Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 7651–7655, December 1984
Neurobiology

Molecular cloning and sequence determination of rat preproenkephalin cDNA: Sensitive probe for studying transcriptional changes in rat tissues

(precursor protein/neuropeptides)

RICHARD D. HOWELLS*, DANIEL L. KILPATRICK*, RAM BHATT†, JOHN J. MONAHAN†, MOHINDAR POONIAN†, AND SIDNEY UDENFRIEND*

Roche Research Center, *Roche Institute of Molecular Biology and †Department of Molecular Genetics, Hoffmann-La Roche, Inc., Nutley, NJ 07110

Contributed by Sidney Udenfriend, August 16, 1984

ABSTRACT **A cDNA probe was prepared to investigate the regulation of proenkephalin biosynthesis in the rat. This was necessary because human and bovine proenkephalin cDNA were not sensitive enough for the accurate detection of preproenkephalin mRNA in tissues that contain low copy numbers of this message, such as the adrenal gland. The rat probe was prepared in the following manner. Preproenkephalin mRNA was enriched by sucrose gradient centrifugation of poly(A)-containing mRNA from rat brain and was used as a template for double-stranded cDNA synthesis. The resulting cDNA was inserted into the plasmid pBR322, and recombinant plasmids were used to transform *Escherichia coli* RR1 cells. A synthetic oligodeoxyribonucleotide (30 bases long) with a sequence that had previously been shown to be identical in bovine and human preproenkephalin cDNA was prepared to screen the clone bank. The plasmid with the longest cDNA insert (about 1200 bases) from the positive clones was isolated, and the sequence of the entire protein coding region was determined. Like the bovine and human gene products, rat preproenkephalin contains four [Met]enkephalin sequences and one copy each of [Leu]enkephalin, [Met]enkephalin-Arg^{6}-Gly^{7}-Leu^{8}, and [Met]enkephalin-Arg^{6}-Phe^{7}. Rat preproenkephalin is 80% and 83% homologous to the bovine and human forms, respectively, at the nucleotide level and is 82% homologous to both species at the amino acid level. Rat preproenkephalin contains 269 amino acid residues, making it larger than the human (267 residues) and bovine (263 residues) precursors. The sensitivity for detection of rat preproenkephalin mRNA with the rat cDNA was several times greater than with the corresponding cDNAs from bovine and human sources.**

Since the discovery of the enkephalin pentapeptides nearly a decade ago (1), much has been learned about the biosynthesis of [Met]- and [Leu]enkephalin. In addition to the isolation and sequence determination of many enkephalin-containing peptides from bovine adrenal medulla (see ref. 2 for review), the mRNA from bovine adrenal medulla and human pheochromocytoma, which encodes the precursor protein preproenkephalin, has been cloned and sequenced (3–5).

The exact function of the enkephalin-containing peptides is still not known even after many years of intensive investigation. It is likely that knowledge concerning the mechanism of action of these peptides and the regulation of their metabolism will come from research on the rat, the animal that has provided most of our knowledge concerning opiate pharmacology. Studies in our laboratory on the stimulation of proenkephalin synthesis in the rat adrenal gland following denervation (6–8) have been hampered because human and bovine cDNA were found not to be sufficiently sensitive as probes for detecting preproenkephalin mRNA in this rat tissue, which has a low copy number of the message under normal conditions. Rat preproenkephalin cDNA was cloned to provide a homologous and, quite likely, a more sensitive probe for monitoring transcriptional changes in rat tissues resulting from pharmacological or surgical manipulations. This report describes the molecular cloning and sequencing of the protein coding region of rat prepropenkephalin cDNA. Evidence is presented that rat preproenkephalin cDNA is indeed a more sensitive probe than either the bovine or human analogs for detecting preproenkephalin mRNA in rat tissues.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

T_4 polynucleotide kinase and DNA polymerase I were purchased from Boehringer Mannheim. [γ-^{32}P]ATP (specific activity $\geq$ 5000 Ci/mmol; 1 Ci = 37 GBq) and [α-^{32}P]dCTP (3000 Ci/mmol) were products of Amersham. Proteinase K and CsCl were from E. M. Reagents, Darmstadt, FRG. Oligo(dT)-cellulose and oligo(dT) primers were obtained from P-L Biochemicals. Reverse transcriptase from avian myeloblastosis virus was purchased from Life Sciences. Ribonuclease H was from Bethesda Research Laboratories. Terminal deoxynucleotidyl transferase was a product of United States Biochemical, Cleveland, OH. All restriction endonucleases were from New England Biolabs. Deoxyribonuclease I was obtained from Worthington.

