Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 19–23, January 1986
Biochemistry

Two prohormones for gastrin-releasing peptide are encoded by two mRNAs differing by 19 nucleotides

(bombesin/alternative splicing)

ELIOT R. SPINDEL*, MARYA D. ZILBERBERG*, JOEL F. HABENER†, AND WILLIAM W. CHIN*

*Section on Molecular Genetics, Department of Medicine, Brigham and Women's Hospital, 75 Francis St., Howard Hughes Medical Institute and Joslin Diabetes Center, Harvard Medical School, Boston, MA 02115; and †Laboratory of Molecular Endocrinology, Massachusetts General Hospital and Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02114

Communicated by Francis O. Schmitt, August 19, 1985

ABSTRACT **In our studies on the molecular biology of human gastrin-releasing peptide (GRP), we have discovered an example of a change in translational reading frame apparently produced through alternative RNA splicing. Complementary DNAs prepared from a pulmonary carcinoid tumor rich in GRP immunoreactivity had one of two different-sized internal DNA fragments after digestion with the restriction enzyme *Pvu* II. Nucleotide sequences of the two DNA fragments were identical except for 19 additional nucleotides present in the larger fragment. The region of the mRNA containing the 19 nucleotides corresponded to the carboxyl-terminal region of the human GRP precursor. The resulting shift in reading frame causes a difference of 10 amino acids in size and an overall sequence difference of 27 amino acids between the two GRP prohormones so formed. The change in reading frame described here is unusual in eukaryotes and is yet another mechanism to produce diversity in the generation of biological peptides.**

Alternative RNA processing is an important mechanism for tissue-specific gene expression (1–3). Such processing was observed in viral systems (2) and more recently in eukaryotes (1, 3). Understanding the variety of ways in which RNA can be spliced appears to be especially important in characterizing the expression of the genes encoding neuropeptides. For example, alternative splicing is responsible for the tissue-specific expression of calcitonin or calcitonin gene-related peptide (3) or for the variable expression of substance P and substance K (4). In this paper we report that alternative RNA splicing is apparently responsible for the production of two prohormones that encode the neuropeptide gastrin-releasing peptide (GRP).

GRP, the 27-residue mammalian homolog of the amphibian tetradecapeptide bombesin (5), is widely distributed throughout the nervous system, gastrointestinal tract, and pulmonary tract (6–8). Within the central nervous system, GRP causes hypothermia and hyperglycemia and stimulates sympathetic outflow (9). Within the gastrointestinal tract GRP potently releases almost every known gastrointestinal hormone (10). In the lung GRP is present in pulmonary neuroendocrine cells and in small cell carcinomas (11), where it appears to be a significant autocrine growth factor (12, 13).

Previously, we have characterized cDNAs from a pulmonary carcinoid tumor rich in GRP immunoreactivity and have demonstrated that GRP is encoded in a typical polypeptide precursor consisting of a signal peptide, GRP itself, and a carboxyl-terminal extension peptide (14). Genomic DNA blot analysis is consistent with a single gene encoding human GRP (hGRP) (14).

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.

METHODS

A cDNA library was prepared from polyadenylylated RNA obtained from a pulmonary carcinoid tumor rich in GRP immunoreactivity as described previously (14). Colony screening (15) was performed using a previously isolated GRP cDNA (14) labeled by nick-translation (16). DNA was sequenced primarily by the chemical cleavage procedure (17) with some confirmatory sequencing using the phage M13–dideoxy procedure (18, 19).

Thirty-two-base oligonucleotides were synthesized with the Applied Biosystems DNA synthesizer and purified by ion-exchange HPLC (20). Tumor RNA was fractionated on formaldehyde/agarose gels (21), transferred to nitrocellulose (22), and hybridized to 5′-end-labeled oligonucleotides (17) as described below.

DNA and protein database analyses were conducted through Genbank and Bionet and with the Beckman Microgenie sequence analysis software.

