

## Cloning and characterization of cDNAs encoding human gastrin-releasing peptide

(bombesin/mRNA/neuropeptide)

ELIOT R. SPINDEL\*, WILLIAM W. CHIN\*, JANET PRICE†, LESLEY H. REES†, GORDON M. BESSER‡, AND JOEL F. HABENER\*

\*Laboratory of Molecular Endocrinology, Massachusetts General Hospital and Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02114; and Departments of †Chemical Endocrinology and ‡Endocrinology, St. Bartholomew's Hospital, London, England

Communicated by Herman M. Kalckar, June 6, 1984

**ABSTRACT** We have prepared and cloned cDNAs derived from poly(A)<sup>+</sup> RNA from a human pulmonary carcinoid tumor rich in immunoreactivity to gastrin-releasing peptide, a peptide closely related in structure to amphibian bombesin. Mixtures of synthetic oligodeoxyribonucleotides corresponding to amphibian bombesin were used as hybridization probes to screen a cDNA library prepared from the tumor RNA. Sequencing of the recombinant plasmids shows that human gastrin-releasing peptide (hGRP) mRNA encodes a precursor of 148 amino acids containing a typical signal sequence, hGRP consisting of 27 or 28 amino acids, and a carboxyl-terminal extension peptide. hGRP is flanked at its carboxyl terminus by two basic amino acids, following a glycine used for amidation of the carboxyl-terminal methionine. RNA blot analyses of tumor RNA show a major mRNA of 900 bases and a minor mRNA of 850 bases. Blot hybridization analyses using human genomic DNA are consistent with a single hGRP-encoding gene. The presence of two mRNAs encoding the hGRP precursor protein in the face of a single hGRP gene raises the possibility of alternative processing of the single RNA transcript.

Gastrin-releasing peptide (GRP) is a mammalian equivalent of the amphibian tetradecapeptide bombesin (1). GRP was originally isolated from porcine stomach by using the release of gastrin as a bioassay (2). Peptides related in structure have been isolated from avian proventriculus (3) and canine small intestine (4). Bombesin and the GRPs share similarities in their sequences and contain a common carboxyl-terminal heptapeptide.

GRP and bombesin have similar biological effects (5). Infusion of nanogram amounts of either peptide to dogs or humans increases plasma immunoreactive levels of gastrin, pancreatic polypeptide, glucagon, gastric inhibitory peptide, and insulin (5, 6). Bombesin is also a potent neuropeptide; intraventricular injections of 1 pM bombesin into rats elevates blood sugar (7) and produces hypothermia (8). These potent effects coupled with the wide distribution of bombesin-like peptides throughout the mammalian gastrointestinal tract (9) and nervous system (10) support the role of bombesin as an important neuroregulatory peptide.

Bombesin-like immunoreactivity is also detected in the neuroendocrine cells of the lung (11). In humans, the highest levels are found in the lung just after birth (12) and levels then decrease in parallel with the observed decrease in the number of pulmonary neuroendocrine cells (13). Bombesin immunoreactivity is also found in high levels in pulmonary carcinoid tumors as well as in many small cell carcinomas of the lung (14, 15). Hence, bombesin-like immunoreactivity is a potential chemical marker for these tumors.

As is characteristic of many neuropeptides, mammalian bombesin-like peptides exist in multiple forms. Bombesin-related peptides of 10 and 27 amino acids have been isolated from porcine (16) and canine (4) tissues. Analyses by high-pressure liquid chromatography of fetal lung (17) and lung tumors (15, 18, 19) yield multiple peaks of immunoreactivity; one peak elutes near bombesin and one elutes near GRP. Multiple immunoreactive forms are also detected in gastrointestinal tissue (9, 20). In brain, immunohistochemistry reveals some distinct regions that react to bombesin antisera and other regions that react only with antisera to ranatensin (21), another amphibian peptide structurally related to bombesin.

To clarify these complexities of the bombesin-like family of peptides, we wish to analyze the gene(s) encoding members of this peptide family. As an initial step in these studies, we describe here the molecular cloning and characterization of cDNAs corresponding to mRNAs encoding the precursor of human GRP (hGRP; pre-proGRP).

