cDNA sequence of human β-preprotachykinin, the common precursor to substance P and neurokinin A

A.J. Harmar, A. Armstrong, J.C. Pascall, K. Chapman, R. Rosie, A. Curtis, J. Going*, C.R.W. Edwards† and G. Fink

*MRC Brain Metabolism Unit, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF, Scotland, *University Department of Pathology, Medical Buildings, Teviot Place, Edinburgh EH8 9AG, Scotland* and †*University Department of Medicine, Western General Hospital, Edinburgh EH4 2XU, Scotland*

Received 14 August 1986

The nucleotide sequence of cDNA encoding the human substance P precursor, β-preprotachykinin (β-PPT), has been determined. The source of mRNA was a human laryngeal carcinoid tumour that contained a high concentration of immunoreactive substance P. The human β-PPT polypeptide is 129 amino acids long and contains regions encoding substance P and neurokinin A, each flanked by basic amino acid residues. Residues 72–107 of the human β-PPT polypeptide encode the sequence of neuropeptide K, an N-terminally extended form of neurokinin A recently isolated from porcine brain.

Substance P *Substance K* *Neuropeptide K* *Tachykinin* *mRNA* *Carcinoid tumour*

1. INTRODUCTION

Substance P was, until recently, the only known representative in mammals of the tachykinins, a family of neuropeptides characterised by the common C-terminal amino acid sequence -Phe-X-Gly-Leu-Met-NH_2, where X is a hydrophobic or aromatic residue [1]. Three further mammalian tachykinins have now been described: NkA [2], neurokinin B and NpK [3], an N-terminally extended form of NkA. In bovine brain, SP is derived from two closely related precursor polypeptides (α- and β-PPT [4]). In addition to a region encoding the sequence of SP, β-PPT contains the sequences of NkA and NpK. SP and NkA are thought to exert their actions through separate classes of tachykinin receptors ('SP-P' and 'SP-K' receptors, respectively), which have different distributions in the nervous system [5].

Abbreviations: cDNA, complementary DNA; mRNA, messenger RNA; NkA, neurokinin A (formerly known as substance K, neuromedin L or neurokinin α); NpK, neuropeptide K; PPT, preprotachykinin; SP, substance P

We report here the nucleotide sequence of cDNA encoding human β-PPT. Human carcinoid tumours have been shown to contain and secrete large amounts of tachykinin-immunoreactive peptides, including SP, NkA [6] and a peptide with chromatographic and immunochemical properties similar to those of NpK [7]. Accordingly, for the cDNA cloning, we have used as the source of mRNA a metastatic deposit from a laryngeal carcinoid tumour which contained abundant (~1 μg/g wet wt) SP-like immunoreactivity. The tumour caused hypersalivation and local pain in the patient, possibly as a consequence of ectopic SP production [8]. Human β-PPT cDNA shows extensive homology to bovine β-PPT cDNA [2] and includes regions encoding the sequences of SP, NkA and NpK.

2. EXPERIMENTAL

2.1. *cDNA synthesis and construction of recombinant plasmids*

Poly(A)$^+$ RNA was isolated from 3.5 g tumour tissue using the guanidine thiocyanate procedure

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

[9] and chromatography on oligo(dT) cellulose [10]. Double-stranded cDNA was synthesised using either oligo(dT)$_{12-18}$ or the synthetic heptadecamer probe I as primer (fig.1: supplied by Celltech). The procedures of Craig et al. [11] were used for the first strand and those of Gubler and Hoffmann [12] for the second strand. Double-stranded cDNA was tailed with approx. 25 dC residues using terminal deoxynucleotidyl transferase as described by the suppliers (Bethesda Research Laboratories).