Brains were removed from male Fisher 344 rats (Taconic Farms, Germantown, NY), and the cortex and cerebellum were removed and discarded, since the enkephalin-containing peptide content in these areas is relatively low (9, 10). Total RNA was extracted from the remaining brain tissue by the method of Hall *et al.* (11). Poly(A)-containing mRNA was obtained by passage over oligo(dT)-cellulose as described (12), except that the intermediate 0.1 M KCl wash was omitted. The mRNA was further fractionated on 0–30% sucrose gradients that were centrifuged at 38,000 rpm in a SW 40 rotor for 18 hr at 4°C. An aliquot from each fraction was subjected to electrophoresis on agarose/methylmercuric hydroxide gels (13) and transferred to nitrocellulose filters (14). The probe used for hybridization to the filters was bovine proenkephalin cDNA (3), approximately 1 kilobase in length, that was kindly provided by U. Gubler. The cDNA was nick-translated (15) with [α-^{32}P]dCTP and hybridized to the nitrocellulose blots at 42°C as described (14).

Rat cDNA was synthesized by the method of Gubler and Hoffman (16) using the fraction of the sucrose gradient that contained the most preproenkephalin mRNA. The double-stranded cDNA was poly(dG)-tailed with terminal deoxynucleotidyl transferase (17) and hybridized to poly(dC)-tailed pBR322 that had been linearized by digestion with *Eco*RV restriction endonuclease. Under the conditions used, 20–30 mol of dGMP or dCMP were added per mol of cDNA or vector DNA. The tailed cDNA and pBR322 were annealed at

58°C for 1 hr at a ratio of approximately 1:25, based on weight. The recombinant plasmids were used to transform *E. coli* RR1 cells as described (18).

Transformed bacteria were transferred to nitrocellulose (19), and those containing preproenkephalin cDNA were identified by *in situ* hybridization with a proenkephalin-specific 30-base-long oligodeoxyribonucleotide that was synthesized by the procedure of Miyoshi *et al.* (20) and purified on reverse-phase HPLC. The filters were hybridized at 30°C as described (14) with the synthetic probe that had been labeled with [γ^{32}P]ATP and polynucleotide kinase (21). Chimeric plasmids were isolated from the positive clones by the method of Birnboim and Doly (22) and analyzed to determine which had the longest cDNA. The plasmid of interest (hereafter referred to as pRPE-1) was amplified by treating the bacteria with chloramphenicol (23) and was sequenced by the technique of Maxam and Gilbert (21).

Dot-blot assays with rat, bovine, and human preproenkephalin cDNAs were performed as described by White and Bancroft (24). Bacteria containing human preproenkephalin chimeric plasmids (pHPE-9) were kindly provided by M. Comb and E. Herbert. A 918-base-pair *Hin*cII fragment derived from pHPE-9 was prepared for use as a probe. Rat preproenkephalin cDNA for use as a probe was obtained by digesting the plasmid pRPE-1 with *Pvu* II and then isolating the 435-base-pair fragment, hereafter referred to as pRPE-1-(165–600). The cDNAS were nick-translated and hybridized to the nitrocellulose dot blots at 42°C as described (14).

RESULTS AND DISCUSSION

Design and Synthesis of a Proenkephalin-Specific Oligodeoxyribonucleotide Probe. Our initial attempts at screening the rat brain cDNA library with either bovine or human proenkephalin cDNA were unsuccessful because of the high background obtained with these probes. Therefore, we synthesized a single-stranded oligodeoxyribonucleotide probe that was specific for proenkephalin. The sequence that was chosen for synthesis (Fig. 1) was complementary to the portion of mRNA that coded for the central region of peptide E, a potent opioid peptide that had been isolated from bovine adrenal medulla (25). Since this 30-base-long sequence was identical in both bovine and human preproenkephalin cDNA (3–5), it was thought that there would be a high probability that it would be identical in the rat as well.

