Volume 10 Number 24 1982 Nucleic Acids Research

The structure and expression of the preproenkephalin gene

Steve Legon, David M. Glover, John Hughes, Philip J. Lowry[1], Peter W.J. Rigby and Christine J. Watson

Department of Biochemistry, Imperial College of Science and Technology, London SW7 2AZ, UK

Received 16 November 1982; Accepted 30 November 1982

ABSTRACT

Enkephalins are pentapeptides with opioid activity which are found in a wide variety of tissues. Studies of enkephalin-containing peptides from the adrenal gland have established that the mature pentapeptides are derived by proteolytic processing of a precursor protein. We have shown that human adrenal medullary tumours contain mRNA which can be translated *in vitro* to yield a single major enkephalin precursor. The sequence of cloned cDNA shows that the preproenkephalin mRNA encodes four copies of met-enkephalin, two copies of met-enkephalin extended sequences and one copy of leu-enkephalin; each copy is flanked by paired basic amino acids which are presumably recognised by the processing protease. We have used the cloned human cDNA as a hybridisation probe to detect the corresponding mRNAs in rat adrenal gland and, in smaller amounts, in rat brain. We have been unable to detect in brain any other cross-hybridising mRNAs which might encode other putative precursor proteins.

INTRODUCTION

Met-enkephalin (Tyr-Gly-Gly-Phe-Met) and leu-enkephalin (Tyr-Gly-Gly-Phe-Leu) are pentapeptides which compete with and mimic the effects of opiate drugs (1). Although interest in enkephalins stems largely from their possible role in the brain, the richest source of these peptides is the adrenal gland (2) and most work has concentrated on this tissue. Analyses of enkephalin-containing peptides isolated from adrenal glands indicate that the peptides are derived by proteolytic cleavage of a precursor protein (3). The most direct route to the structure of this proenkephalin is to clone and sequence cDNA copies of the relevant mRNA. We hoped that such an approach might, in addition to defining the structure of proenkephalin, also lead to the characterisation of novel peptide hormones found within the same precursor, as occurred previously in the case of pro-opiomelanocortin (4). Moreover, the availability of a cloned probe would enable us to study enkephalin synthesis in brain from which it is difficult to obtain sufficient material for direct peptide analysis.

Although most previous work has concentrated on the proenkephalin of

0305-1048/82/1024-7905$2.00/0

the bovine adrenal gland, we decided, in view of the possible role of enkephalins in pain control and in psychiatric disease (5), to study the corresponding human protein. Phaeochromocytomas, which are tumours of the human adrenal medulla, synthesise very large amounts of enkephalin (6) and therefore seemed likely to be a useful source of the corresponding mRNA. We have used *in vitro* translation techniques to show that such tumours synthesise a single major enkephalin precursor and have cloned and sequenced the cDNA encoding this protein thus defining the complete structure of a human preproenkephalin. The cloned cDNA has been used as a probe in transfer hybridisation experiments to analyse enkephalin precursor synthesis in rat adrenal gland and brain.

MATERIALS AND METHODS

Nucleic Acids. The DNAs of plasmid vectors and of recombinant plasmids were isolated by the alkaline lysis procedure of Birnboim and Doly (7) from saturated cultures of appropriate strains of *Escherichia coli*. The plasmid vector pCD5 (M.R.D. Scott and P.W.J. Rigby, manuscript in preparation) is derived from pXf3 (D. Hanahan, personal communication). pXf3 is derived from pBR322 by two deletions, from nucleotide 1426 to nucleotide 2521 and from nucleotide 3102 to nucleotide 3223 in the sequence of Sutcliffe (8). To construct pCD5 a further deletion, encompassing nucleotides 380 to 649, was introduced. This has the effect of relocating the single *Sal*I site of pBR322 to be only six nucleotides from the single *Bam*I site. cDNA inserted between the *Hin*dIII and *Bam*I sites of pCD5 can then be excised by digestion with *Sal*I (or *Taq*I) and *Cla*I (or *Eco*RI). Plasmid DNAs were labeled with ^{32}P *in vitro* by nick translation (9). Oligonucleotides were labeled to specific activities of approximately 2 x 10^8 c.p.m./µg in 50µl reactions containing 1µg of oligonucleotide, 2.6µM [$\gamma^{32}P$]ATP (>3,000 Ci/mmole, Amersham International), 5 units of T4 polynucleotide kinase (Boehringer Mannheim) and the buffer described by Maxam and Gilbert (10). After 30 min. at 37°C the oligonucleotides were separated from unincorporated ATP by gel filtration on columns of Sephadex G25 run in 10mM Tris HCl, pH7.5, 100mM NaCl, 1mM EDTA (TSE).

mRNA was isolated from frozen tissue by grinding it to a powder under liquid nitrogen and then dispersing it in 100mM NaCl, 1mM $MgCl_2$, 1mM spermidine, 2mg./ml. heparin, 0.5% (v/v) NP40. Nuclei were removed by centrifugation at 5,000rpm. for 5 min. at 4°C in a Sorvall HB4 rotor and the supernatant was adjusted to be 1% (w/v) in SDS and 10mM in EDTA and then

extracted with phenol. RNA was precipitated from the aqueous phase with ethanol and polyadenylated RNA was selected by either two cycles of chromatography on oligodT-cellulose (tumours C and M and all brain RNAs) or a single cycle on polyU-sepharose (tumour B and rat adrenal RNA). mRNA was isolated from cultured KBG cells by extraction of cell lysates with hot phenol (M.R.D. Scott and P.W.J. Rigby, manuscript in preparation).

