*Proc. Natl. Acad. Sci. USA*
Vol. 79, pp. 1049–1053, February 1982
Biochemistry

# Molecular cloning and nucleotide sequence of full-length cDNA coding for porcine gastrin

(recombinant DNA/colony hybridization/protein precursor processing)

O. JOON YOO, C. THOMAS POWELL, AND KAN L. AGARWAL*

Departments of Biochemistry, Chemistry, and Biophysics and Theoretical Biology, The University of Chicago, Chicago, Illinois 60637

*Communicated by Josef Fried, October 27, 1981*

ABSTRACT **We have cloned in *Escherichia coli* a cDNA copy of mRNA coding for the porcine antral mucosal hormone preprogastrin. Full-length double-stranded cDNA was synthesized and inserted into the *Pst* I endonuclease site in plasmid pBR322 by using a homopolymeric extension technique. A partial cDNA clone was used as a probe to identify a complete cDNA clone in a cDNA library by colony hybridization. Four positive clones were isolated, one of which corresponded to porcine preprogastrin mRNA. The nucleotide sequence of the cDNA insert (602 nucleotides) revealed 312 nucleotides in the entire mRNA coding region, 61 nucleotides in the 5′ untranslated region, 86 nucleotides in the 3′ untranslated region, and a poly(A) tail of 86 nucleotides. Gastrin is located near the carboxyl end of preprogastrin and is flanked at both its amino and carboxyl ends by a pair of basic amino acid residues. The presence of glycine and a pair of basic amino acid residues adjacent to the carboxyl-terminal phenylalanine of gastrin indicates that the glycine and a pair of basic amino acid residues may be required for the enzymatic amidation of phenylalanine to phenylalanine amide.**

Gastrin, a polypeptide hormone, exhibits a wide range of biological actions, the most potent of which is stimulation of gastric acid secretion and antral smooth muscle activity (1). Hypersecretion of gastrin in the Zollinger–Ellison syndrome is well recognized and is often associated with severe peptic ulcer disease, diarrhea, or both. This differential regulation of gastrin synthesis in normal and defective tissue is not understood. Studies on gastrin biosynthesis will further our understanding of the molecular events controlling eukaryotic gene expression and may reveal the mechanism(s) through which normal control elements are altered in mutant tissue. Because much of our knowledge on hormone precursors has come from studies with mRNA preparations used in *in vitro* translation systems (2) and from the primary structure of the DNA complementary to mRNA (3), we have begun to study gastrin mRNA structure as a first step in understanding the regulation of the biosynthesis of gastrin.

In this communication, we describe cloning in pBR322 and sequence analysis of gastrin cDNA synthesized from porcine antral poly(A)-RNA. The method employed in the present study allowed the synthesis and cloning of full-length double-stranded (ds) cDNA corresponding to a mRNA species present in less than 0.01% of the total poly(A)-RNA (gastrin mRNA constitutes 0.006% of the total poly(A)-RNA). The recombinant cDNA library was screened with a partial cDNA probe in order to obtain a complete cDNA clone. The nucleotide sequence analysis of the gastrin cDNA revealed the entire coding sequence of preprogastrin together with the structure of 5′ and 3′ untranslated regions.

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

**Enzymes and Reagents.** DNA polymerase I (large fragment) and *Pst* I endonuclease were purchased from Boehringer Mannheim. $(dT)_{15}$ and calf thymus terminal deoxynucleotidyltransferase were from P-L Biochemicals. Other restriction endonucleases were from New England BioLabs. Exonuclease-free T4 phage polynucleotide kinase (32,000 units/mg) was isolated by the modified procedure of Panet *et al.* (4). Reverse transcriptase (RNA-dependent DNA polymerase) from avian myeloblastosis virus was obtained from J. W. Beard (Life Sciences, Inc., St. Petersburg, FL).

**Nucleic Acids and Bacterial Strains.** Total RNA was extracted from the antral mucosa of freshly slaughtered hogs as described (5) and was centrifuged through 5.7 M CsCl to remove DNA and proteins (6). Poly(A)-enriched RNA was isolated by repeated chromatography on oligo(dT)-cellulose (7). pBR322 DNA was isolated from *Escherichia coli* LE392 (8). HB101, a certified EK2 host, was a gift from Howard Goodman.