Detection of Preproenkephalin mRNA in Brain. Blot-hybridization analysis of poly(A)-containing mRNA from rat brain indicated that preproenkephalin mRNA could be assayed by using the synthetic proenkephalin-specific ^{32}P-labeled probe (Fig. 2). Under these conditions, preproenkephalin mRNA could be detected in 5 μg of poly(A)-containing mRNA isolated from rat brain. As expected, poly(A)-containing mRNA isolated from rat liver did not contain detectable amounts of preproenkephalin mRNA. Rat preproenkephalin mRNA was approximately the same size (about 1400 nucleotides) as bovine preproenkephalin mRNA, based on its migration relative to the rRNA markers (Fig. 2) and from blot-hybridization analysis of the two species together (data not shown).

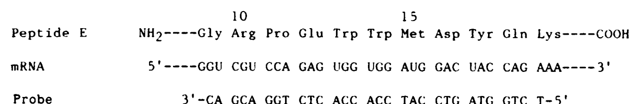
10 15
Peptide E NH2----Gly Arg Pro Glu Trp Trp Met Asp Tyr Gln Lys----COOH
mRNA 5'----GGU CGU CCA GAG UGG UGG AUG GAC UAC CAG AAA----3'
Probe 3'-CA GCA GGT CTC ACC ACC TAC CTG ATG GTC T-5'

FIG. 1. Sequence of the synthetic oligodeoxyribonucleotide used for probing cloned cDNA for preproenkephalin. The partial amino acid sequence of peptide E and the corresponding codons that were shown to be present in both bovine and human preproenkephalin mRNA (3–5) are shown on top. Below is the complementary sequence of the 30-base-long oligodeoxyribonucleotide probe that was synthesized for use as a probe. Numbers above the amino acids refer to their positions within the sequence of peptide E, which contains 25 amino acids.

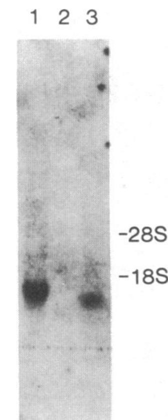

FIG. 2. Hybridization of the synthetic oligodeoxyribonucleotide probe to poly(A)-containing RNA from rat brain and liver. RNA samples were subjected to electrophoresis on 1% agarose gels containing 10 mM CH_3HgOH and then transferred overnight onto nitrocellulose sheets with 3 M sodium acetate buffer (pH 7.0). The nitrocellulose blot was baked, prehybridized, then hybridized as described (12) at 30°C with 3.7×10^7 cpm of the synthetic probe (kinase-labeled to a specific activity of 5×10^8 cpm/μg). Washes were 2× NaCl/Cit (1× NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate, pH 7) containing 0.1% $NaDodSO_4$ followed by 0.1× NaCl/Cit containing 0.1% $NaDodSO_4$ at room temperature. Autoradiography was for 27 hr at −70°C. The positions of rat 28S and 18S rRNA, visualized by ethidium bromide staining of the gel after transfer, are indicated. Lanes: 1, 11.3 μg of brain poly(A)-containing RNA: 2, 11.2 μg of liver poly(A)-containing RNA; 3, 5.6 μg of brain poly(A)-containing RNA.

Molecular Cloning of Preproenkephalin cDNA. An enriched fraction containing preproenkephalin mRNA from sucrose gradient centrifugation of rat brain poly(A)-containing mRNA was used to direct the synthesis of double-stranded cDNA. The yield was 1.1 μg of cDNA from 11 μg of mRNA. The terminal deoxynucleotidyl transferase reaction was optimized to produce poly(dG)·poly(dC) tails of 20–30 bases. This tail length has been shown previously to be the most efficient for transforming *E. coli* RR1 cells (18). Approximately 110,000 transformants were generated from 0.3 μg of cDNA. Sixty-three ampicillin-resistant colonies were tested for tetracycline sensitivity, and not a single colony grew on tetracycline, indicating that the clone bank contains very little native circular pBR322. About 28,000 bacterial clones were grown on eight 130-mm Petri dishes, transferred to nitrocellulose, and then screened by *in situ* hybridization. Ten positive clones were isolated, and the one with the longest cDNA insert was selected for DNA sequencing.