In vitro Translation Assay for Proenkephalin mRNA. Translation was performed in a nuclease-treated rabbit reticulocyte lysate (11) in the presence of 1mCi/ml. [^{35}S]methionine (Amersham International) and 50µg./ml. mRNA. Translation products were analysed by SDS-polyacrylamide gel electrophoresis (12). Proteins eluted (13) from slices of such gels were dissolved in 200µl 50mM Tris HCl, pH8.0. Methionine residues were oxidised to the sulphoxide by adjusting the solution to be 1M in H_2O_2 and incubating at room temperature for 1h. Excess peroxide was removed by boiling for 10 min. and adjusting the solution to be 100mM in DTT. The samples were then made 1mM in $CaCl_2$ and digested with 50µg./ml. trypsin (TLCK-treated, Worthington) for 2h. at 37°C. After inactivating the trypsin by boiling for 5 min. the peptides were further digested with 20µg./ml. carboxypeptidase B (Sigma) for 1.5h at 37°C and the second enzyme was similarly inactivated. Samples were diluted twofold in NENT buffer (10mM Tris HCl, pH8.0, 150mM NaCl, 0.1mM EDTA, 0.05% (v/v) NP40) and split into two equal aliquots which were incubated overnight at 4°C with 1µl of either control serum or anti-met-enkephalin-sulphoxide (14) antiserum. Immunoprecipitates were collected on inactivated *S. aureus* bacteria (15), washed five times in NENT buffer and then eluted by boiling in 1% (w/v) SDS, 100mM 2-mercaptoethanol. The eluted radioactivity was measured by scintillation counting.

cDNA Cloning Procedures. cDNA was synthesised from mRNA from a human adrenal medullary phaeochromocytoma (tumour B) as follows. 10µg polyadenylated RNA was reverse transcribed in a 200µl reaction containing 50mM Tris HCl, pH8.3, 50mM ammonium acetate, 6mM $MgCl_2$, 10mM DTT, 5µg/ml. oligodT, 100µg/ml. actinomycin D, 1mM each of dATP, dGTP and TTP, 200µM [α^{32}P]dCTP, 200 units human placental RNase inhibitor (Biotec Inc., Madison, Wisconsin) and 100 units AMV reverse transcriptase (Life Sciences Inc., St. Petersburg, Florida). After 2h. at 42°C the reaction mixture was adjusted to be 10mM in EDTA and 0.1% (w/v) in SDS and incubated for a further 15 min. with 10µg proteinase K (British Drug Houses). It was then adjusted to be 0.3N in NaOH, incubated for 30 min. at 60°C, neutralised by the addition of solid MOPS buffer and deproteinised by extraction with phenol/chloroform. Phenol was removed by

ether extraction and the cDNA was purified by gel filtration on a 2ml. column of Sephadex G100 run in TSE buffer. The second cDNA strand was synthesised using the large subfragment of *E. coli* DNA polymerase I and self-priming (16). The terminal hairpin loops were cleaved with endonuclease S1 (Sigma) and the cDNA was tailed (17) with approximately 25 dGMP residues using terminal transferase (Ratliff Biochemicals, Los Alamos, New Mexico). DNA of the plasmid vector pCD5 was digested with *Bst*I and *Hin*dIII and tailed with a similar number of dCMP residues. cDNA and vector were annealed in TSE buffer at an initial temperature of 65°C with cooling to 40°C over 4h. The resultant hybrid DNA molecules were introduced by transformation into competent *E. coli* DH1 cells (D. Hanahan, personal communication) using a transformation protocol developed in this laboratory (M.R.D. Scott and P.W.J. Rigby, manuscript in preparation). Transformants were plated directly onto nitrocellulose filters, selection being for the ampicillin resistance determinant of pCD5. We obtained 20,000 clones from 0.3µg cDNA; these were generally grown at a density of 3,000 colonies per 90mm plate. Replica filters were prepared for screening as described by Hanahan and Meselson (18).

DNA Sequencing. The chemical sequencing procedures of Maxam and Gilbert (10) were employed. DNA fragments were labeled at the vector *Eco*RI and *Sal*I sites by incubating the fragments (50µg/ml.) in 50mM Tris HCl, pH8.3, 6mM magnesium acetate, 20mM DTT, 0.5µM [α^{32}P]dNTP (>3,000 Ci/mmole, Amersham International) and 40 units/ml. reverse transcriptase. When necessary, unlabeled triphosphate was present at 250µM. Labeling at *Sma*I or *Pst*I sites in the cDNA insert was performed by incubating DNA (500µg/ml) in 30mM Tris HCl, pH7.4, 140mM potassium cacodylate, 1mM $CoCl_2$, 1mM DTT, 1µM [α^{32}P] cordycepin (Amersham International) and 0.5 units/µl terminal transferase (19).