### METHODS

Polyadenylated RNA was obtained from a hepatic metastasis of a pulmonary carcinoid tumor rich in bombesin-like immunoreactivity (§), by using the guanidine thiocyanate procedure (23) and oligo(dT) chromatography (24). Double-stranded cDNA was prepared enzymatically as described by Crabtree and Kant (25). A recombinant cDNA library was prepared using homopolymeric dC-dG extensions and *Pst* I cut pBR322 (New England Nuclear). *Escherichia coli* MC1061 were transformed by the calcium-shock procedure (26).

Oligodeoxynucleotide probes were synthesized by the phosphate triester method (27). The probes (Fig. 1) corresponded to two different regions of amphibian bombesin and ranatensin to allow double screening and thus decrease the detection of false-positive recombinants. Duplicate high- (28) and low-density (29) colony screening was carried out with 5'-end-labeled (30) probes using hybridization conditions as described by Hanahan and Meselson (28). Cell-free translations using wheat germ extract and reticulocyte lysate, plasmid preparations, RNA hybridization, DNA hybridization, and DNA chemical sequence analyses were carried out by standard techniques (31-37).

RNA blots were hybridized for 18 hr at 42°C in 50% formamide/5× NaCl/Cit (1× NaCl/Cit = 0.15 M NaCl/0.015 M Na citrate)/5× Denhardt's solution (1× Denhardt's solution = 0.02% Ficoll-400/0.02% bovine serum albumin/0.02% polyvinylpyrrolidone-40)/0.5% sodium dodecyl sulfate/sonicated denatured salmon sperm (100 µg/ml). Southern blots

Abbreviations: GRP, gastrin-releasing peptide; hGRP, human GRP; pre-proGRP, precursor of hGRP.

§Price, J., Nieuwenhuijsen-Kruseman, A. C., Clement-Jones, V. & Rees, L. H. (1983) Endocrine Society, 65th Annual Meeting, San Antonio, TX, June 8-10, 188 (abstr.).

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.

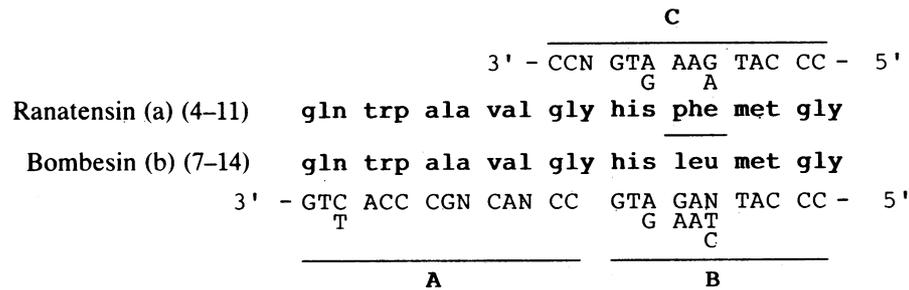


FIG. 1. Design of oligodeoxyribonucleotide probes. Probe A was prepared as two 16-fold mixed sets ( $A_1$  and  $A_2$ );  $A_1$  used the GTC and  $A_2$  used the GTT complement to the glutamine codon. Probe B was synthesized as a 32-fold mixed set. Probe C was made to ranatensin-(8-11), which is closely related to bombesin-(11-14), but requires only a 16-fold mixed synthesis. N indicates the inclusion of all 4 bases. Hybridization temperatures for probes A, B, and C were 41°C, 28°C, and 37°C, respectively. Sequences are ranatensin-(4-11) (a) and bombesin-(7-14) (b).

were hybridized for 36 hr at 45°C in the same hybridization solution.

## RESULTS

From 3 g of tumor, 100  $\mu$ g of poly(A)<sup>+</sup> RNA was prepared. Addition of 1  $\mu$ g of RNA to cell-free translation systems resulted in 10- to 30-fold increases in protein synthesis, thus demonstrating the integrity of the tumor mRNA. By using 10  $\mu$ g of poly(A)<sup>+</sup> RNA, 0.8  $\mu$ g of double-stranded cDNA was synthesized as described by Crabtree and Kant (25). Transformation of *E. coli* with 100 ng of cDNA produced 20,000 colonies. High- followed by low-density screening with probes  $A_2$  and C (Fig. 1) yielded 15 hybridizing colonies (probe C hybridized to essentially the same colonies as probe B but with less background). Sequence analyses of the recombinant plasmids obtained from these colonies (Fig. 2) yielded close to the entire sequences for the mRNA encoding pre-proGRP (Fig. 3). Both probes  $A_2$  and C hybridized appropriately despite single base mismatches.