Poly(dC)-tailed cDNA and poly(dG)-tailed, *Pst*1-restricted pBR322 DNA (New England Nuclear) were annealed [13] at final concentrations of 8 and 400 ng/ml, respectively. The resulting chimeric plasmids were then used to transform *E. coli* RR1 [14]. Aliquots of the transformed cells were plated onto L-agar containing 10 μg/ml tetracycline and incubated overnight at 37°C. Tetracycline-resistant colonies were then picked in triplicate onto Biodyne filters overlaying agar and allowed to grow overnight at 37°C prior to storage or colony filter hybridisation.

2.2. *Colony screening*

To identify clones containing PPT sequences, filters were grown overnight on L-agar plates containing tetracycline (10 μg/ml) and chloramphenicol (250 μg/ml) prior to screening with synthetic 17-mer probes (fig.1). Bacteria were lysed and the DNA was fixed onto the filters as described by Grunstein and Hogness [15]. Filters were prehybridised for 18 h at 37°C in a solution containing 0.9 M NaCl, 60 mM EDTA, 19.8 mM Tris-HCl (pH 8) and 0.1% each of bovine serum albumin, polyvinyl pyrrolidone and Ficoll (5 × Denhardt's reagent) then hybridised in an identical solution containing hybridisation probe, labelled at the 5′-end with [γ-^{32}P]ATP [16]. Filters were hybridised for 18 h at 37°C and were then washed five times (30 min per wash) with 0.9 M NaCl, 0.09 M sodium citrate, pH 7 (6 × SSC), at room temperature and twice with 6 × SSC/0.1% SDS at 35°C. Filters were dried and exposed to Kodak X-Omat S film for 24–72 h using Kodak X-Omatic intensifying screens.

2.3. *Nucleotide sequence determination*

The sequences of portions of 2 clones containing β-PPT cDNA (phE293 and ph8.216) were determined by the method of Sanger et al. [17] after subcloning of suitable restriction fragments into bacteriophages M13 mp18 or M13 mp19.

2.4. *Northern blot analysis of human β -PPT mRNA*

Polyadenylated RNA (1 μg) isolated from the tumour was fractionated by electrophoresis on a formaldehyde-containing 1% (w/v) agarose gel and the RNA blotted onto a nylon membrane [18]. The membrane was hybridised with ^{32}P-labelled, nick-translated plasmid phE293 and washed as described [18].

3. RESULTS

3.1. *Identification and sequence of β-PPT cDNA*

A cDNA library of 8700 transformants was generated from total poly(A)$^+$ RNA using cDNA synthesised with oligo(dT)$_{12-18}$ as the primer. Identification of one cDNA clone (phE293) encoding PPT was achieved by screening duplicate filters of the library with ^{32}P-labelled probe I, which is complementary to both the SP and NkA coding regions of bovine β-PPT. Sequencing of the cDNA insert in phE293 revealed that it consisted of 904 bp. The insert included an open reading frame which encoded a PPT, extending for 321 bp from the 5′-end, followed by 553 bp of 3′-untranslated region and 30 nucleotides of poly(A)$^+$ tail.

The 5′-end of the insert in phE293 interrupted the open reading frame coding for PPT, indicating that reverse transcription had not proceeded to the 5′-end of the mRNA and that phE293 did not contain the entire β-PPT cDNA. In addition, Northern blot analysis demonstrated that the smallest and most abundant PPT mRNA in human tissue consisted of approx. 1200 nucleotides, some 300 nucleotides larger than the insert in phE293 (fig.2). Therefore an additional cDNA library was constructed from the tumour poly(A)$^+$ RNA to obtain the region corresponding to the 5′-end of the PPT mRNA. To enrich for the desired clones, cDNA was synthesised using probe I as a PPT-specific primer. A second 17-mer (probe II, fig.1; supplied by Dr S. Minter, UMIST) complementary to a region close to the 5′-end of the insert in phE293 was used to screen 2700 recombinants; one clone (ph8.216) gave a strong hybridisation signal. The

5′-TTTGCCCAT$^{C}_{T}$A$^{A}_{G}$TCCAA-3′ PROBE I

5′-GTCGTACCAGTCGGACC-3′ PROBE II

Fig.1. Sequences of oligonucleotides used for screening of human cDNA libraries and as primers for cDNA synthesis. Probe I is complementary to parts of the SP- and NkA-coding regions of bovine β-PPT cDNA (nucleotides 337–353 and 454–470 of the published sequence [4]). Probe II is complementary to a region of human β-PPT cDNA close to the 5′-end of the insert in phE293 (nucleotides 173–189 in fig.3).

