The EMBO Journal vol.3 no.4 pp.909–912, 1984

A cDNA encoding a small common precursor for human pancreatic polypeptide and pancreatic icosapeptide

E. Boel*, T.W. Schwartz[1], K.E. Norris and N.P. Fiil

Laboratory of Genetics, NOVO Research Institute, DK-2880 Bagsvaerd, and [1]Laboratory of Molecular Endocrinology, University Department of Clinical Chemistry, Rigshospitalet 6321, DK-2100 Copenhagen, Denmark

*To whom reprint requests should be sent

Communicated by N.P. Fiil

A cDNA for the hormone, human pancreatic polypeptide (PP), was isolated by oligodeoxynucleotide screening from a cDNA library constructed from normal human pancreatic mRNA. The primary structure of the precursor protein as deduced from the cDNA sequence is 95 amino acids long and is composed of a typical, but rather long signal peptide of 29 residues, followed by the sequence of the 36 amino acid human pancreatic polypeptide, which again is separated from the human pancreatic icosapeptide sequence by a classic cleavage and amidation site, Gly-Lys-Arg. The precursor terminates in a heptapeptide which is cleaved from the icosapeptide at a monobasic processing site. Both the size and the structure of the PP precursor was supported by the results of peptide analysis of biosynthetically labeled pro-PP isolated from canine PP cells in which processing was prevented by the arginine analogue canavanine. It is concluded that the precursor for mammalian PP gives rise to two peptide products, the well preserved, carboxyamidated PP and an icosapeptide which is preserved only in its COOH-terminal end, plus a small highly variable COOH-terminal oligopeptide.

Key words: amino acid analogues/DNA sequence/molecular cloning/oligodeoxynucleotide screening/peptide biosynthesis

Introduction

Pancreatic polypeptide (PP) was the first identified product of the dominating endocrine cell type of the duodenal pancreas (Chance *et al.*, 1975; Kimmel *et al.*, 1975; Larsson *et al.*, 1976). This 36 amino acid, carboxyamidated peptide is one of a family of biologically active peptides which at present also includes peptide YY and neuropeptide Y (Tatemoto, 1982a, 1982b). PP itself seems to be a hormone involved in the regulation of the exocrine pancreatic secretion, and in the regulation of biliary tract motility (Schwartz, 1983). PP might also be of importance in the regulation of food intake as genetically obese laboratory animals have altered PP release (Malaisse-Lagae *et al.*, 1977; Gates and Lazarus, 1977; Larsson *et al.*, 1977); and, at least in New Zealand obese mice, the weight gain can be cured by infusions of PP (Gates and Lazarus, 1977). In humans the most clear cut analogy is the Prader-Willi syndrome; obese children with this disease also demonstrate a blunted PP secretion (Zipf *et al.*, 1981).

With primary cultures of endocrine cells from the canine duodenal pancreas it was shown that PP is synthesized as the NH_2-terminal part of a relatively small precursor (Schwartz *et al.*, 1980). From the COOH-terminal region of this precursor a newly identified icosapeptide is derived initially as a COOH-terminally extended form (Schwartz and Tager, 1981). The evidence which indicated that the pancreatic icosapeptide and PP are synthesised in a common precursor stems solely from peptide chemistry and includes pulse-chase experiments, immunoprecipitation, peptide mapping, microsequencing, and immunocytochemistry (Schwartz *et al.*, 1980, 1984; Schwartz and Tager, 1981). Here we describe the cloning of a cDNA complementary to human PP mRNA which also codes for the pancreatic icosapeptide. Furthermore, we present evidence that the biosynthetically labeled precursor, which accumulates in the canine cells when processing is prevented by incorporation of a basic amino acid analogue, has a structure similar to the one deduced from the sequence of the human PP cDNA.