Transfer Hybridisation. RNAs were fractionated by electrophoresis in 1.2% (w/v) agarose gels containing formaldehyde and then transferred to nitrocellulose filters (20). Human, yeast and *Drosophila* rRNAs were used as size markers. Hybridisations with oligonucleotide probes (approximately 2ng/ml.) were performed in 10mM Tris HCl, pH7.5, 500mM NaCl, 1mM EDTA, 1 x Denhardt's solution (21), 100µM ATP at 15°C for 16h. Filters were washed in several changes of hybridisation buffer at 25°C for 2h. Hybridisations with plasmid DNA probes were performed in 5 x SSC, 50% (v/v) formamide, 0.1% (w/v) SDS, 1 x Denhardt's solution at 42°C or 30°C, as indicated, for 16h. Filters were washed in hybridisation buffer at 30°C for 1h. then in 1 x SSC (150mM

NaCl, 15mM trisodium citrate) at 30°C for 1h. In some cases, indicated in the figure legends, filters were subsequently washed in 1 x SSC for 10 min. at a higher temperature.

RESULTS

In vitro Translation Assay for Proenkephalin mRNA. The ability to synthesise proenkephalin *in vitro* could provide a means of screening cDNA clones and would also be of value in comparing the primary translation products of adrenal gland and brain mRNAs. In developing such an assay we have used mRNA from human adrenal medullary phaeochromocytomas, since such tumours over-produce enkephalins (6) and are thus likely to be a rich source of proenkephalin mRNA. Unfortunately, none of the many anti-enkephalin antisera that we have tested immunoprecipitate proenkephalin and we have therefore had to resort to an indirect assay which depends upon the fact that enkephalins can be liberated from proenkephalin by digestion with trypsin and carboxypeptidase B. The *in vitro* translation products were fractionated by electrophoresis in an SDS-polyacrylamide gel, eluted from slices and digested to release enkephalins for immunoassay. Figure 1 shows that we detect a single peak of immunoprecipitable met-enkephalin-sulphoxide corresponding to a proenkephalin of apparent molecular weight 36,500. This peak represents 1% of the total methionine incorporated into protein and the sensitivity of this assay allows detection of proenkephalin in unfractionated *in vitro* translation products (data not shown). These data

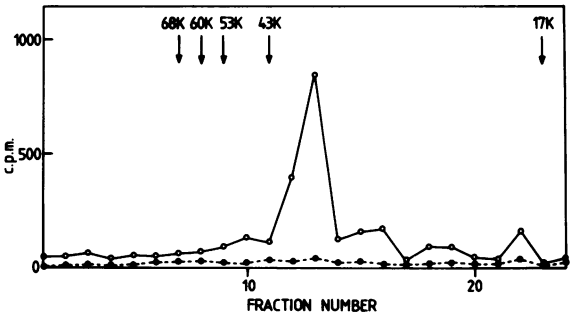

Figure 1. Synthesis of proenkephalin *in vitro* in response to phaeochromocytoma mRNA. [^{35}S]methionine-labeled proteins were fractionated by SDS-polyacrylamide gel electrophoresis and enkephalin sequences were located by radioimmunoassay. Marker proteins were: BSA (68k), catalase (60k), glutamate dehydrogenase (53k), ovalbumin (43k) and globin (17k). O——O, anti-enkephalin serum; ●-----●, control serum.

gave us an estimate of the abundance of proenkephalin mRNA and defined the size of the precursor protein. Using this procedure we were unable to detect enkephalin amongst the *in vitro* translation products of whole rat brain RNA.

Synthetic Oligonucleotide Probes for Proenkephalin Coding Sequences. Although the *in vitro* translation assay could be used to screen cDNA clones a more direct screening protocol was afforded by the observation that bovine adrenal gland proenkephalin contains the sequence Trp-Trp-Met-Asp-Tyr (22), which is highly favourable for oligonucleotide synthesis. We reasoned that this unusual sequence was likely to be conserved between the bovine and human proteins and therefore used the corresponding oligonucleotides, TAATCCATCCACC and TAGTCCATCCACC, as hybridisation probes. The labeled oligonucleotides were used as probes in transfer hybridisation experiments to detect proenkephalin mRNA in electrophoretically fractionated phaeochromocytoma RNA (Figure 2). At equivalent autoradiographic exposures one of the oligonucleotides (TAGTCCATCCACC) gave a strongly hybridising band at about 1.4Kb while the other hybridised only weakly. This mRNA is not present in a control RNA from the human tumour cell line KBG and is of an appropriate size to code for the proenkephalin detected by *in vitro* translation.