RESULTS

Screening of 50,000 cDNA clones resulted in the isolation of 20 GRP-encoding cDNAs. Restriction enzyme analysis revealed two classes of cDNAs characterized by an internal *Pvu* II fragment of either, approximately 200 or 220 bases (Fig. 1). DNA sequence analysis revealed that the internal *Pvu* II fragments were identical except for 19 extra bases present in the longer form (Fig. 2). The sequences of the two fragments preceding and following this 19-nucleotide insertion were otherwise identical. Further restriction enzyme and sequence analyses showed that the entire clones were otherwise identical. The sequence in the area of the insertion was determined both chemically and enzymatically and on multiple plasmids. Of the 20 cDNA clones isolated, 11 contained the internal *Pvu* II fragment. Of those 11, 6 contained the insert and 5 did not.

Inasmuch as the insert occurs within the protein-encoding region of pre-pro-GRP and 19 is not a multiple of 3, a shift of translational reading frame results. The predicted effect of the insert-induced frameshift on the proteins resulting from the translation of the mRNAs is shown in Fig. 3. Translation of the GRP mRNA without the insert reaches a stop codon after amino acid 138. Translation of the RNA with the insert terminates after amino acid 148 with three successive stop codons. In all, the 19-nucleotide insertion causes the insert-positive prohormone to contain 27 amino acids not found in the insert-negative prohormone. The deduced molecular weights for the two prohormones so formed are ≈15,000 and 16,000.

Blot analysis of the GRP mRNAs was performed with two 32-base oligonucleotides designed to differentially hybridize

Abbreviations: GRP, gastrin-releasing peptide; hGRP, human GRP.