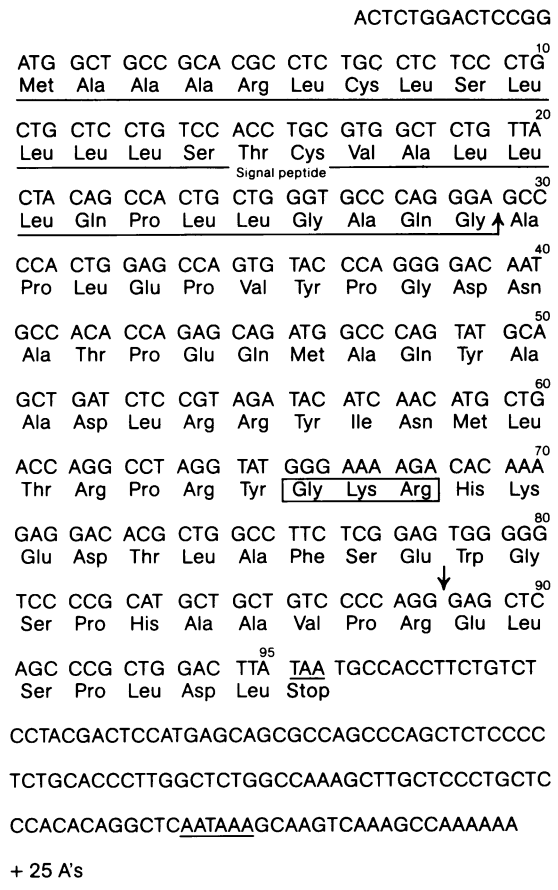

Fig. 1. Sequence of the human PP cDNA insert in clone pHPP165. The stop codon and the polyadenylation site are underlined. The amino acid sequence deduced from the cDNA sequence of the common precursor for human PP and pancreatic icosapeptide is shown. Amino acids are numbered from the initiating methionine. In the protein sequence the putative signal peptide is underlined, and the combined cleavage and amidation site between the PP (30–65) and the icosapeptide sequence (69–88) has been put in a box. The arrow indicates the cleavage site between the icosapeptide and the terminating heptapeptide (89–95). The deduced protein sequence confirms the primary structure of human PP (Chance *et al.*, 1975) and that of human pancreatic icosapeptide (Schwartz *et al.*, 1984) as determined by peptide chemistry.