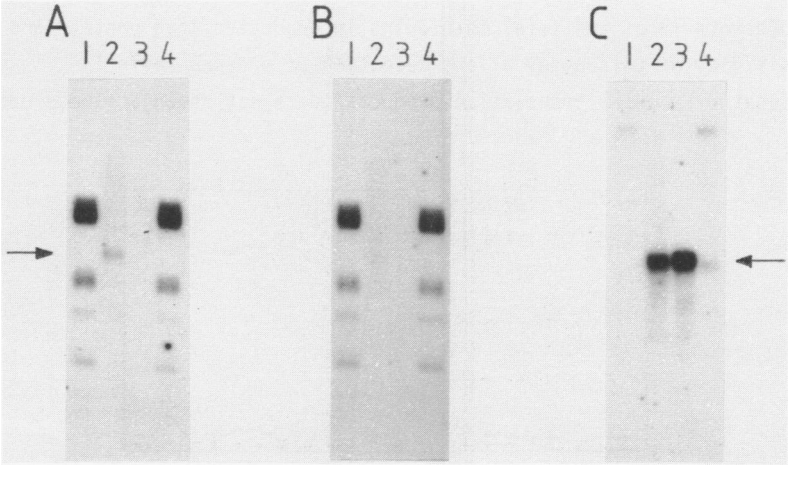

Figure 2. Transfer hybridisation analysis of phaeochromocytoma mRNA. Panel A: Hybridisation with ^{32}P-labeled TAGTCCATCCACC; samples were, (1) DNA size markers; (2) phaeochromocytoma RNA; (3) KBG RNA; (4) DNA size markers. Panel B: Hybridisation with ^{32}P-labeled TAATCCATCCACC; samples as in Panel A. Panel C: Hybridisation at 42^{o}C with ^{32}P-labeled preproenkephalin cDNA clone; samples were, (1) KBG RNA; (2) tumour C RNA; (3) tumour M RNA; (4) tumour B RNA. The arrows indicate preproenkephalin mRNA.

The oligonucleotide TAGTCCATCCACC was therefore labeled and used to screen a library of 20,000 cDNA clones derived from phaeochromocytoma mRNA. 30 positive signals were obtained (Figure 3). One positive signal came from a filter carrying only 300 colonies (Figure 3A) and the relevant colony could thus be picked from the master plate. In the other cases the colony density was tenfold higher and a second round of screening was required. Plasmid DNA was isolated from the clone picked from the low density plate, labeled and hybridised to gel-fractionated phaeochromocytoma mRNA (Figure 2C). The cloned cDNA and the oligonucleotide detect mRNAs of the same size. The additional faint bands in Figure 2C are due to hybridisation of the dG:dC tails in the plasmid to contaminating rRNA.

It is clear that the abundance of proenkephalin mRNA varies considerably between different tumours (Figure 2C). Such variation accounts for the discrepancy between the frequency (1 in 700) at which we recovered proenkephalin cDNA clones from a library constructed from mRNA from tumour B and our estimate (1%) of the abundance of proenkephalin mRNA, which derives from *in vitro* translation of mRNA from tumours C and M.

Sequence Analysis of Cloned Proenkephalin cDNA. We first determined the size of the cDNA insert in each of our thirty clones. This was straightforward because the structure of the plasmid vector pCD5 allows the insert to be excised by digestion with *Eco*RI and *Sal*I. This structural feature of

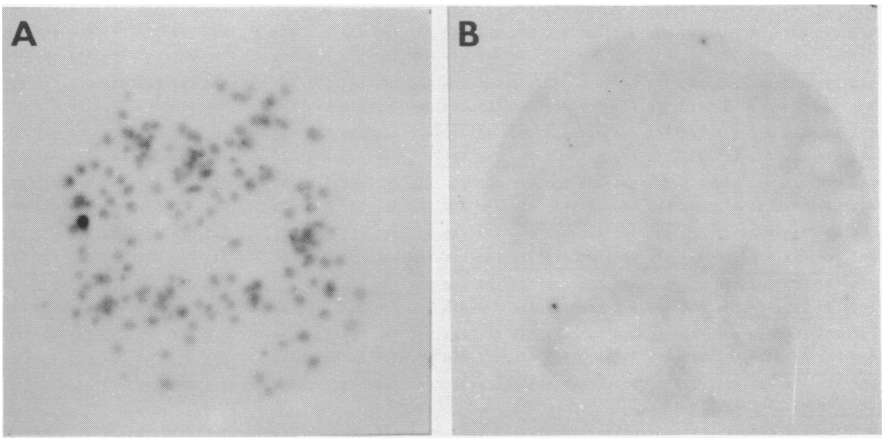

Figure 3. Colony screening with labeled oligonucleotides. Filter A carried 300 colonies, one of which was positive; Filter B carried 3000 colonies, two of which were positive. The probe was ^{32}P-labeled TAGTCCATCCACC.