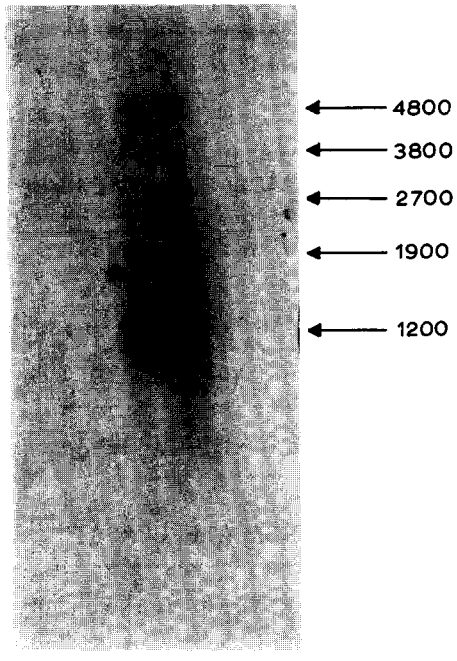


Fig.2. Northern blot analysis of 1 μg of polyadenylated RNA isolated from tumour tissue. The sizes (in base pairs) of the principal hybridising RNA species, determined by comparison with restriction endonuclease *Pst*I-cut bacteriophage λ DNA fragments electrophoresed in parallel, are indicated.

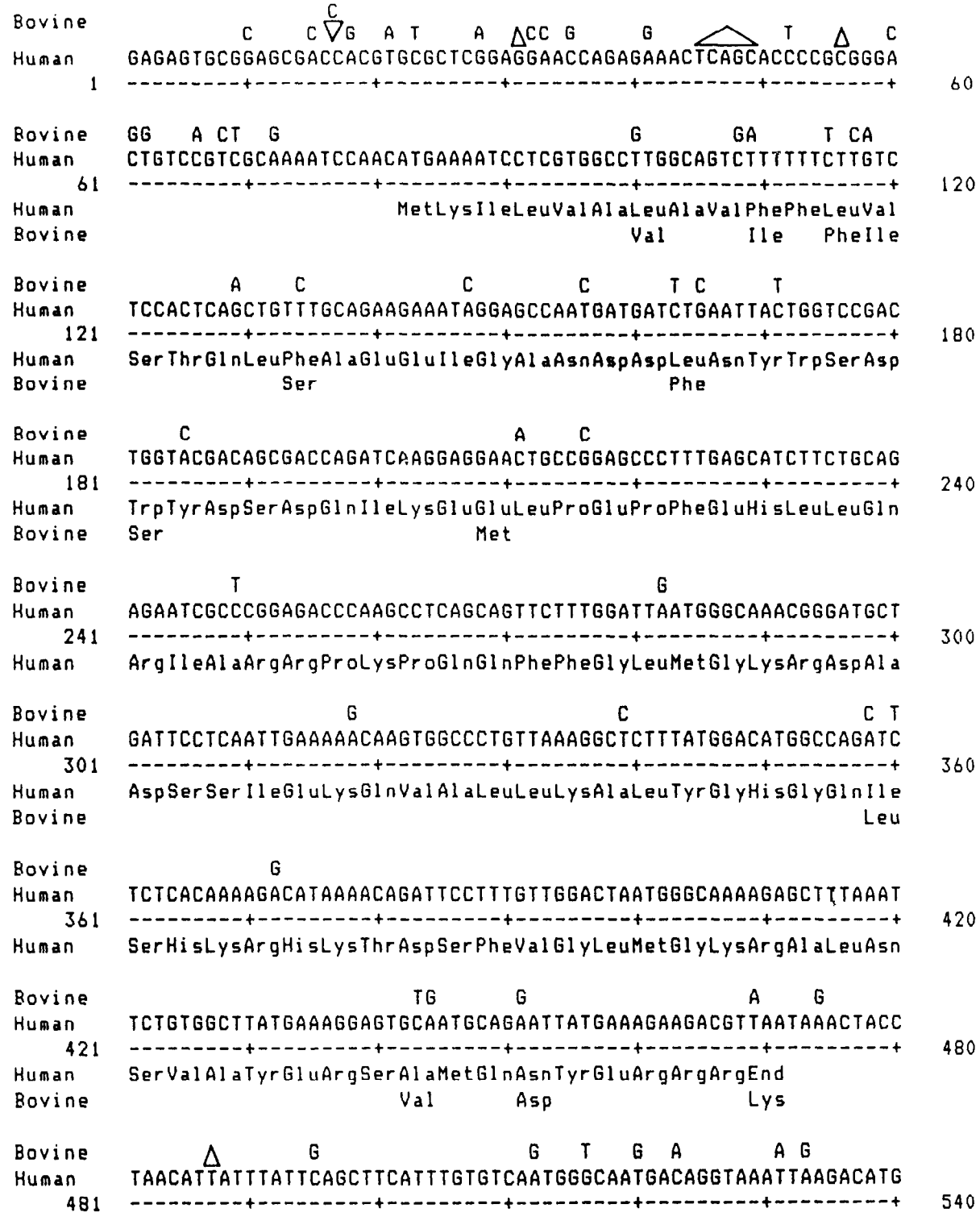
insert contained in ph8.216 was sequenced: it consisted of 260 bp including 113 bp of the PPT sequence that had been determined in phE293, together with a further 147 bp of 5′-sequence.