Sequence Determination of Rat Preproenkephalin cDNA. The plasmid pRPE-1 contained the largest cDNA, which was approximately 1200 bases long and included the entire coding region for rat preproenkephalin (Fig. 3). The nucleotide sequence was determined from both strands of the cDNA. The initiation site for translation has been tentatively assigned to the methionine codon at positions 1–3, based on the amino acid sequences of bovine and human preproenkephalin (Fig. 4). Rat preproenkephalin would consist then of 269 amino acids with a molecular weight of 30,936. This compares to 267 residues for human and 263 residues for bovine preproenkephalin. The amino terminus of rat preproenkephalin is most likely the aspartic acid residue (coded for by GAC) at position 24 of the amino acid sequence by analogy to the bovine and human proteins. In those species the amino-terminal amino acid is glutamic acid, which is coded for

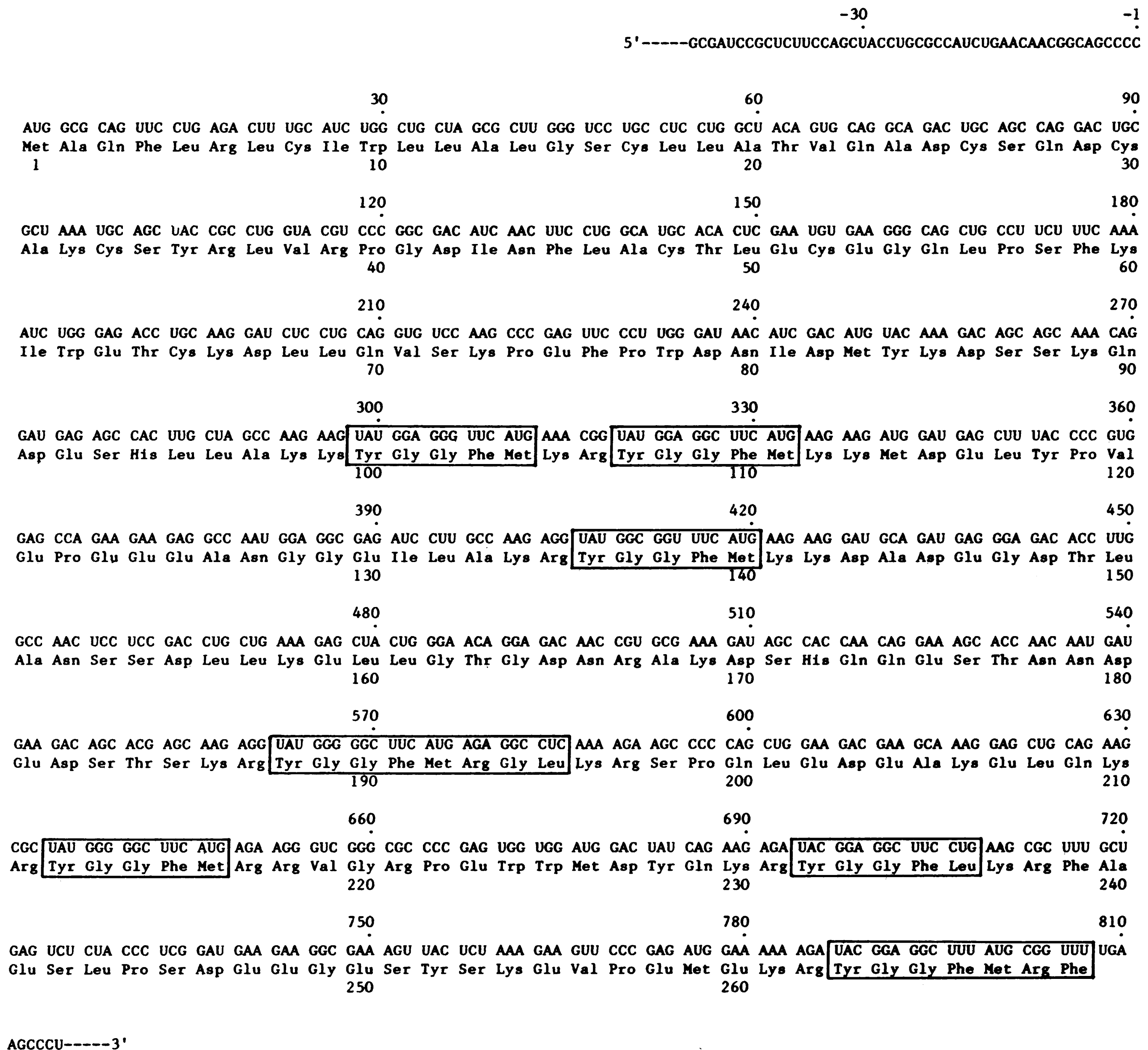

FIG. 3. Partial nucleotide sequence of rat preproenkephalin mRNA. Nucleotides are numbered in the 5'-to-3' direction beginning with the first residue of the initiator codon AUG; nucleotides preceding the AUG codon are given negative numbers. The sequence is incomplete at both the 5' and 3' termini of the mRNA. The predicted amino acid sequence is displayed below the mRNA. Amino acids are numbered beginning with the methionine residue coded for by nucleotides 1–3. The sequences of [Met]enkephalin, [Leu]enkephalin, [Met]enkephalin-Arg^6-Gly^7-Leu^8 and [Met]enkephalin-Arg^6-Phe^7 are boxed.