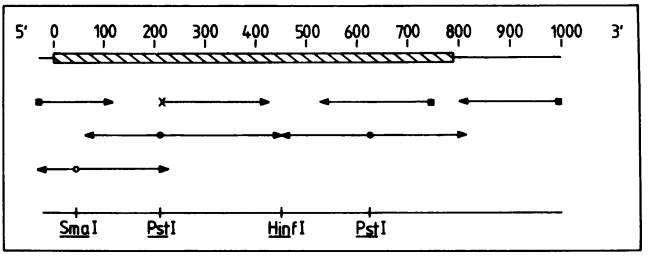

Figure 4. DNA sequencing strategy. The entire sequence is shown in the top line with the translated region cross-hatched. Sites at which labeling was performed are indicated thus: ■ , *Sal*I; X, *Eco*RI; ●, *Pst*I; O, *Sma*I. Fragments labeled at the vector *Sal*I and *Eco*RI sites were derived from different clones, those labeled at *Pst*I and *Sma*I sites from the largest clone. 5' and 3' indicate the sense of the mRNA.

the vector was also exploited in designing our sequencing strategy (Figure 4). Because the cDNA inserts are of different lengths the clones themselves provide the overlapping series of DNA fragments required for sequencing. Most of the sequence could thus be determined by sequencing from the

```
AATTGGCCTGCTCCATCCGAACAGCGTCAACTCC  ATG GCG CGG TTC CTG ACA CTT TGC ACT TGG CTG CTG TTG CTC GGC CCC GGG CTC
                                    Met Ala Arg Phe Leu Thr Leu Cys Thr Trp Leu Leu Leu Leu Gly Pro Gly Leu

CTG GCG ACC GTG CGG GCC GAA TGC AGC CAG GAT TGC GCG ACG TGC AGC TAC CGC CTA GTG CGC CCG GCC GAC ATC AAC TTC
Leu Ala Thr Val Arg Ala Glu Cys Ser Gln Asp Cys Ala Thr Cys Ser Tyr Arg Leu Val Arg Pro Ala Asp Ile Asn Phe

CTG GCT TGC GTA ATG GAA TGT GAA GGT AAA CTG CCT TCT CTG AAA ATT TGG GAA ACC TGC AAG GAG CTC CTG CAG CTG TCC
Leu Ala Cys Val Met Glu Cys Glu Gly Lys Leu Pro Ser Leu Lys Ile Trp Glu Thr Cys Lys Glu Leu Leu Gln Leu Ser

AAA CCA GAG CTT CCT CAA GAT GGC ACC AGC ACC CTC AGA GAA AAT AGC AAA CCG GAA GAA AGC CAT TTG CTA GCC AAA AGG
Lys Pro Glu Leu Pro Gln Asp Gly Thr Ser Thr Leu Arg Glu Asn Ser Lys Pro Glu Glu Ser His Leu Leu Ala Lys Arg

TAT GGG GGC TTC ATG AAA AGG TAT GGA GGC TTC ATG AAG AAA ATG GAT GAG CTT TAT CCC ATG GAG CCA GAA GAA GAG GCC
Tyr Gly Gly Phe Met Lys Arg Tyr Gly Gly Phe Met Lys Lys Met Asp Glu Leu Tyr Pro Met Glu Pro Glu Glu Glu Ala

AAT GGA AGT GAG ATC CTC GCC AAG CGG TAT GGG GGC TTC ATG AAG AAG GAT GCA GAG GAG GAC GAC TCG CTG GCC AAT TCC
Asn Gly Ser Glu Ile Leu Ala Lys Arg Tyr Gly Gly Phe Met Lys Lys Asp Ala Glu Glu Asp Asp Ser Leu Ala Asn Ser

TCA GAC CTG CTA AAA GAG CTT CTG GAA ACA GGG GAC AAC CGA GAG CGT AGC CAC CAC CAG GAT GGC AGT GAT AAT GAG GAA
Ser Asp Leu Leu Lys Glu Leu Leu Glu Thr Gly Asp Asn Arg Glu Arg Ser His His Gln Asp Gly Ser Asp Asn Glu Glu

GAA GTG AGC AAG AGA TAT GGG GGC TTC ATG AGA GGC TTA AAG AGA AGC CCC CAA CTG GAA GAT GAA GCC AAA GAG CTG CAG
Glu Val Ser Lys Arg Tyr Gly Gly Phe Met Arg Gly Leu Lys Arg Ser Pro Gln Leu Glu Asp Glu Ala Lys Glu Leu Gln

AAG CGA TAT GGG GGC TTC ATG AGA AGA GTA GGT CGC CCA GAG TGG TGG ATG GAC TAC CAG AAA CGG TAT GGA GGT TTC CTG
Lys Arg Tyr Gly Gly Phe Met Arg Arg Val Gly Arg Pro Glu Trp Trp Met Asp Tyr Gln Lys Arg Tyr Gly Gly Phe Leu

AAG CGC TTT GCC GAG GCT CTG CCC TCC GAC GAA GAA GGC GAA AGT TAC TCC AAA GAA GTT CCT GAA ATG GAA AAA AGA TAC
Lys Arg Phe Ala Glu Ala Leu Pro Ser Asp Glu Glu Gly Glu Ser Tyr Ser Lys Glu Val Pro Glu Met Glu Lys Arg Tyr

GGA GGA TTT ATG AGA TTT TAATATTTTTCCCACTAGTGGCCCCAGGCCCCAGCAAGCCTCCCTCCATCCTCCAGTGGGAAACTGTTGATGGTGTTTTATT
Gly Gly Phe Met Arg Phe

GTCATGTGTTGCTTGCCTTGTATAGTTGACTTCATTGTCTGGATAACTATACAACCTGAAAACTGTCATTTCAGGTTCTGTGCTCTTTTTGGAGTCTTTAAGCTCAG
TATT
```