The sequence of 1021 nucleotides of human PPT cDNA, determined from clones phE293 and ph8.216, is shown in fig.3, together with the predicted amino acid sequence of the translated product. The human β-PPT polypeptide consists of 129 amino acids (M_r 14820).

4. DISCUSSION

We have established the sequence of a human β-PPT cDNA cloned from carcinoid tumour tissue. The cDNA consists of 81 nucleotides of 5′-flanking sequence, the complete coding region (387 bp) and 533 nucleotides of 3′-untranslated sequence. As we have determined the sequence of a total of 1051 nucleotides of PPT cDNA, and Northern blot analysis indicated that the mature PPT mRNA is 1200 nucleotides long, we consider that we have cloned cDNA corresponding to at least 85% of the complete β-PPT mRNA. As well as the mature PPT mRNA at 1200 nucleotides, several larger RNA species were detected by Northern blot analysis. In bovine brain, SP is derived from two closely related PPTs, α and β [4] which are generated from a single PPT gene as a result of tissue-specific RNA splicing [19]. Southern blot analysis (J.P., unpublished) indicated that human PPTs are also encoded by a single gene. These larger mRNA species may therefore correspond to intermediates in the RNA processing pathway of β-PPT mRNA. The human PPT cDNA sequence described here is closely homologous to the β-form of bovine PPT cDNA [4]. Amino acid sequences corresponding to SP and NkA (nucleotides 265–297 and 385–414, respectively) are flanked by sequences which are sites of post-translational processing. At the N-terminus, the SP sequence is flanked by an arginine residue and the NkA sequence by a pair of basic amino acids (Lys-Arg). At the C-terminus, both peptides are flanked by a sequence (-Gly-Lys-Arg) from which the C-terminal amide groups of SP and NkA are generated [20].