The deduced amino acid sequence shows that hGRP is part of a 148-amino acid polypeptide that is delimited by an AUG start codon and three successive stop codons. The amino-terminal portion of this precursor contains a cluster of hydrophobic amino acids consistent with the "pre" or signal,

sequence (38). At the carboxyl terminus, hGRP is flanked by two lysines following a glycine, which likely provides the signal for amidation of the terminal methionine residue (39). Pre-proGRP does not code for any additional GRP-like or bombesin-like peptides, nor are there any other peptides set off by two consecutive basic amino acids characteristic of sites that are cleaved during the post-translational processing of prohormones. The protein-coding region is followed by 289 untranslated bases before the poly(A) tail commences; a typical A-A-A-T-A-A polyadenylation signal (40) occurs 19 bases before the poly(A) tail.

Hybridization analysis (RNA blot) demonstrates the presence of two mRNAs: a predominant mRNA of 900 bases and a minor RNA of approximately 850 bases (Fig. 4). Hybridization analysis (Southern blot) of genomic DNA from human lymphocytes shows only a single band with *Bam*HI, *Hind*III, and *Eco*RI digests, consistent with a single gene coding for pre-proGRP.

## DISCUSSION

Cloning of cDNA prepared from a pulmonary carcinoid tumor demonstrates that hGRP is encoded in a precursor protein. There is a hydrophobic signal sequence (38), although the exact cleavage point of the "pre" sequence is not deter-

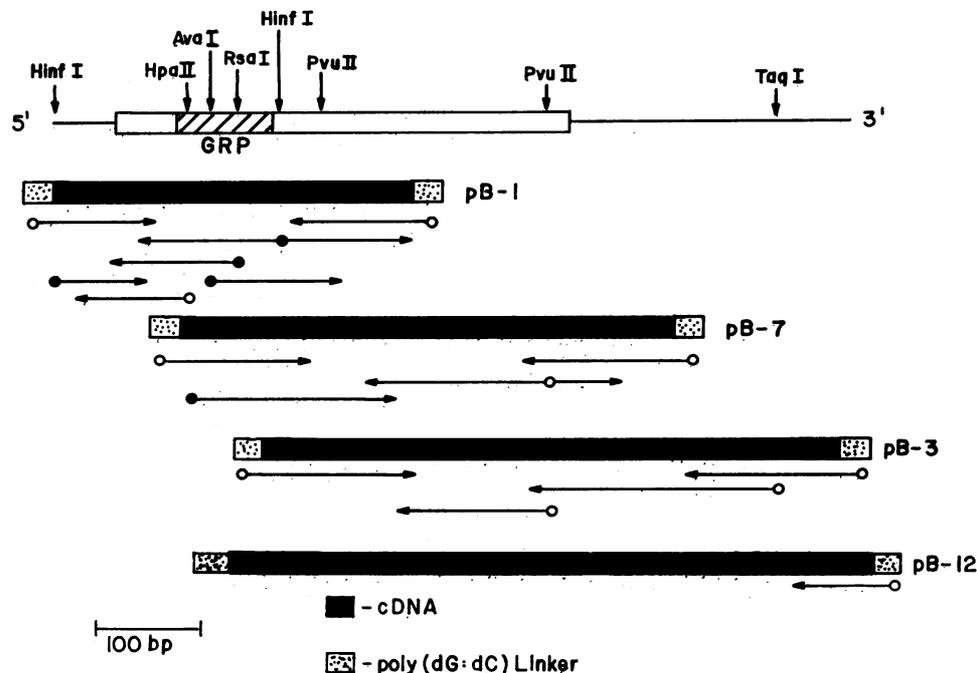


FIG. 2. DNA sequencing strategy for cDNAs encoding pre-proGRP. DNA fragments were produced by the indicated restriction enzymes and were either 5'- (●) or 3'- (○) end-labeled and sequenced by the method of Maxam and Gilbert (30). All DNA sequences were confirmed by sequencing on two different recombinant plasmids or on both strands of a single plasmid.