Figure 5. Sequence of human preproenkephalin. Enkephalin sequences are boxed. The sequence complementary to our oligonucleotide probe is underlined.

flanking vector sites into the inserts carried by a series of the clones. The remainder of the sequence was determined by labeling at internal restriction sites (Figure 4). It is 1,028 base pairs in length and includes the entire coding sequence of preproenkephalin (Figure 5).

Preproenkephalin mRNA in Rat Adrenal Gland and Brain. The predicted amino acid sequence of human adrenal preproenkephalin accounts for all of the precursor peptides detected in adrenal gland and most of those from brain. However, it has been reported that brain contains a proenkephalin of apparent molecular weight 92,000 (23). Such a protein could be derived from the same genomic sequence as the adrenal gland precursor if differential splicing were to operate, it could be the product of a distinct but evolutionarily related gene or it could be the result of convergent evolution and have no sequence homology to the preproenkephalin we have characterised. To explore the synthesis of preproenkephalin in brain we

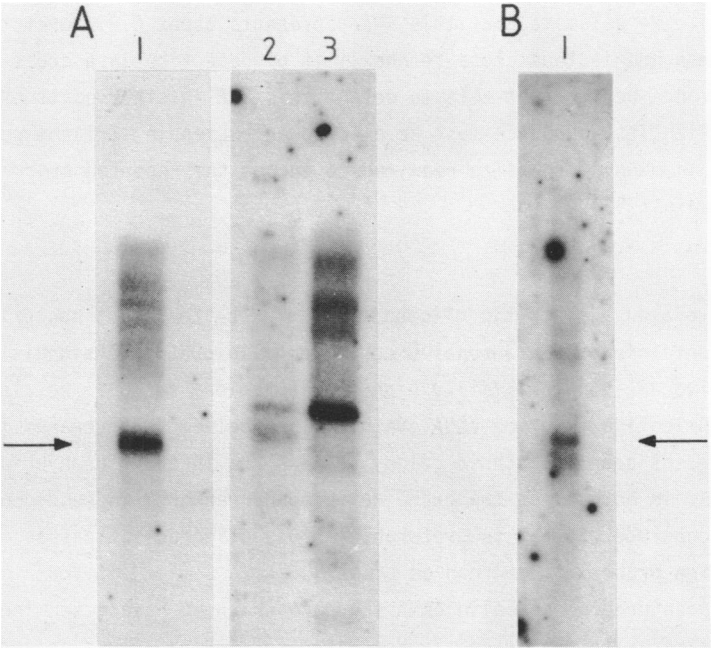

Figure 6. Transfer hybridisation analysis of rat adrenal gland and brain mRNA. Panel A: samples are, (1) human phaeochromocytoma mRNA; (2) rat adrenal gland mRNA; (3) rat brain mRNA. Hybridisation with ^{32}P-labeled human preproenkephalin cDNA was at 30°C; the final wash was at 45°C. Track 1 was exposed for a shorter time than tracks 2 and 3. Panel B: rat striatum mRNA. In this case the final wash was at 55°C. The arrows indicate preproenkephalin mRNAs.

have used our cDNA clone to probe brain RNA in transfer hybridisation experiments.

Suitable human tissue was not readily available so we chose to analyse mRNA from rat adrenal gland and brain. Figure 6A shows that in rat adrenal gland mRNA there are two prominent bands, one the same size as human adrenal gland preproenkephalin mRNA, the other slightly larger, about 1.6Kb. This latter RNA is also found in rat brain but has no counterpart in human adrenal glands. Other, faint, bands are also observed. However, hybridisation with labeled, tailed vector shows that all bands apart from that of the same size as human adrenal gland preproenkephalin mRNA result from hybridisation to the oligo dG:dC tails in the recombinant plasmid. The background due to these artefactual hybridisations can be greatly reduced by washing the filters at higher temperatures. In this way we were able to detect in rat striatum, the part of the brain richest in enkephalin (31), an mRNA of the same size as the human adrenal gland preproenkephalin mRNA (Figure 6B). We estimate that this RNA represents about 0.01% of total striatal mRNA and is thus close to the limit of detection in a cross-species hybridisation. We were not able to detect this RNA in other parts of the brain in which its abundance must be even lower. Even in striatum we could not detect an mRNA of the size required to encode the reported precursor of molecular weight 92,000.

DISCUSSION

We have used a synthetic oligonucleotide probe to detect bacterial colonies containing human adrenal preproenkephalin cDNA; the signals obtained were unequivocal even in a fairly high density colony screen. We also used the oligonucleotide to prime cDNA synthesis. Labeled cDNA prepared in this way was used as a probe both in colony screens and in transfer hybridisation experiments; in both cases the probe detected non-enkephalin sequences. We therefore conclude that it is preferable to use oligonucleotides as direct hybridisation probes rather than as primers.