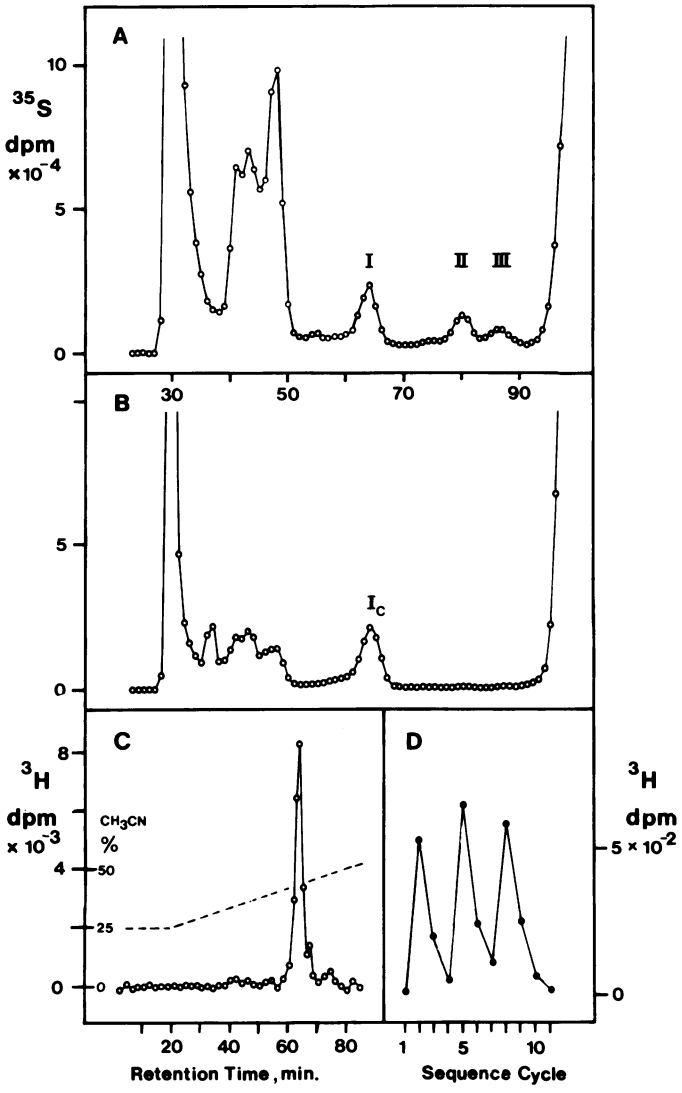

Fig. 2. Characterization of biosynthetically labeled pro-PP from canine PP cells. **(A)** Gel filtration profile of biosynthetically labeled peptides extracted from a primary culture of PP cells pre-incubated for 6 h followed by an incubation for 3 h. The column of BIO-GEL P-30 (1.6 x 95 cm) was eluted with 0.5 M acetic acid. It has previously been shown that peptide I is a common precursor for peptide II, PP, and peptide III, pancreatic icosapeptide. **(B)** A similar experiment to that shown in **A**, except that these cells were exposed to a medium in which arginine was replaced by the analogue canavanine in order to prevent processing and thereby to obtain pro-PP, the primary translation product minus the signal peptide. Note that there is no radioactivity eluting as PP or icosapeptide, however, there is a peptide co-eluting with peptide I, peptide I_c. **(C)** Biosynthetically labeled peptide I_c characterized on h.p.l.c. using a VYDAC C_{18} column and elution with a gradient of acetonitrile and trifluoroacetic acid, 0.1%, in water. **(D)** Release of radioactivity during automated sequential Edman degradation of the biosynthetically labeled and purified peptide I_c. In this case the cells had been incubated with [^{3}H]proline.

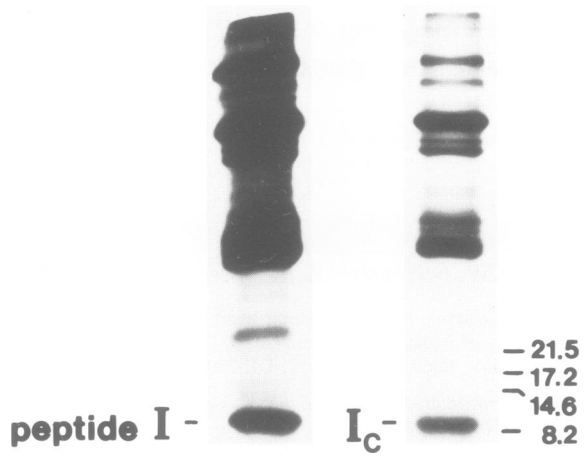

Fig. 3. Autoradiograph from SDS-gel electrophoresis of biosynthetically labeled peptides extracted from canine pancreatic polypeptide cells incubated with [^{35}S]methionine for 3 h. The left lane shows peptides extracted from cells incubated in the presence of all essential amino acids and the right lane shows peptides from cells incubated in the presence of the arginine analogue, canavanine. The position taken by peptide I and peptide I_c (from Figure 2) is indicated and the position taken by calibration peptides of mol. wt. 21 500 (21.5), 17 200 (17.2), 14 600 (14.6), and 8300 (8.3) is indicated at the right.