by GAA (Fig. 4). The signal peptide then would contain 24 largely hydrophobic amino acids and terminate in alanine. This is typical of signal peptides (26). The rat signal peptide contains two cysteine residues at positions 8 and 17, whereas the human and bovine forms contain only a single cysteine at position 8. All three species contain six additional cysteine residues (excluding those in the signal peptide), which are located in the amino-terminal region of proenkephalin; all occur at exactly the same positions (amino acid residues 26, 30, 33, 48, 52, and 65). The amino acid sequence of rat proenkephalin contains, like bovine and human proenkephalin, four copies of [Met]enkephalin and one copy each of [Leu]enkephalin, [Met]enkephalin-Arg^6-Gly^7-Leu^8 and [Met]enkephalin-Arg^6-Phe^7. Each of these sequences is bounded by dibasic amino acids, except the heptapeptide, which is followed by the termination codon UGA. It should be noted that the sequence of the highly active peptide E is entirely conserved in all three species, even though there are a significant number of single-base changes (compare Fig. 3 and refs. 3–5). The amino acid sequence of rat preproenkephalin is 82% homologous to both bovine and human preproenkephalin, allowing for the additional amino acids at positions 170 and 182 in the rat molecule relative to the human and bovine molecules and at positions 85–87 and 183 relative to the bovine form. Rat preproenkephalin cDNA is 80% and 83% homologous to the bovine and human cDNAs, respectively, at the nucleotide level.

The ability of the rat, bovine, and human preproenkephalin cDNAs to detect preproenkephalin mRNA in extracts of rat brain was compared (Fig. 5) in a dot-blot assay. When the band intensities were compared by scanning laser densitometry, the rat cDNA, pRPE-1-(165–600), was shown to be 4