Using the techniques of Edman sequencing, amino acid analysis and fast atom bombardment mass spectrometry, Roth et al. [6] have demonstrated the presence of both SP and NkA in human carcinoid tumour tissue. Our results indicate that in man, SP and NkA are synthesised from a single polypeptide precursor highly homologous to that found in bovine tissues. Nucleotides 295–402 of human β-PPT cDNA encode a peptide with a sequence identical to that of porcine NpK [3], which differs at a single position ($Leu^{22} \longrightarrow Ile$) from the corresponding region of bovine β-PPT. These studies establish that human NpK has an identical sequence to the porcine peptide [3] as


Bovine
Human
GAGAGTGCGGAGCGACCACGTGCGCTCGGAGGAACCAGAGAAACTCAGCACCCCGCGGGA
1
60
CTGTCCGTCGCAAAATCCAACATGAAAATCCTCGTGGCCTTGGCAGTCTTTTTTCTTGTC
61
120
MetLysIleLeuValAlaLeuAlaValPhePheLeuVal
Val Ile PheIle
TCCACTCAGCTGTTTGCAGAAGAAATAGGAGCCAATGATGATCTGAATTACTGGTCCGAC
121
180
SerThrGlnLeuPheAlaGluGluIleGlyAlaAsnAspAspLeuAsnTyrTrpSerAsp
Ser Phe
TGGTACGACAGCGACCAGATCAAGGAGGAACTGCCGGAGCCCTTTGAGCATCTTCTGCAG
181
240
TrpTyrAspSerAspGlnIleLysGluGluLeuProGluProPheGluHisLeuLeuGln
Ser Met
AGAATCGCCCGGAGACCCAAGCCTCAGCAGTTCTTTGGATTAATGGGCAAACGGGATGCT
241
300
ArgIleAlaArgArgProLysProGlnGlnPhePheGlyLeuMetGlyLysArgAspAla
GATTCCTCAATTGAAAAACAAGTGGCCCTGTTAAAGGCTCTTTATGGACATGGCCAGATC
301
360
AspSerSerIleGluLysGlnValAlaLeuLeuLysAlaLeuTyrGlyHisGlyGlnIle
Leu
TCTCACAAAAGACATAAAACAGATTCCTTTGTTGGACTAATGGGCAAAAGAGCTTTAAAT
361
420
SerHisLysArgHisLysThrAspSerPheValGlyLeuMetGlyLysArgAlaLeuAsn
TCTGTGGCTTATGAAAGGAGTGCAATGCAGAATTATGAAAGAAGACGTTAATAAACTACC
421
480
SerValAlaTyrGluArgSerAlaMetGlnAsnTyrGluArgArgArgEnd
Val Asp Lys
TAACATTATTTATTCAGCTTCATTTGTGTCAATGGGCAATGACAGGTAAATTAAGACATG
481
540