Results

Cloning and identification of human PP cDNA

Double-stranded cDNA made from total human pancreas poly(A)$^+$ RNA was inserted into the *Pst*I site of the cloning vector pBR327 and cloned in *Escherichia coli*. About 1000 colonies containing recombinant plasmids were immobilized on filter paper and hybridized with a 5′ end-labeled heptadecanucleotide mixture specific for PP as described in Materials and methods. The sequence of one of the 32 oligonucleotides in the mixture is complementary to the human PP mRNA in the region coding for Glu-Gln-Met-Ala-Gln-Tyr. One clone, pHPP165, hybridized to the probe mixture and was found to carry an insert of ~530 bp. The nucleotide sequence of the insert of pHPP165 was determined from a *Hind*III site as well as from four internal *Bst*NI sites in the recombinant. The resulting nucleotide sequence (Figure 1) was in total agreement with the known amino acid sequence of human PP (Chance *et al.*, 1975) and the COOH-terminal icosapeptide (Schwartz *et al.*, 1984).

Structure of human prepro-PP mRNA

The insert of pHPP165 (Figure 1) contains 14 nucleotides as a part of a 5′-untranslated region, 87 nucleotides corresponding to a typical hydrophobic signal peptide of 29 amino acid residues, 108 nucleotides coding for human PP, nine nucleotides coding for Gly-Lys-Arg, 60 nucleotides coding for the human icosapeptide, and 21 nucleotides coding for a COOH-terminal heptapeptide before a 3′-untranslated region of 126 nucleotides. This 3′ end had a typical polyadenylation signal, AATAAA, located 13 nucleotides before the first A in the poly(A) chain. pHPP165 contained 31 A residues from this chain and had ~30 residues in each poly[d(G-C)] tail. Human pancreatic polypeptide is therefore synthesized as a 95 amino acid long precursor of mol. wt. 10 433. The PP, 36 residues long, is followed by a glycine in position 66 known to be important in the amidation process (Bradbury *et al.*, 1982). The basic residues Lys-Arg at positions 67 and 68 constitute a classical processing site for a dibasic specific endopeptidase, that would cleave PP from the icosapeptide (Schwartz *et al.*, 1984); this is followed in the human precursor by a COOH-terminal heptapeptide before the stop-codon, TAA.

Peptide analysis of pro-PP

To prevent cleavage of pro-PP the amino acid analogue, canavanine, was substituted for arginine in the medium of primary cultures of endocrine cells from dog duodenal pancreas. Canavanine can be esterified to arginine tRNA, and is thereby readily incorporated into nascent polypeptide chains