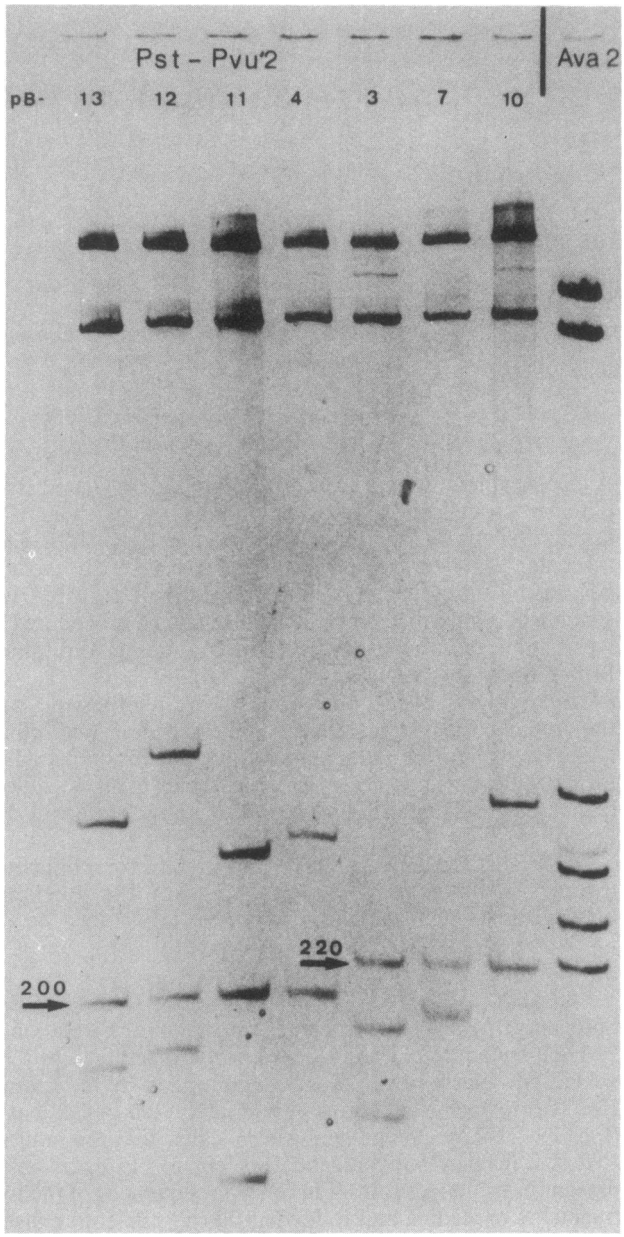


FIG. 1. *Pvu* II/*Pst* I restriction enzyme digests of GRP-encoding cDNAs. The two different-sized internal *Pvu* II fragments are indicated by arrows. cDNAs were inserted into the *Pst* I site of pBR322 as described previously (14). DNA was digested to completion with *Pst* I and *Pvu* II, separated on 5% polyacrylamide, and stained with ethidium bromide. Size markers (nucleotides) are *Ava* II-digested pBR322.

to either insert-positive or insert-negative mRNAs. These were synthesized as indicated by the underlined sequences in Fig. 2 *Lower*. Hybridization conditions that allowed the desired differential hybridization were determined by preliminary dot blot analysis (data not shown) of the oligonucleotide probes against insert-positive and insert-negative cDNA clones. Under conditions as described in the legend of Fig. 4, the probes were specific for either the presence or the absence of the 19 nucleotides and would not hybridize to microgram quantities of the inappropriate cDNA. Under these conditions, RNA blot analysis showed that both oligonucleotides hybridized to GRP mRNAs (Fig. 4), though no differences in the size of the main hybridizing bands were seen because a difference of 19 nucleotides out of 950 is below the resolving power of the agarose gel. Significantly, only the oligonucleotide complementary to the insert hybridized to high molecular weight RNA forms (Fig. 4). No large RNA forms lacking the insert were detected.

DISCUSSION

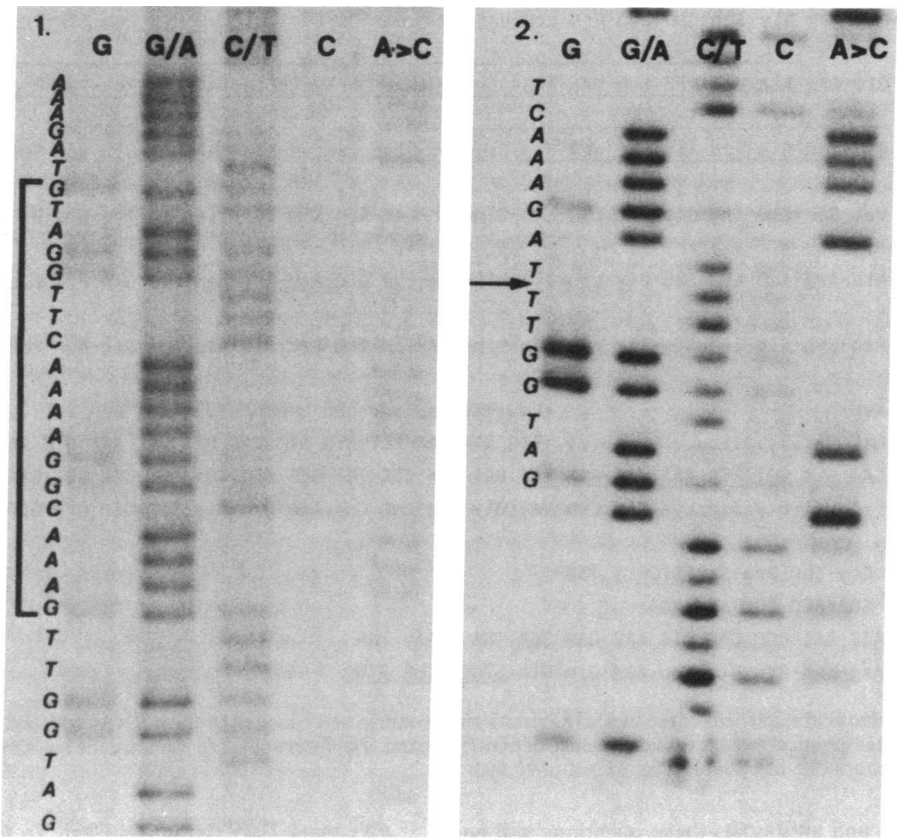
Restriction enzyme and sequence analyses of cDNA clones indicate that there are two species of RNA present in a pulmonary carcinoid tumor. These RNAs differ by the presence or absence of 19 nucleotides. The near-equal distribution of insert-positive and insert-negative cDNAs suggests that the RNAs occur with similar frequencies.