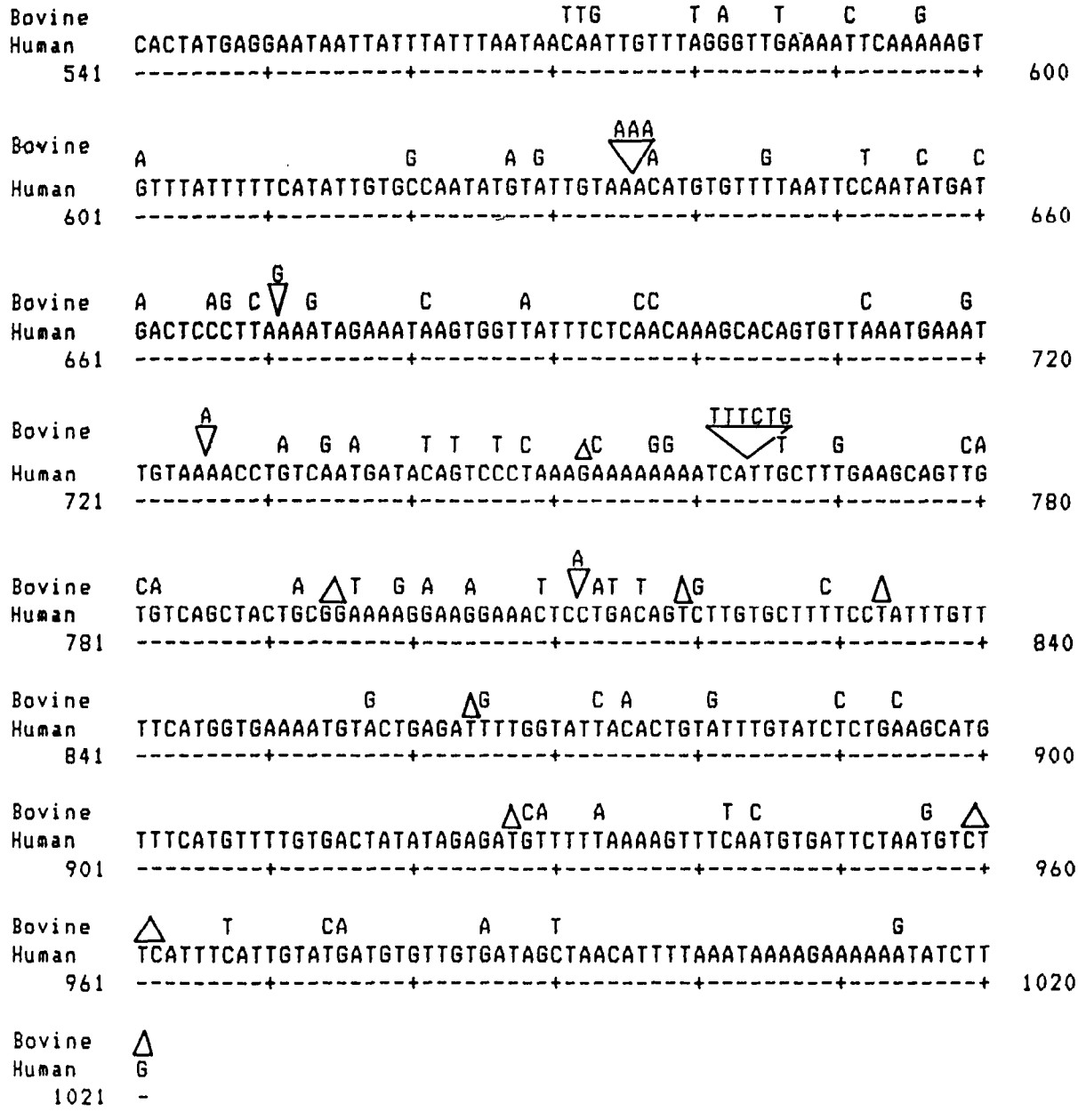


Fig.3. The sequence of cDNA for human β-PPT and the deduced sequence of the precursor polypeptide. Numbering of nucleotides begins with the first nucleotide of cDNA sequence. The sequence of nucleotides 1–260 was determined from clone ph8.216 and nucleotides 148–1021 were sequenced from clone phE293. A sequence of 30 poly(A) residues at the end of the sequence of phE293 (immediately 3′ to nucleotide 1021) is not shown. Nucleotides which differ in bovine β-PPT [4] are shown above the main nucleotide sequence: homology between the sequences was maximised by inserting gaps [22] indicated by the triangular symbols. Amino acid differences between bovine and human β-PPT are shown below the main polypeptide sequence.

suggested by the chromatographic studies of Theodorsson-Norheim et al. [7].