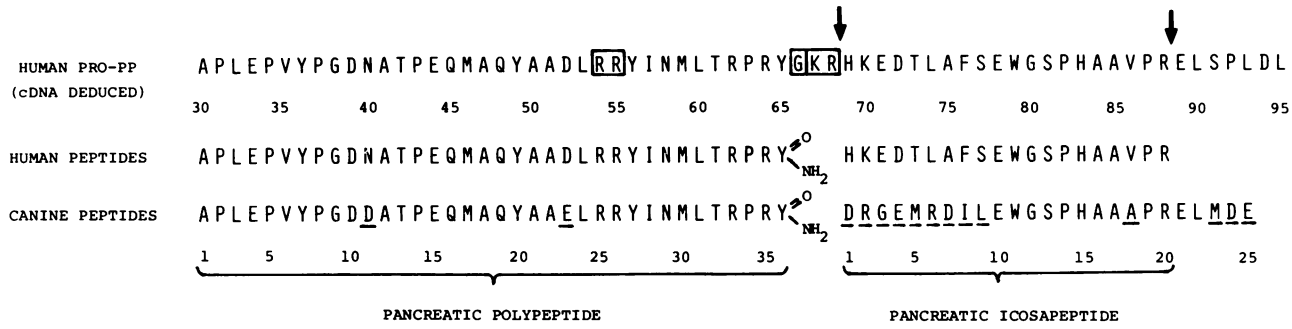

Fig. 4. Amino acid sequence of the common precursor for human PP and pancreatic icosapeptide as deduced from the PP cDNA sequence compared with the amino acid sequence of human and canine peptide products characterized through peptide chemistry. Amino acids of the pro-PP are numbered from the initiating methionine and the different peptides have, in addition, been numbered from their respective NH_2-terminal residue: the combined cleavage and amidation site and a possible dibasic cleavage site have been placed in boxes. The arrows indicate where cleavage actually takes place according to peptide chemistry information (the possible processing site $R_{54}R_{55}$ is apparently not used in the pancreas, unpublished information). The COOH-terminal extension of the biosynthetic intermediate for the canine icosapeptide is also shown (unpublished information). Amino acid residues that differ between the two species are underlined in the canine sequences.

(Attias *et al.*, 1969; Pines *et al.*, 1981). However, peptide precursors containing the analogue are not processed, e.g., insulin cells will synthesize and release only pro-insulin when exposed to canavanine (Noe, 1981; Halban, 1982). PP cells normally incorporate [^{35}S]methionine into a precursor, peptide I Figure 2A, which is cleaved into the PP segment (peptide II) and the icosapeptide segment, (peptide III) (Schwartz *et al.*, 1980). It is not known, however, whether the precursor, peptide I, is in fact identical to pro-PP, i.e., the primary translation product minus the signal peptide.

When PP cells were exposed to canavanine instead of arginine, there was, as expected, no evidence of the production of labeled peptides corresponding to PP or the icosapeptide. However, a peptide, peptide I_C, was observed, which both on gel filtration (Figure 2B) and SDS-gel electrophoresis (Figure 3) behaved as peptide I. On SDS-gel electrophoresis peptide I_C, the anticipated canine pro-PP, migrated with an apparent mol. wt. of 9500 (Figure 3) which is slightly larger than expected from a peptide with a mol. wt. of 7300 as deduced from the cDNA. However, the PP molecule itself migrates on SDS-gel electrophoresis with an apparent mol. wt. larger than the theoretical (unpublished observation). If the cells were exposed to canavanine, 4 mmol/l, in the presence of arginine, 1 mmol/l, a normal pattern of peptide labeling and processing was observed (data not shown). Peptide I_C, biosynthetically labeled with [^{3}H]proline was further purified by reverse phase h.p.l.c., where it appeared as one peak (Figure 2C). Sequence analysis, by sequential Edman degradation of peptide I_C showed that [^{3}H]proline appeared in sequence cycles 2, 5 and 8 in accordance with the proline residues in the NH_2-terminal sequence of PP (Figure 2D). These data indicate that the PP sequence is located at the far NH_2-terminal end of pro-PP as also deduced from the structure of the human PP cDNA. Thus, the peptide analysis of canine pro-PP corroborates the data deduced from the human PP cDNA both in respect of size and structure of the precursor for PP.

Discussion

Here a cDNA for human PP is identified through oligodeoxynucleotide screening of a cDNA library constructed on the basis of mRNA isolated from total human pancreas. The deduced protein structure of the precursor for PP confirms previous cell biological studies indicating that PP is synthesized together with a newly identified icosapeptide (Schwartz *et al.*, 1980; Schwartz and Tager, 1981). Furthermore, peptide analysis of biosynthetically labeled pro-PP from dog PP cells is in full agreement with the size and structure of the PP precursor inferred from the human PP cDNA.

The coding region of the human PP mRNA is 285 nucleotides long and accordingly the primary translation product for PP is a 95 amino acid precursor. Although there is no definitive amino acid sequence which indicates the cleavage site of signal peptides, there are certain general rules (von Heijne, 1983). According to these rules the peptide bond just before the actual PP sequence, between Gly_{29} and Ala_{30} of the primary translational product, constitutes an almost perfect point for processing by the signal peptidase. This notion is supported by the peptide analysis which indicates that the canine PP sequence is located at the far NH_2-terminal end of the precursor, when further cleavage has been blocked (Figure 2). The resulting pro-peptide is among the shortest which have been described in eukaryotes, only 66 amino acids; on the other hand, the putative signal peptide of 29 amino acids is among the longest known.