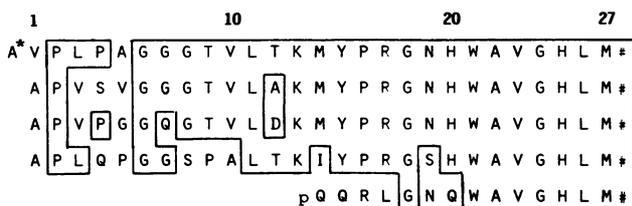


FIG. 5. Comparison of the amino acid sequence of hGRP with related peptides. Solid boxes enclose regions of homology with hGRP. M# indicates methioninamide. pQ indicates pyroglutamate. A\* indicates that hGRP could (though probably does not) have an additional amino-terminal alanine. Amino acids are designated by standard one-letter abbreviations. Sequences of the following peptides are shown (from top to bottom): hGRP, porcine GRP, canine GRP, avian GRP, amphibian bombesin.

tively, the smaller GRP-like peptide could be encoded in a second mRNA. Conceivably, the smaller GRP-like immunoreactive forms could be artifacts of peptide extraction, but the consistency with which these forms are observed and their variation in concentration in different tissues makes this circumstance unlikely (15, 17–20, 44).

RNA blot analysis shows a major mRNA of  $\approx 900$  bases and a minor mRNA of  $\approx 850$  bases (Fig. 4). The size of the major mRNA is consistent with the 800-base cDNA sequence shown in Fig. 3. Additional cloned recombinant plasmids are being screened to determine whether the smaller mRNA encodes a second GRP-like peptide in possible analogy with the two mRNAs that encode substance P and the related peptide substance K (22).

Preliminary genomic Southern blot analysis shows only single bands in restriction enzyme digests with six-base recognition sites. This pattern is consistent with the existence of a single gene and suggests that the two GRP-encoding mRNAs may arise from alternative exon splicing of a single transcript, alternative sites of initiation of transcription, or alternative sites of transcription termination and polyadenylation.

Inasmuch as GRP is produced by pulmonary carcinoid tumors and small cell carcinomas of the lung, it is a potentially important tumor marker. Because patterns of immunoreactivity observed after chromatography of tumor extracts reveal multiple GRP-like peptides, different processing of pre-proGRP may occur in different tissues. Further molecular analysis of tumor and normal GRP-encoding mRNAs will clarify these issues. The exact relationship between the various forms of GRP-like, bombesin-like, and ranatensin-like immunoreactivity in brain, lung, and gastrointestinal tract remains to be determined. The availability of cDNAs encoding pre-proGRP will greatly facilitate these studies.

**Note Added in Proof.** Orloff *et al.* (45) have reported a partial amino acid sequence of hGRP isolated from a pulmonary carcinoid tumor. The first amino acid is valine; thus, hGRP appears to be a 27 amino acid peptide beginning with valine.

We thank Richard H. Goodman for his assistance during these studies, and we thank Rudy Tansi and James Gusella for their assistance with genomic Southern blots. We also thank Allen Oser and Jin Sing Moon for their technical assistance in this project.

- Anastasi, A., Erspamer, V. & Bucci, M. (1971) *Experientia* **27**, 166–167.
- McDonald, T. J., Jornvall, H., Nilsson, G., Vagne, M., Ghatei, M. & Bloom, S. R. (1979) *Biochem. Biophys. Res. Commun.* **90**, 227–233.
- McDonald, T. J., Jornvall, H., Ghatei, M., Bloom, S. R. & Mutt, V. (1980) *FEBS. Lett.* **122**, 45–48.