The human preproenkephalin cDNA sequence which we have determined is generally similar to that recently described for bovine adrenal gland preproenkephalin (27, 28). During the latter stages of this work Comb *et al.* reported on the sequence of human adrenal gland preproenkephalin, also using a cDNA clone derived from phaeochromocytoma mRNA (29). Their sequence differs from ours in three positions none of which affects the predicted amino acid sequence. While this manuscript was in preparation Noda *et al.*

reported the sequence of the human preproenkephalin gene (30). Their sequence differs from ours in just one position and shows that part of the cDNA clone of Comb *et al.* results from an artefactual inversion. We have sequenced both strands of our cDNA clone in the regions of these discrepancies and are confident of our assignments. These differences could reflect errors introduced during cloning or could result from polymorphisms present in the human population. The latter possibility is particularly likely as restriction fragment length polymorphisms are commonly detected in the human population using the cloned preproenkephalin cDNA as a probe (M. Hill and R. Williamson, personal communication).

In view of the general agreement between the three sequences we will restrict our discussion to those areas in which our work adds to our understanding of the enkephalin system, namely the disputed issues of the size of the enkephalin precursor protein in human adrenal glands (25) and the possibility that the pathway of enkephalin biosynthesis in the brain might be quite different (23). We refer readers to other publications of the preproenkephalin sequence (29, 30) for a full discussion of its properties.

Most studies of enkephalin biosynthesis have used adrenal glands because they are the richest source of these peptides (2). It has been a disturbing possibility that such studies might prove to be relevant only to the adrenal gland and that the pathway of enkephalin biosynthesis in brain might be quite different, proceeding *via* a major precursor protein of molecular weight 92,000 (23). Even in the case of the human adrenal gland there are reports of abundant precursors with molecular weights of between 70,000 and 180,000 (25). Our *in vitro* translation studies of human adrenal gland mRNA reveal no such precursors (Figure 1) and it is likely that the previously reported precursors were in fact aggregates which were not disrupted by the non-denaturing gel filtration conditions employed (25). Our data on the translation of human adrenal gland mRNA are in agreement with those of Dandekar and Sabol on the bovine system (24); they similarly found no evidence for very high molecular weight precursors. The estimated abundance and size of the precursor detected by *in vitro* translation are sufficiently close to the figures obtained by cloning and sequencing to make us confident that the sequenced cDNA represents the major preproenkephalin mRNA in the human adrenal gland.

The relevance of studies on the adrenal gland to the biosynthesis of enkephalin in the brain can still not be completely assessed. Our *in vitro*

translation assay is not sufficiently sensitive to detect precursor synthesis from brain mRNA so we have had to resort to transfer hybridisation experiments using the adrenal gland cDNA as probe. Experiments with rat adrenal gland mRNA show sufficient sequence homology between rat and human to allow cross-species hybridisation and further showed that rat and human adrenal gland preproenkephalin mRNAs are of essentially the same size (Figure 6A). We could not detect an adrenal gland type mRNA in whole rat brain (Figure 6A) but when mRNA isolated from various anatomical regions of the brain was analysed it was possible to detect hybridisation to mRNA extracted from the striatum (Figure 6B). This mRNA is the same size as adrenal gland preproenkephalin mRNA and its detection only in the striatum is significant as this is the part of the brain known to be richest in enkephalin (31). This is a direct indication that work on the adrenal gland is relevant to the brain and shows clearly that the gene encoding the adrenal gland preproenkephalin is also expressed in brain striatum.

However, our data should not be taken to imply that the mRNA we have detected encodes the only, or even the major, preproenkephalin in brain. Recently a precursor to leu-enkephalin of molecular weight 29,000 has been reported (32). The sequence of this preproenkephalin B is unrelated to that of the preproenkephalin we have described and it remains possible that the precursor of molecular weight 92,000 (23) is encoded by yet another unrelated gene. The examples seen in other systems in which related proteins are generated by differential splicing of the transcript of a single gene or by the duplication of an ancestral gene with subsequent sequence divergence are not necessarily helpful in this case because the active enkephalin sequence is so short that it could have independently arisen several times during evolution. Occassionally the occurrence of this sequence will have conferred useful properties on the gene and, in the case of the preproenkephalin described here, local duplications have resulted in multiple enkephalin sequences within a single gene. This is evident from the conservation seen in these sequences, with four of the first five enkephalins using the same codons and with the fifth differing in only a single residue. In other cases, for example β-endorphin (4), a single met-enkephalin sequence confers opiate activity upon a larger peptide containing no duplications. Given the potent activity and small size of neuroactive peptides it seems likely that there are many other cases of families of genes which are related not by their evolutionary background but only by the function of their encoded peptides.