The PP sequence of 36 residues is followed by a typical cleavage and amidation site, Gly_{66}-Lys_{67}-Arg_{68} (Figure 4). At this site the sequential action of three processing enzymes is supposed to take place. First, a dibasic specific endopeptidase would cleave after Arg_{68}, then a carboxypeptidase would remove the basic residues, and finally an amidating enzyme would transform the Gly_{66} residue into the α-carboxyamide group on Tyr_{65} (Steiner *et al.*, 1980; Bradbury *et al.*, 1982). The COOH-terminal part of the precursor contains the sequence of the human pancreatic icosapeptide (Schwartz *et al.*, 1984) with a COOH-terminal extension of seven amino acids. In pulse-chase experiments the dog icosapeptide is excised from the PP precursor as a COOH-terminally extended intermediate form (Schwartz and Tager, 1981, 1982). Recently, this intermediate form has been characterized as a 25 amino acid peptide in which the processing occurs at a monobasic cleavage site between Arg_{20} and Glu_{21} (unpublished observations, Figure 4). As deduced from the cDNA sequence, this cleavage site and the penultimate residues are identical in the human and canine peptides whereas the COOH-terminal extension otherwise varies both in respect of length and sequence.

We conclude that three different peptides are made from

the mammalian PP precursor: (i) an evolutionary well preserved 36 amino acid peptide, PP itself; (ii) an icosapeptide which is well conserved in its COOH-terminal, but not in its NH_2-terminal part; (iii) a small highly variable oligopeptide of five to seven residues. This conclusion can only be reached through the knowledge of both the structure of the total primary translation product derived from the cDNA, and knowledge of the structure of the peptides, which are the final products. One obvious processing site, Arg_{54}-Arg_{55}, is apparently not used and another non-obvious processing site, Arg_{88}, is used.

Materials and methods

Enzymes and reagents

Reverse transcriptase (RNA-dependent DNA polymerase, EC 2.7.7.49; from avian myeloblastosis virus) was purchased from J.W. Beard, Life Sciences (St. Petersburg, FL). DNA polymerase I 'Klenow fragment' (EC 2.7.7.7; from *E. coli*), polynucleotide kinase (ATP: 5'-dephosphopolynucleotide 5'-phosphotransferase, EC 2.7.1.78; from *E. coli* strain B infected with phage T4) and $(dT)_{12-18}$ were purchased from P-L Biochemicals (Milwaukee, WI). Terminal deoxynucleotidyl transferase (EC 2.7.7.31) was from Bethesda Research Laboratories (Gaithersburg, MD). Nuclease S1 (EC 3.1.30.1; from *Aspergillus oryzae*) was purchased from Sigma. RNasin (human placental ribonuclease inhibitor) was from BioTec (Madison, WI). Restriction endonucleases were from New England BioLabs (Beverly, MA). Oligo(dT)-cellulose was type T-2 from Collaborative Research (Waltham, MA). [α-^{32}P]dTTP (3200 Ci/mmol; 1 Ci = 3.7 x 10^{10} Bq), [α-^{32}P]dCTP (3200 Ci/mmol) and [γ-^{32}P]ATP (7500 Ci/mmol), L-[^{35}S]methionine (>800 Ci/mmol), and L-[2,3,4,5-^{3}H]proline (>100 Ci/mmol), were obtained from New England Nuclear.

Construction of a human pancreatic cDNA library

mRNA from a human pancreas was isolated, purified and used as template to construct an oligo(dT)-primed cDNA library in *E. coli* on *Pst*I linearized pBR327 by methods previously described (Boel *et al.*, 1983). As host strain a hsdR$^-$,M$^+$ derivative of MC 1000 (Casadaban and Cohen, 1980) was used.