- Reeve, J. R., Walsh, J. H., Chew, P., Clark, B., Hawke, D. & Shively, J. E. (1983) *J. Biol. Chem.* **258**, 5582–5588.
- McDonald, T. J., Ghatei, M. A., Bloom, S. R., Track, N. S., Radziuk, J., Dupre, J. & Mutt, V. (1981) *Reg. Pept.* **2**, 293–304.
- Bruzzone, R., Tamburrano, G., Lala, A., Mauceri, M., Annibale, B., Severi, C., de Magistris, L., Leonetti, F. & Fave, G. D. (1983) *J. Clin. Endocrinol. Metab.* **56**, 643–647.
- Brown, M., Tache, Y. & Fisher, D. (1979) *Endocrinology* **105**, 660–665.
- Brown, M., Rivier, J. & Vale, W. (1977) *Science* **196**, 998–1000.
- Dockray, G. J., Valliant, C. & Walsh, J. H. (1979) *Neuroscience* **4**, 1561–1568.
- Moody, T. W. & Pert, C. B. (1979) *Biochem. Biophys. Res. Commun.* **90**, 7–14.
- 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.
- Track, N. S. & Cutz, E. (1982) *Life Sci.* **30**, 1553–1556.
- Johnson, D. E., Lock, J. E., Elde, R. P. & Thompson, T. R. (1982) *Pediatr. Res.* **16**, 446–454.
- Moody, T. W., Pert, C. B., Gazdar, A. F., Carney, D. N. & Minna, J. D. (1981) *Science* **214**, 1246–1248.
- Erisman, M. D., Linnoila, R. I., Hernandez, O., DiAugustine, R. P. & Lazarus, L. H. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2379–2383.
- Minamino, N., Kangawa, K. & Matsuo, H. (1983) *Biochem. Biophys. Res. Commun.* **114**, 541–548.
- Price, J., Penman, E., Bourne, G. L. & Rees, L. H. (1983) *Reg. Pept.* **7**, 315–322.
- Roth, K. A., Evans, C. J., Weber, E., Barchas, J. D., Bostwick, D. G. & Bensch, K. G. (1983) *Cancer Res.* **43**, 5411–5415.
- Moody, T. W., Russell, E. K., O'Donohue, T. L., Linden, C. D. & Gazdar, A. F. (1983) *Life Sci.* **32**, 487–493.
- Price, J., Penman, E., Wass, J. A. H. & Rees, L. H. (1984) *Reg. Pept.*, in press.
- Knight, R. M., Chronwall, B. M., Pisano, J. Y., Moody, T. W. & O'Donohue, T. L. (1983) *Soc. Neurosci. Abstr.* **9**, 294.5.
- Nawa, H., Hirose, T., Takashima, H., Inayama, S. & Nakaniishi, S. (1983) *Nature (London)* **306**, 32–36.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
- Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
- Crabtree, G. R. & Kant, J. A. (1981) *J. Biol. Chem.* **256**, 9718–9723.
- Dagert, M. & Ehrlich, S. D. (1979) *Gene* **6**, 23–28.
- Narang, S. A., Brousseau, R., Hsiung, H. M. & Michniewicz, J. J. (1980) *Methods Enzymol.* **65**, 610–620.
- Hanahan, D. & Meselson, M. (1983) *Methods Enzymol.* **100**, 333–341.
- Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3966.
- Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
- Roberts, B. E. & Patterson, B. M. (1973) *Proc. Natl. Acad. Sci. USA* **70**, 2330–2334.
- Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* **67**, 247–256.
- Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1520.
- Clewell, D. B. & Helsinki, D. R. (1970) *Biochemistry* **9**, 4428–4440.
- Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5350–5354.
- Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. (1977) *Biochemistry* **16**, 4743–4751.
- Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
- Perlman, D. & Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 391–409.
- Bradbury, A. F., Finnie, M. D. A. & Smyth, D. G. (1982) *Nature (London)* **298**, 686–688.
- Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.

41. Goodman, R. H., Aron, D. C. & Roos, B. A. (1983) *J. Biol. Chem.* **258**, 5570–5573.
42. Gubler, U., Monahan, J. J., Lomedico, P. T., Bhatt, R. S., Collier, K. J., Hoffman, B. J., Bohlen, P., Esch, F., Ling, N., Zeytin, F., Brazeau, P., Poonian, M. S. & Gage, L. P. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4311–4314.
43. Land, H., Schultz, G., Schmale, H. & Richter, D. (1982) *Nature (London)* **295**, 299–303.
44. Yoshizaki, K., deBock, V. & Solomon, S. (1984) *Life Sci.* **34**, 835–843.
45. Orloff, M. S., Reeve, J. P., Ben-Avram, C. M., Shively, J. E. & Walsh, J. H. (1984) *Peptides* **12**, in press.