The protein-coding region of the sequence of human β-PPT mRNA is very similar to that of the corresponding region of bovine β-PPT: the translation products predicted from the two sequences differ by only 13/130 amino acids (90% homology) and there is over 92% homology at the nucleotide level. Human and bovine β-PPTs also exhibit significant homology in the 5′- and 3′-untranslated regions (68 and 81%, respectively). The close similarity of the human and bovine β-PPT polypeptides suggests that these precursors may give rise not only to SP, NpK and NkA, but also to other biologically active peptides. The most likely site for the cleavage of the signal sequence from human β-PPT is between Ala^{19} and Glu^{20} [21], which would liberate a prohormone of 110 amino acids (M_r 12868). Post-translational processing of this prohormone could generate two novel peptides which may possess biological activity: (i) a peptide of approx. 38 amino acids extending from the signal peptidase cleavage site to the arginine residue immediately N-terminal to the SP sequence and (ii) a 19 amino acid peptide, with three basic amino acids (Arg-Arg-Arg) at the C-terminus, corresponding to nucleotides 412–468 of the human β-PPT cDNA sequence.

Some of the hitherto unexplained clinical features of the carcinoid syndrome may be due to the secretion of novel peptides derived from β-PPT. The isolation of cDNA clones encoding human β-PPT will permit the characterisation of these peptides and facilitate studies of the regulation of PPT gene expression in normal and ectopic sites in man.

ACKNOWLEDGEMENTS

We thank Norma Brearley for the careful preparation of the manuscript, Helen Stanley and Adrian Pierotti for constructive criticism and Dr A.F.W. Coulson for help and advice in the use of computer methods for DNA sequence analysis.

REFERENCES

[1] Erspamer, V. (1981) Trends Neurosci. 4, 267–269.
[2] Kimura, S., Oada, M., Sugita, Y., Kanazawa, I. and Munekata, E. (1983) Proc. Jap. Acad. Ser. B 59, 101–104.
[3] Tatemoto, K., Lundberg, J.M., Jornvall, H. and Mutt, V. (1985) Biochem. Biophys. Res. Commun. 128, 947–953.
[4] Nawa, H., Hirose, T., Takashima, H., Inayama, S. and Nakanishi, S. (1983) Nature 306, 32–36.
[5] Buck, S.H., Burcher, E., Shults, C.W., Lovenberg, W. and O'Donohue, T.L. (1984) Science 226, 987–989.
[6] Roth, K.A., Makk, G., Beck, O., Faull, K., Tatemoto, K., Evans, C.J. and Barchas, J.D. (1985) Regul. Pept. 12, 185–199.
[7] Theodorsson-Norheim, E., Norheim, I., Oberg, K., Brodin, E., Lundberg, J.M., Tatemoto, K. and Lindgren, P.G. (1985) Biochem. Biophys. Res. Commun. 131, 77–83.
[8] Going, J.J., Harmar, A.J., Gow, I.F. and Edwards, C.R.W. (1985) J. Endocrinol. 104, 52.
[9] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 18, 5294–5299.
[10] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.
[11] Craig, R.K., Hall, L., Parker, D. and Campbell, P.N. (1981) Biochem. J. 194, 989–998.
[12] Gubler, V. and Hoffmann, B.J. (1983) Gene 25, 263–269.
[13] Peacock, S.L., McIver, C.M. and Monahan, J.J. (1981) Biochim. Biophys. Acta 655, 243–250.
[14] Dagert, M. and Ehrlich, S.D. (1979) Gene 6, 23–28.
[15] Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 872, 3961–3965.
[16] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[17] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[18] Taylor, J.B., Craig, R.K., Beale, D. and Ketterer, B. (1984) Biochem. J. 219, 223–231.
[19] Nawa, H., Kotani, H. and Nakanishi, S. (1984) Nature 312, 729–734.
[20] Eipper, B.A., Mains, R.E. and Glembotski, C.C. (1983) Proc. Natl. Acad. Sci. USA 80, 5144–5148.
[21] Von Heijne, G. (1983) Eur. J. Biochem. 133, 17–21.
[22] Needleman, S.B. and Wunsch, C.D. (1970) J. Mol. Biol. 58, 443–453.