Synthesis and end-labeling of oligodeoxyribonucleotides

A mixture of 32 oligodeoxyribonucleotides each 17 bases long, d(TA^C_TTGNG-CCAT^C_TTG^C_TTC), one of which is complementary to human PP mRNA in the region coding for the amino acids Glu-Gln-Met-Ala-Gln-Tyr, was synthetised on a DNA synthesizer (model 380 A, Applied Biosystems, Inc., Foster City, CA) using phosphoramidite chemistry (Beaucage and Caruthers, 1981). The oligonucleotide mixture was purified by electrophoresis on a 20% polyacrylamide gel containing 7 M urea and labeled to a specific activity of 2 x 10^7 c.p.m./pmol by T4 polynucleotide kinase and [γ-^{32}P]ATP as described by Boel *et al.* (1983).

Colony hybridization

Colonies with recombinant plasmids were transferred in an ordered array to Whatman 540 filter paper (82 mm diameter) and immobilized according to Gergen *et al.* (1979). Thirteen filters were hybridized with 4 pmol of 5' end-labeled oligonucleotide mixture for 2 h at 48°C in 15 ml of: 900 mM NaCl, 90 mM Tris-HCl (pH 8.3), 1 mM EDTA, 0.1% SDS with 0.2% each of bovine serum albumin (BSA), Ficoll and polyvinylpyrrolidone. The filters were washed in 900 mM NaCl/90 mM sodium citrate (pH 7) at 42°C with four changes of each 15 min and then exposed to X-ray films for 16 h.

Plasmid DNA isolation and DNA sequence analysis

Plasmid DNA was isolated as described by Birnboim and Doly (1979). DNA restriction fragments generated by *Hind*III and *Bst*NI cleavage of a pancreatic polypeptide specific recombinant were sequenced according to Maxam and Gilbert (1977).

Biosynthetic studies in isolated dog PP cells

A preparation of single cells were obtained from dog duodenal pancreas by trypsin digestion in Ca^{2+} free medium as described previously (Schwartz *et al.*, 1980). Endocrine cells were purified by Percoll density gradient centrifugation (Schwartz and Tager, 1981). Cells were pre-incubated for 6 h in Eagle's basal medium with Hank's salts, and all essential amino acids, and for 1 h in a medium lacking methionine. During the last hour of pre-incubation some cell preparations were incubated in a medium also lacking arginine, but supplemented with L-canavanine, 4 mmol/l (Sigma, C1625) in which they stayed for the rest of the experiment. The labeling was performed with 0.25 mCi L-[^{35}S]methionine and 0.2 mCi L-[2,3,4,5-^{3}H]proline. After 3 h of incubation the cells were centrifuged and washed twice before extraction in 1 ml of 3 M acetic acid and frozen.

Characterization of biosynthetically labeled peptides

The cell extracts were centrifuged and applied to 1.5 x 95 cm Bio-Gel P-30 columns and eluted with 0.5 M acetic acid containing 10 mg/l BSA. Aliquots from the cell extracts were also subjected to SDS-polyacrylamide gel electrophoresis (16% acryl amide) according to Laemmli (1970); the gels were dried and autoradiography was performed. Gel filtration fractions corresponding to the elution position of the pro-PP were pooled, dried and purified on high performance liquid chromatography using a VYDAC C_{18} column (10 x 240 mm, 10 micron particles, and 300 Å pores) which was eluted in trifluoroacetic acid (Pierce, sequanal grade), 0.1% w/v in redistilled water and a gradient of acetonitrile (Merck, Art. No. 30). Biosynthetically labeled pro-PP was subjected to automated sequence analysis by sequential Edman degradation on an Applied Biosystems 470A gas-phase sequenator using the MHTFA1 programming of M. Hunkapiller and hepthane as a supplementary solvent, S_1.