The 19-nucleotide insertion/deletion occurs in the region of the GRP carboxyl-terminal extension peptide and causes a shift in protein reading frame. The 19-nucleotide change causes a 10 amino acid difference in size and a 27 amino acid difference between the sequences of the two prohormones. Preliminary cell-free translation studies detect two immunoprecipitable bands consistent with the predicted molecular weights for the two prohormones (data not shown). The changes in the GRP prohormones that result from the 19-nucleotide insertion do not create any obvious new neuropeptides as indicated by dibasic amino acid cleavage points. Formal search of a protein data base did not find any strong homologies with other known proteins, nor did comparison of the 19-nucleotide insert with a DNA data base find any strong homologies with known DNA sequences. Computer analysis also did not detect any sequences in the GRP mRNA itself likely to form a stem-loop structure in the region of the insert. Hence, the insert is unlikely to be a cloning artifact caused by mRNA secondary structure and incorrect reverse transcription.

The 19-nucleotide insert is bounded by the dinucleotides GT and AG, the mandatory intron donor and acceptor consensus sequences (23). In addition, five out of the first six nucleotides of the 5′ end of the insert match the intron donor consensus sequence (23). The 3′ end of the insert matches the intron consensus acceptor sequence only with the mandatory AG.

Blot analysis performed with oligonucleotides synthesized to hybridize respectively (under appropriately stringent conditions) to either insert-positive or insert-negative RNA demonstrated that the two forms exist in roughly similar amounts (Fig. 4). Only the probe designed to hybridize to insert-positive mRNA hybridized to high molecular weight RNA (Fig. 4). If the two GRP mRNAs originate from two separate genes, with and without the 19 nucleotides, then both oligonucleotides should have hybridized to precursor forms. The fact that no high molecular weight RNAs hybridized to the insert-negative probe suggests that the 19-nucleotide insert is removed during RNA processing and that the two RNA forms arise from alternative splicing of a single gene transcript. This is in agreement with genomic DNA blot analyses (14).

The putative splice could involve an alternative donor site or an alternative acceptor site, or the insert itself could be a small exon or an unspliced mini-intron. The GT and AG dinucleotides bordering the insert give it resemblance to a small intron; if so, then the GRP gene would be processed analogously to the γ-fibrinogen gene, in which the final intron is not removed from the mature mRNA 10% of the time (24). If the insert is actually a small exon, the situation would be analogous to the αA-crystallin gene, in which a small exon, lying within an intron, is alternatively spliced (25). The most common alternative RNA processing involves the use of an alternative acceptor site, as occurs with mRNAs for rat fibronectin (26), rat proopiomelanocortin (27), and human growth hormone (28). This alternative splicing pattern appears unlikely to occur with the GRP transcript because the 3′ end of the 19-nucleotide insert has only weak homology




```
     263
I:   CAGCTGAGAG AGTACATCAG GTGGGAAGAA GCTGCAAGGA ATTTGCTGGG TCTCATAGAA GCAAAGGAGA ACAGAAACCA CCAGCCACCT
II:  CAGCTGAGAG AGTACATCAG GTGGGAAGAA GCTGCAAGGA ATTTGCTGGG TCTCATAGAA GCAAAGGAGA ACAGAAACCA CCAGCCACCT
     353
I:   CAACCCAAGG CCTTGGGCAA TCAGCAGCCT TCGTGGGATT CAGAGGATAG CAGCAACTTC AAAGAT.....................TTGGT
II:  CAACCCAAGG CCTTGGGCAA TCAGCAGCCT TCGTGGGATT CAGAGGATAG CAGCAACTTC AAAGATGTAG GTTCAAAAGG CAAAGTTGGT
                                                                        |---- 19 bases -----|
     443                                          485
I:   AGACTCTCTG CTCCAGGTTC TCAACGTGAA GGAAGGAACC CCC
II:  AGACTCTCTG CTCCAGGTTC TCAACGTGAA GGAAGGAACC CCC
```

