 M NH₄HCO₃ at a flow rate of 3.6 ml per h. Fractions of 1.2 ml were collected and analysed by specific radioimmunoassays.

ucts were subsequently cloned into the *EcoRI* and *BamHI* site of plasmid pGEM-3.

3.3. Structural features of cionin cDNA and precionin

One cDNA clone from the gastrointestinal tract, designated p3/Cio-21, was sequenced to completion on both strands and found to contain a sequence of 510 bp (Fig. 2). A hexanucleotide sequence, AATAAA (poly(A) addition signal), was found 15 nucleotides upstream from the poly(A)⁺ tail. Northern blot analysis confirmed the expression of a ~500 nucleotide cionin mRNA in the gastrointestinal tract (Fig. 3). Moreover, IPCR cloning confirmed the established DNA sequences from clones pCIII_M and pCIII_G as obtained by the IPPC technique (see section 2). PCR amplification of neuronal ganglion cDNA, using gene-specific prim-

ers #4 and #9 in a first round, and #4 and #10 (Table I) in a second round PCR amplification, respectively, revealed that cionin mRNA is expressed in intestinal, as well as in ganglionic, tissue.

The deduced precionin has 128 amino acids residues and the 484 bp coding sequence begins with AGCATGG, which is in agreement with the consensus sequence for translation initiation sites in eukaryotic cells [17]. The signal peptide of 20 amino acid residues is similar in size to those of preproCCK and preprogastrin. The structure of precionin does not resemble that of the procaeruleins, having only one copy of cionin per propeptide. Instead, the overall sequence of precionin is also similar to that of proCCK in terms of C-terminal position of cionin in the propeptide between a single basic residue, Arg, and a classic amidation site, Gly-Lys-Arg.

Cionin	Met	Asp	Arg	Asn	Tyr	Tyr	Gly	Trp	Met	Asp	Phe	Gly	Lys	Arg
CCK	Ser	-	-	Asp	-	Met	-	-	-	-	-	-	Arg	-
Gastrin	Glu	Glu	Glu	Glu	Ala	-	-	-	-	-	-	-	Arg	-

Cionin		Ala	Ile	Glu	Asp	Val	Asp	Tyr	Glu	Tyr		
CCK	Ser	-	-	-	-	-	-	-	-	Pro	Ser	
Gastrin	-	-	-	-	-	-	Glu	Glu	Asp	Gln	Tyr	Asn

Fig. 5. Comparison of the C-terminal parts of precionin, rat proCCK, and rat progastrin. Gaps (blank spaces) are introduced to maximize homology between the sequences. Dashes indicate identity between the sequences. Gly is the amide donor for the mature peptides.

3.4. Expression and processing of procionin

In view of the similarity of procionin with proCCK as well as the dual occurrence of cionin mRNA in the neural ganglion and intestinal tract, a corresponding dual occurrence of procionin and its products was also examined. Extracts of the neural ganglion contained only the cionin octapeptide amide without traces of precursors or processing intermediates. In contrast, the gastrointestinal extracts also contained significant concentrations of glycine-extended cionin (Fig. 4). Moreover, antibody G160, which is specific for tyrosyl-sulfated CCK-like peptides (such as the sulfated neuronal cionin), did not recognize the gastrointestinal cionin peptide (Fig. 4). Thus, in gastrointestinal cionin, the tyrosine in position seven is not sulfated. Hence, the neuronal processing of procionin is more complete than that of gastrointestinal cells.

In addition to identification of the primary structure of preprocionin the present study has answered three questions raised by the original identification of cionin. First, procionin resembles mammalian proCCK more than progastrin. Accordingly, procionin displays only slight similarity with the procaeruleins [4]. Second, the pentapeptide sequence of procionin immediately N-terminal to the two O-sulfated tyrosyl residues in the cionin sequence contains two Arg and only one Asp (Fig. 2). Hence, the procionin sequence challenges the proposed consensus rules for tyrosyl O-sulfation [5], and suggests a different substrate specificity of protochordate tyrosyl-protein sulfotransferase. Third, cionin mRNA and peptides are present also in the *Ciona* gut, although intestinal procionin is less processed than the neuronal procionin (Fig. 4). Thus, also with respect to tissue-specific expression and processing, the cionin system resembles the mammalian CCK system.

Cionin was originally identified in extracts from the neuronal ganglion [3]. The finding of a gastrin/CCK-like peptide in this ganglion is in accordance with earlier immunochemical studies [14,18,19]. Interestingly, Thorndyke and Dockray [19], using antibodies raised against mammalian gastrin-17, could not confirm an earlier observation of gastrin immunoreactivity in gut cells [20]. Since cionin reacts with C-terminal-directed gastrin antibodies, the intestinal immunoreactivity is presumably due to cionin. Our study supports such a contention, as PCR amplification of a cionin mRNA isolated from the gut confirmed the expression of cionin mRNA in the gut. Moreover, cionin peptides were also identified in gut tissue (Fig. 4).