Acknowledgements

We wish to thank Ms. Iben Hjort and Ms. Bente Jørgensen for outstanding technical assistance and Ms. Linda Mygil and Ms. Jytte Ulrich for excellent secretarial assistance. Dr. Mogens Trier Hansen, NOVO, kindly provided us with the hsdR$^-$,M$^+$ derivative of MC 1000. This study was supported by grants 12-3316 and 12-4558 from the Danish Medical Research Council and grants 11-3486 and 11-4027 from the Danish Science Council (to T.W.S.).

References

Attias,J., Schlesinger,M.J. and Schlesinger,S. (1969) *J. Biol. Chem.*, **244**, 3810-3817.
Beaucage,S.L. and Caruthers,M.H. (1981) *Tetrahedron Lett.*, **22**, 1859-1869.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513-1523.
Boel,E., Vuust,J., Norris,F., Norris,K., Wind,A., Rehfeld,J.F. and Marcker,K.A. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2866-2869.
Bradbury,A.F., Finnie,M.D.A. and Smyth,D.G. (1982) *Nature*, **298**, 686-688.
Casadaban,M.J. and Cohen,S.N. (1980) *J. Mol. Biol.*, **138**, 179-207.
Chance,R.E., Moon,N.E. and Johnson,M.G. (1975) in Jaffe,B.M. and Behreman,H.R. (eds.), *Methods of Hormone Radioimmunoassay*, Academic Press, pp. 657-672.
Gates,R.J. and Lazarus,N.R. (1977) *Hormone Res.*, **8**, 189-202.
Gergen,J.P., Stern,R.H. and Wensink,C. (1979) *Nucleic Acids Res.*, **7**, 2115-2135.
Halban,P.A. (1982) *J. Biol. Chem.*, **257**, 13177-13180.
Kimmel,J.R., Hayden,L.J. and Pollock,H.G. (1975) *J. Biol. Chem.*, **250**, 9369-9376.
Laemmli,U.K. (1970) *Nature*, **277**, 680-685.
Larsson,L.-I., Sundler,F. and Håkanson,R. (1976) *Diabetologia*, **12**, 211-216.
Larsson,L.-I., Boder,G.B. and Shaw,N.N. (1977) *Lab. Invest.*, **36**, 593-598.
Malaisse-Lagae,F., Carpentier,J.-L., Patel,Y.C., Malaisse,W.J. and Orci,L. (1977) *Experientia*, **33**, 915-916.
Maxam,A. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-564.
Noe,B.D. (1981) *J. Biol. Chem.*, **256**, 4949-4946.
Pines,M., Rosenthal,G.A. and Applebaum,S.W. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 5480-5483.
Schwartz,T.W. (1983) *Gastroenterology*, **85**, 1411-1425.
Schwartz,T.W. and Tager,H.S. (1981) *Nature*, **294**, 589-591.
Schwartz,T.W. and Tager,H.S. (1982) *Hoppe-Seyler's Z. Physiol. Chem.*, **363**, 889-890.
Schwartz,T.W., Gingerich,R.L. and Tager,H.S. (1980) *J. Biol. Chem.*, **255**, 11494-11498.
Schwartz,T.W., Hansen,H.F., Håkanson,R., Sundler,F. and Tager,H.S. (1984) *Proc. Natl. Acad. Sci. USA*, in press.
Steiner,D.F., Quinn,P.S., Chan,S.J., Marsh,J. and Tager,H.S. (1980) *Ann. N.Y. Acad. Sci.*, **343**, 1-18.
Tatemoto,K. (1982a) *Proc. Natl. Acad. Sci. USA*, **79**, 2514-2518.
Tatemoto,K. (1982b) *Proc. Natl. Acad. Sci. USA*, **79**, 5485-5489.
von Heijne,G. (1983) *Eur. J. Biochem.*, **133**, 17-21.
Zipf,W.B., O'dorisio,T.M., Cataland,S. and Sotos,J. (1981) *J. Clin. Endocrinol. Metab.*, **52**, 1264-1265.

Received on 24 January 1984