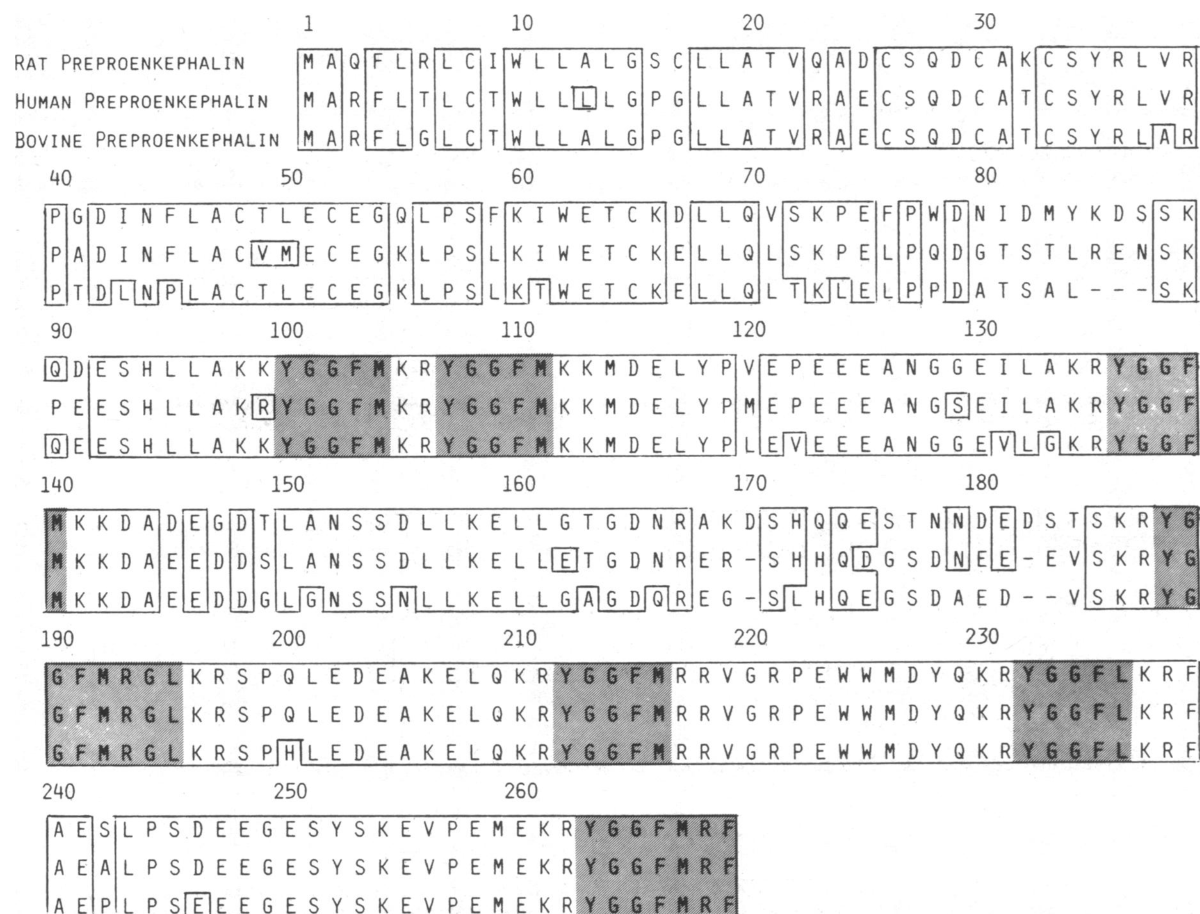

FIG. 4. Comparative amino acid sequences of rat, human, and bovine preproenkephalin (top, middle, and bottom lines, respectively, of each row). The human and bovine proteins are compared with rat preproenkephalin; regions of homology with the rat sequence are boxed. Deletions in the human and bovine sequences are denoted with a line under the corresponding rat amino acid. Numbering is as in Fig. 3. The sequences of [Met]enkephalin, [Leu]enkephalin, [Met]enkephalin-Arg[6]-Gly[7]-Leu[8] and [Met]enkephalin-Arg[6]-Phe[7] are shaded.