The procionin structure shows other remarkable features. Procionin contains mono- and dibasic processing sites similar to those of proCCK and progastrin ([1,21,22], and Fig. 2). Moreover, alignment of the C-terminal sequences of procionin, proCCK, and progastrin shows that procionin is more closely related to proCCK than to progastrin (Fig. 5). In known progas-

trin and proCCKs, the C-terminal amidation site, Gly-Arg-Arg, is followed by the tripeptide Ser-Ala-Glu. It has been proposed that phosphorylation of the Ser adjacent to the amidation site is of regulatory significance [23]. However, the procionin amidation site Gly-Lys-Arg is followed by the tripeptide Ala-Ile-Glu rather than Ser-Ala-Glu (Fig. 5, lower panel). This indicates that processing at the procionin amidation site does not require phosphorylation.

Acknowledgements The skilful technical assistance of Mette Simons, Peter Nørhave, Alice Lieth, the secretarial help of Charlotte Vienberg, and critical reading of the manuscript by Dr Jan Zabielski are gratefully acknowledged. For help in collecting *Ciona intestinalis* we thank Dr. Ib Svane and the Marine Biology Station in Kristineberg, Sweden. This study was supported by the Danish MRC, the Danish Cancer Society, the Lundbeck Foundation (to R.F.) and the Danish governmental educational support (to J.U.T.). The sequence reported in this paper has been deposited in the GenBank data base, accession no. X69130.

REFERENCES

- [1] Rehfeld, J.F. (1988) *J. Mol. Endocrinol.* 1, 87-94.
- [2] Larsson, L.I. and Rehfeld, J.F. (1977) *Nature* 269, 335-338.
- [3] Johnsen, A.H. and Rehfeld, J.F. (1990) *J. Biol. Chem.* 265, 3054-3058.
- [4] Richter, K., Egger, R. and Kreil, G. (1986) *J. Biol. Chem.* 261, 3676-3680.
- [5] Huttner, W. and Baeuerle, P.A. (1988) *Modern Cell Biol.* 6, 97-140.
- [6] Schjoldager, B., Park, J., Johnsen, A.H., Yamada, T. and Rehfeld, J.F. (1991) *Am. J. Physiol.* 260, G977-G982.
- [7] Frohman, M.A., Dush, M.K. and Martun, G.R. (1988) *Proc. Natl. Acad. Sci. USA* 85, 8998-9002.
- [8] Monstein, H.-J., Geijer, T. and Bakalkin, G.Y. (1992) *BioTechniques* 13, 842-844.
- [9] Huang, S.H., Hu, Y., Wu, Sh. and Holcenberg, J. (1990) *Nucleic Acids. Res.* 18, 1922.
- [10] Chomczynski, T. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156-159.
- [11] Monstein, H.-J. (1992) *Methods Mol. Cell. Biol.* 3, 204-205.
- [12] Hilsted, L. and Rehfeld, J.F. (1986) *Anal. Biochem.* 152, 119-126.
- [13] Rehfeld, J.F. (1978) *J. Biol. Chem.* 253, 4016-4021.
- [14] Conlon, J.M., Schwartz, T.W. and Rehfeld, J.F. (1988) *Regul. Pept.* 20, 241-250.
- [15] Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y. and Matsubara, K. (1985) *J. Biol. Chem.* 260, 2605-2608.
- [16] Ohara, O., Dorit, R.L. and Gilbert, W. (1989) *Proc. Natl. Acad. Sci. USA* 86, 5673-5677.
- [17] Kozak, M. (1986) *Cell* 44, 283-292.
- [18] Thorndyke, M.C. (1982) *Regul. Pept.* 3, 281-288.
- [19] Thorndyke, M.C. and Dockray, G.J. (1986) *Regul. Pept.* 16, 269-279.
- [20] Fritsch, H.A.R., Noorden, S.V. and Pearse, A.G.E. (1978) *Cell Tissue Res.* 186, 181-185.
- [21] Gubler, U., Chua, A., Hoffman, B.J., Collier, K.J. and Eng, J. (1984) *Proc. Natl. Acad. Sci. USA* 81, 4307-4310.
- [22] Rehfeld, J.F., Hilsted, L. and Johnsen, A.H. (1991) in: *The Stomach as an Endocrine Organ* (Håkanson, R. and Sundler, F. eds.) Fernström Foundation series vol. 15, pp. 179-195.
- [23] Dockray, G.J., Varro, A., Watkinson, A. and Dimaline, R. (1991) in: *The Stomach as an Endocrine Organ* (Håkanson, R. and Sundler, F. eds.) Fernström Foundation Series, vol. 15, pp. 197-210.