FIG. 2. Sequences of *Pvu* II fragments. (*Upper*) Autoradiogram showing the difference in sequence of the *Pvu* II fragments. Gel 1, sequence of the DNA fragment containing the 19-nucleotide insert (indicated by bracket). Gel 2, sequence of the DNA fragment without the insert. Arrow indicates where the insert would be. The internal *Pvu* II fragments were 3′-end labeled with [^{32}P]cordycepin, and the strands were separated and sequenced according to Maxam and Gilbert (17). Labels atop the lanes indicate the base cleavage reaction used. (*Lower*) Complete sequences for the internal *Pvu* II fragments. The sequence of each fragment was confirmed on at least two plasmids. Fragment II contains 19 extra nucleotides, but the fragments are otherwise identical. The underlined sequences indicate the complements of the two 32-base synthetic oligonucleotides used in Fig. 4. Compared to our previously published sequence (14), nucleotide 456 has been corrected from a T to a C.

with the consensus sequence of the intron acceptor site. The 5′ end of the insert, however, displays strong homology with the intron donor site consensus sequence, a situation that suggests the use of an alternative donor site such as occurs with the immunoglobulin μ gene (29), in some forms of β-thalassemia (30), and with the simian virus 40 big and little T (tumor) antigens (31).

Though homology with consensus sequences makes an alternative donor site splice most likely, the exact mechanism will be confirmed only when the sequence of the hGRP gene is determined. While alternative splicing of the type described here is not unusual in itself, it is unusual to have the splice shift the reading frame such that the small insertion causes a relatively large difference between the two prohormones. This is thus an additional way for alternative splicing to cause diversity in biologically active peptides.

The physiologic significance of this alternative splicing is dependent upon the answers to a number of questions. First, does this splice occur in all tissues that express GRP? One suggestion that it may occur commonly is the observation of the identical splice in several GRP-producing small cell carcinomas and cell lines (34). The extent to which normal lung, gut, and brain tissues will also exhibit this splicing remains to be determined. The extent to which there is regulation or tissue specificity of this phenomenon also must be determined. As yet the function of the GRP carboxyl-terminal extension peptide is unknown. It will be important to see if its structure is conserved among other species when other GRP-encoding mRNAs are cloned. The alternative splice would be expected to alter any biological function the carboxyl-terminal extension peptide might have. Another possibility is that the alternative splice could modulate the processing of pre-pro-GRP to either GRP or its derivative, the GRP carboxyl-terminal decapeptide, GRP-10 (32, 33).

```
I:   Met-Arg-Gly-Ser-Glu-Leu-Pro-Leu-Val-Leu-Leu-Ala-Leu-Val-Leu-Cys-Leu-Ala-Pro-Arg-  20
II:   -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
I:   Gly-Arg-Ala-Val-Pro-Leu-Pro-Ala-Gly-Gly-Gly-Thr-Val-Leu-Thr-Lys-Met-Tyr-Pro-Arg-  40
II:   -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
I:   Gly-Asn-His-Trp-Ala-Val-Gly-His-Leu-Met-Gly-Lys-Lys-Ser-Thr-Gly-Glu-Ser-Ser-Ser-  60
II:   -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
I:   Val-Ser-Glu-Arg-Gly-Ser-Leu-Lys-Gln-Gln-Leu-Arg-Glu-Tyr-Ile-Arg-Trp-Glu-Glu-Ala-  80
II:   -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
I:   Ala-Arg-Asn-Leu-Leu-Gly-Leu-Ile-Glu-Ala-Lys-Glu-Asn-Arg-Asn-His-Gln-Pro-Pro-Gln-  100
II:   -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
I:   Pro-Lys-Ala-Leu-Gly-Asn-Gln-Gln-Pro-Ser-Trp-Asp-Ser-Glu-Asp-Ser-Ser-Asn-Phe-Lys-  120
II:   -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -
I:   Asp-                          -Leu-Val-Asp-Ser-Leu-Leu-Gln-Val-Leu-Asn-Val-Lys-Glu-  134
     GAT ..........................  TTG GTA GAC TCT CTG CTC CAG GTT CTC AAC GTG AAG GAA
II:  GAT GTA GGT TCA AAA GGC AAA GTT GGT AGA CTC TCT GCT CCA GGT TCT CAA CGT GAA GGA
      -  Val-Gly-Ser-Lys-Gly-Lys-Val-Gly-Arg-Leu-Ser-Ala-Pro-Gly-Ser-Gln-Arg-Glu-Gly-  140

     Gly-Thr-Pro-Ser STOP   138
I:   GGA ACC CCC AGC TGA
     AGG AAC CCC CAG CTG AAC CAG CAA TGA  TAA  TGA
II:  Arg-Asn-Pro-Gln-Leu-Asn-Gln-Gln STOP STOP STOP   148
```