ACKNOWLEDGEMENTS

We are grateful to Doug Hanahan (Harvard University) for making pXf3 and *E. coli* DH1 available to us, to Mike Eaton (Celltech Ltd) for very generously synthesising the oligonucleotide probes and to Profs. G.M. Besser and L.H. Rees (St. Bartholomew's Hospital) for providing us with tumour samples.

D.M.G. and P.W.J.R. hold Career Development Awards from the Cancer Research Campaign. This research was performed under a contract from the National Research Development Corporation.

[1] Permanent Address: Pituitary Hormone Laboratory, St. Bartholomew's Hospital, 51-53 Bartholomew Close, West Smithfield, London EC1.

REFERENCES

1. Hughes, J., Smith, T.W., Kosterlitz, H.W., Fothergill, L.A., Morgan, B.A. and Morris, H.R. (1979). Nature 258, 577-579.
2. Lewis, R.V., Stern, A.S., Rossier, J., Stein, S. and Udenfriend, S. (1979). Biochem. Biophys. Res. Comm. 89, 822-829.
3. Kimura, S., Lewis, R.V., Stern, A.S., Rossier, J., Stein, S. and Udenfriend, S. (1980). Proc. Natl. Acad. Sci. U.S.A. 77, 1681-1685.
4. Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A.C.Y. and Cohen, S.N. (1979). Nature 278, 423-427.
5. Hughes, J., Beaumont, A., Fuentes, J., Malfoy, B. and Unsworth, C. (1980). J. Exp. Biol. 89, 239-255.
6. Corder, R. and Lowry, P.J. (1980). Biochim. Biophys. Res. Comm. 95, 665-673.
7. Birnboim, H.C. and Doly, J. (1979). Nucl. Acids. Res. 7, 1513-1523.
8. Sutcliffe, J.G. (1979). Cold Spring Harbor Symp. Quant. Biol. 43, 77-90.
9. Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). J. Mol. Biol. 113, 237-251.
10. Maxam, A. and Gilbert, W. (1980). Methods in Enzymology 65, 499-560.
11. Pelham, H.R.B. and Jackson, R.J. (1976). Eur. J. Biochem. 67, 247-256.
12. Laemmli, U.K. (1970). Nature 227, 680-685.
13. Rundell, K., Collins, J.K., Tegtmeyer, P., Ozer, H.L., Lai, C-J. and Nathans, D. (1977). J. Virol. 21, 636-646.
14. Clement-Jones, V., Lowry, P.J., Rees, L.H. and Besser, G.M. (1980). Nature 283, 295-297.
15. Kessler, S.W. (1975). J. Immunol. 115, 1617-1624.
16. Humphries, P., Old, R.W., Coggins, L.W., McShane, T., Watson, C.J. and Paul, J. (1978). Nucl. Acids Res. 5, 905-924.
17. Roychoudhury, R., Jay, E. and Wu, R. (1976). Nucl. Acids Res. 3, 101-116.
18. Hanahan, D. and Meselson, M. (1980). Gene 10, 63-67.
19. Tu, C.P.D. and Cohen, S.N. (1980). Gene 10, 177-183.
20. Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M. and Darnell, J.E. Jr. (1981). Cell 23, 731-739.
21. Denhardt, D.T. (1966). Biochim. Biophys. Res. Comm. 23, 641-646.

22. Kilpatrick, D.L., Taniguchi, T., Jones, B.N., Stern, A.S., Shively, J.E. Hullihan, J., Kimura, S., Stein, S. and Udenfriend, S. (1981). Proc. Natl. Acad. Sci. U.S.A. 78, 3265-3268.
23. Beaumont, A., Fuentes, J.A., Hughes, J. and Metters, K.M. (1980). FEBS Lett. 122, 135-137.
24. Dandekar, S. and Sabol, S.L. (1982). Proc. Natl. Acad. Sci. U.S.A. 79, 1017-1021.
25. Giraud, P. and Eiden, L.E. (1981). Biochem. Biophys. Res. Comm. 99, 969-975.
26. Blobel, G. and Dobberstein, B. (1975). J. Cell Biol. 67, 852-862.
27. Noda, M., Furutani, Y., Takahashi, H., Toyosato, M., Hirose, T., Inayama, S., Nakanishi, S. and Numa, S. (1982). Nature 295, 202-206.
28. Gubler, U., Seeburg, P., Hoffman, B.J., Gage, L.P. and Udenfriend, S. (1982). Nature 295, 206-209.
29. Comb, M., Seeburg, P.H., Adelman, J., Eiden, L. and Herbert, E. (1982). Nature 295, 663-666.
30. Noda, M., Teranishi, Y., Takahashi, H., Toyosato, M., Notake, M., Nakanishi, S. and Numa, S. (1982). Nature 297, 431-434.
31. Lewis, R.V., Stein, S., Gerber, M., Rubinstein, M. and Udenfriend, S. (1978). Proc. Natl. Acad. Sci. U.S.A. 75, 4021-4023.
32. Kakideni, H., Furutani, Y., Takahashi, H., Noda, M., Morimoto, Y., Hirose, T., Asai, M., Inayama, S., Nakanishi, S. and Numa, S. (1982). Nature 298, 245-249.