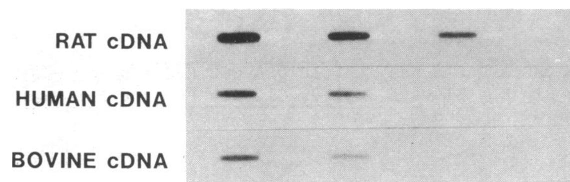

FIG. 5. Comparative hybridization of different preproenkephalin cDNAs to rat brain poly(A)-containing RNA. Rat, human, and bovine cDNAs (250 ng each) were labeled by nick-translation to a specific activity of approximately 10^8 cpm/μg. Rat brain poly(A)-containing RNA was blotted onto nitrocellulose in 1:2 serial dilutions (blots: left, 6 μg; middle, 2 μg; right, 0.67 μg) in triplicate, and each replicate was hybridized with one of the preproenkephalin cDNAs (3×10^6 cpm) as described. Autoradiograms of the washed filters were developed after overnight exposure at −70°C, and relative band intensities were measured by scanning laser densitometry.

times more sensitive as a probe for rat preproenkephalin mRNA than were the heterologous cDNAs. Moreover, only rat cDNA was able to consistently detect the message in normal rat adrenal glands. This laboratory has used the pRPE-1-(165–600) probe to determine the magnitude of the increase in preproenkephalin mRNA following denervation of the rat adrenal gland that we previously described (27).

1. Hughes, J., Smith, T. W., Kosterlitz, H. W., Fothergill, L. A., Morgan, B. A. & Morris, H. R. (1975) *Nature (London)* **258,** 577–579.
2. Udenfriend, S. & Kilpatrick, D. L. (1983) *Arch. Biochem. Biophys.* **221,** 309–323.
3. Gubler, U., Seeburg, P. H., Gage, L. P. & Udenfriend, S. (1982) *Nature (London)* **295,** 206–208.
4. Noda, M., Furutani, Y., Takahashi, H., Toyosato, M., Hirose, T., Inayama, S., Nakanishi, S. & Numa, S. (1982) *Nature (London)* **295,** 202–206.
5. Comb, M., Seeburg, P. H., Adelman, J., Eiden, L. & Herbert, E. (1982) *Nature (London)* **295,** 663–666.
6. Lewis, R. V., Stern, A. S., Kilpatrick, D. L., Gerber, L. D., Rossier, J., Stein, S. & Udenfriend, S. (1981) *J. Neurosci.* **1,** 80–82.
7. Fleminger, G., Lahm, H.-W. & Udenfriend, S. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 3587–3590.
8. Fleminger, G., Howells, R. D., Kilpatrick, D. L. & Udenfriend, S. (1984) *Proc. Natl. Acad. Sci. USA*, in press.
9. Yang, H.-Y., Hong, J. S. & Costa, E. (1977) *Neuropharmacology* **16,** 303–307.
10. Kobayashi, R. M., Palkovits, M., Miller, R. J., Chang, K.-J. & Cuatrecasas, P. (1978) *Life Sci.* **22,** 527–530.
11. Hall, L., Craig, R. K. & Campbell, P. N. (1979) *Nature (London)* **277,** 54–56.
12. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
13. Bailey, J. M. & Davidson, N. (1976) *Anal. Biochem.* **70,** 75–85.
14. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
15. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113,** 237–251.
16. Gubler, U. & Hoffmann, B. J. (1983) *Gene* **25,** 263–269.
17. Michelson, A. M. & Orkin, S. H. (1982) *J. Biol. Chem.* **257,** 14773–14782.
18. Peacock, S. L., McIver, C. M. & Monahan, J. J. (1981) *Biochim. Biophys. Acta* **655,** 243–250.

19. Hanahan, D. & Meselson, M. (1980) *Gene* **10,** 63–67.
20. Miyoshi, K.-J., Arentzen, R., Huang, T. & Itakura, K. (1980) *Nucleic Acids Res.* **8,** 5507–5517.
21. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
22. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513–1522.
23. Clewell, D. B. & Helinski, D. R. (1972) *J. Bacteriol.* **110,** 1135–1146.
24. White, B. A. & Bancroft, F. C. (1982) *J. Biol. Chem.* **257,** 8569–8572.
25. Kilpatrick, D. L., Taniguchi, T., Jones, B. N., Stern, A. S., Shively, J. E., Hullihan, J., Kimura, S., Stein, S. & Udenfriend, S. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 3265–3268.
26. Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. (1980) *Ann. N.Y. Acad. Sci.* **343,** 1–16.
27. Kilpatrick, D. L., Howells, R. D., Fleminger, G. & Udenfriend, S. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 7221–7223.