FIG. 3. Deduced amino acid sequences of the two hGRP-containing prohormones encoded by the mRNAs with and without the 19-nucleotide insert. A hyphen indicates identical residues. The sequence of hGRP is boxed. The insert-positive prohormone has a molecular weight of 16,000. The insert-negative prohormone has a molecular weight of 15,000.

With specific probes and antibodies these questions will be readily answered.

Note Added in Proof. Sequence analysis of a human genomic clone encoding GRP has demonstrated a large intron immediately following the 19-nucleotide insert. Thus the insert most likely arises by an alternative donor site splice with the second donor site provided by the 5′ end of the intron (unpublished data).

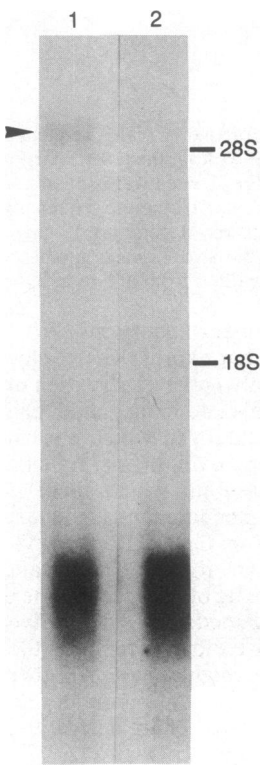


FIG. 4. RNA blot analysis of pulmonary carcinoid tumor RNA. Two micrograms of polyadenylylated RNA was fractionated on a 1.5% formaldehyde/agarose gel and transferred to nitrocellulose. Lane 1 was probed with the 32-base oligonucleotide (see legend of Fig. 2 *Lower*) designed to hybridize to RNA with the insert (insert-positive probe). Arrowhead points to high molecular weight RNA that hybridized. Lane 2 was probed with the 32-base oligonucleotide (see legend of Fig. 2 *Lower*) designed to hybridize to RNA without the insert (insert-negative probe). There were no other hybridizing bands visible below the major hybridizing bands. Hybridization was at 35°C in 50% (vol/vol) formamide/0.75 M NaCl/0.075 M sodium citrate/0.1% sodium dodecyl sulfate/0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone containing salmon sperm DNA at 100 μg/ml. Positions of rRNAs are indicated on the right.

We thank Gary Gryan and Chon Vo for their assistance in the synthesis of oligonucleotides. Bombesin antiserum was generously provided by L. H. Rees and J. Price. We especially thank Anne-Marie Lebacq-Verheyden for confirmatory DNA sequence analysis. This work was supported in part by National Institutes of Health Grant 1R23-CA-39237-01.

1. Douglass, J., Civelli, O. & Herbert, E. (1984) *Annu. Rev. Biochem.* **53**, 665–715.
2. Ziff, E. B. (1980) *Nature (London)* **287**, 491–499.
3. Rosenfeld, M. G., Mermod, J., Amara, S. G., Swanson, L. W., Sawchenko, P. E., Rivier, J., Vale, W. W. & Evans, R. M. (1983) *Nature (London)* **304**, 129–135.
4. Nawa, H., Hirose, T., Takashima, H., Inayama, S. & Nakanishi, S. (1983) *Nature (London)* **306**, 32–36.
5. McDonald, T. J., Jornvall, H., Nilsson, G., Vagne, M., Ghatei, M., Bloom, S. R. & Mutt, V. (1979) *Biochem. Biophys. Res. Commun.* **90**, 227–233.
6. Villarreal, J. A. & Brown, M. R. (1978) *Life Sci.* **23**, 2729–2734.
7. Walsh, J. H., Wong, H. C. & Dockray, G. J. (1979) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **38**, 2315–2319.
8. Wharton, J., Polak, J. M., Bloom, S. R., Ghatei, M. A., Solcia, E., Brown, M. R. & Pearse, A. G. E. (1978) *Nature (London)* **273**, 769–770.
9. Tache, Y. & Brown, M. (1982) *Trends Neurosci.* **5**, 431–433.
10. McDonald, T. J., Ghatei, M. A., Bloom, S. R., Adrian, T. E., Mochizuki, T., Yanaihara, C. & Yanaihara, N. (1983) *Regul. Pept.* **5**, 125–137.
11. Moody, T. W., Pert, C. B., Gazdar, A. F., Carney, D. N. & Minna, J. D. (1981) *Science* **214**, 1246–1248.
12. Weber, S., Zuckerman, J. E., Bostwick, D. G., Bensch, K. G., Sikic, B. I. & Raffin, T. A. (1985) *J. Clin. Invest.* **75**, 306–309.
13. Cuttitta, F., Carney, D. N., Mulshine, J., Moody, T. W., Fedorko, J., Fischler, A. & Minna, J. D. (1985) *Nature (London)* **316**, 823–826.
14. Spindel, E. R., Chin, W. W., Price, J., Rees, L. H., Besser, G. M. & Habener, J. F. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5699–5703.
15. Hanahan, D. & Meselson, M. (1983) *Methods Enzymol.* **100**, 331–341.

16. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113,** 237–251.
17. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
19. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9,** 309–321.
20. Ito, H., Ike, Y., Ikuta, S. & Itakura, K. (1982) *Nucleic Acids Res.* **10,** 1755–1769.
21. Lehrach, H., Diamond, D. & Wozney, J. M. (1977) *Biochemistry* **16,** 4743–4751.
22. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
23. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472.
24. Crabtree, G. R. & Kant, J. A. (1982) *Cell* **31,** 159–166.
25. King, C. R. & Piatigorsky, J. (1983) *Cell* **32,** 707–712.
26. Tamkun, J. W., Schwarzbauer, J. E. & Hynes, R. O. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 5140–5144.
27. Oates, E. & Herbert, E. (1984) *J. Biol. Chem.* **259,** 7421–7425.
28. DeNoto, F. M., Moore, D. D. & Goodman, H. M. (1981) *Nucleic Acids Res.* **9,** 3719–3730.
29. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* **20,** 313–319.
30. Treisman, R., Orkin, S. H. & Maniatis, T. (1983) *Nature (London)* **302,** 591–596.
31. Lebowitz, P. & Weissman, S. M. (1979) *Curr. Top. Microbiol. Immunol.* **87,** 42–172.
32. Reeve, J. R., Jr., Walsh, J. H., Chew, P., Clark, B., Hawke, D. & Shively, J. E. (1983) *J. Biol. Chem.* **258,** 5582–5588.
33. Orloff, M. S., Reeve, J. R., Jr., Ben-Avram, C. M., Shively, J. E. & Walsh, J. H. (1984) *Peptides* **5,** 865–870.
34. Sausville, E. A., Lebacq-Verheyden, A., Spindel, E. R., Cuttitta, F., Gazdar, A. F. & Battey, J. F., *J. Biol. Chem.*, in press.