**Synthesis of ds DNA.** Single-stranded (ss) cDNA was synthesized by using the procedure of Agarwal *et al.* (9). Ten micrograms of porcine antral mRNA and 320 pmol of 5′-$^{32}$P-labeled $(dT)_{15}$ ($2.4 \times 10^5$ cpm/pmol) were incubated in a 50-$\mu$l reaction mixture containing 48 units of reverse transcriptase. After incubation for 60 min at 42°C, the reaction was terminated by addition of EDTA to 20 mM, followed by extraction with phenol/$CHCl_3$/isoamyl alcohol, 25:24:1 (vol/vol). The aqueous phase was chromatographed over a Sephadex G-200 column preequilibrated with 10 mM Tris·HCl (pH 7.6)/10 mM NaCl/1 mM EDTA/0.1% sodium dodecyl sulfate. Fractions containing the 5′-labeled cDNA·RNA hybrid were pooled and the nucleic acids were precipitated with 2.5 vol of ethanol. The pellet was dissolved in 0.3 M NaOH/2 mM EDTA and incubated for 16 hr at 23°C as described (10).

The 3′ end of the cDNA was extended by a poly(dA) homopolymeric tail of approximately 20 nucleotides, using [$^3$H]dATP and terminal transferase (11). The tailed cDNA was made double stranded with reverse transcriptase and $(dT)_{15}$ as described above for ss cDNA synthesis. After incubation for 60 min at 37°C, the reaction mixture was extracted with phenol/$CHCl_3$/isoamyl alcohol, 25:24:1 (vol/vol), and the aqueous solution was chromatographed over a Sephadex G-200 column as described above. Fractions containing ds cDNA were pooled and the nucleic acids were precipitated with 2.5 vol of ethanol.

The ds cDNA was treated with DNA polymerase I (large fragment) in the presence of four deoxynucleoside triphosphates. The reaction mixture contained: 50 mM potassium phosphate (pH 7.4), 6.6 mM $MgCl_2$, 1.5 mM dithiothreitol, 192 $\mu$M each of the four deoxynucleoside triphosphates, and 4 units of DNA

Abbreviations: ss cDNA, single-stranded complementary DNA; ds cDNA, double-stranded complementary DNA.

*To whom reprint requests should be addressed.

polymerase I (large fragment). After incubation for 2 hr at 37°C, the reaction mixture was extracted with an equal volume of phenol/$CHCl_3$/isoamyl alcohol, 25:24:1 (vol/vol). The nucleic acids from the aqueous solution were precipitated with 2.5 vol of ethanol. The blunt-ended ds cDNA was then extended at its 3′ ends with poly(dC) tails of 20 nucleotides. Similarly, poly(dG) tails of 20 nucleotides were added to *Pst* I-cleaved pBR322 DNA that had been through phenol/$CHCl_3$/isoamyl alcohol extraction and ethanol precipitation steps.

**Transformation of *E. coli* HB101.** Recombinant DNA was handled as prescribed by the National Institutes of Health guidelines. Poly(dG)-tailed pBR322 (0.06 pmol) and poly(dC)-tailed ds cDNA (0.06 pmol) were dissolved in 60 μl of 5 mM Tris·HCl (pH 8.0)/100 mM NaCl/0.5 mM EDTA and annealed by heating to 45°C for 10 min and cooling to 23°C over 2 hr and then to 0°C over 2 hr. The mixture was used directly to transform *E. coli* HB101 according to a modified procedure of Goodman and MacDonald (10) which will be described elsewhere. Transformants resistant to tetracycline (25 μg/ml) were grown for 16 hr at 37°C.

**Preparation of Hybridization Probe.** Plasmid pPG 2.1 was prepared by insertion of gastrin-specific ds cDNA at the *Hin*dIII site of pBR322. For colony (12) and Southern hybridization (13), the cDNA insert (220-base-pair fragment) was isolated from the nick-translated pPG 2.1 plasmid by *Hin*dIII digestion followed by purification of the fragment by 5% polyacrylamide gel electrophoresis. Details of ds cDNA synthesis, construction, and characterization of plasmid pPG 2.1 will be described elsewhere.

**Plasmid DNA Isolation and Restriction-Site Analysis.** Closed circular plasmid DNAs were isolated from chloramphenicol-amplified cultures by the cleared lysate technique (8) followed by CsCl/ethidium bromide equilibrium gradients. Plasmid DNA was rapidly isolated from 10-ml cultures by the procedure of Klein *et al.* (14). Plasmid DNAs were digested by several restriction enzymes under the conditions recommended by the suppliers, and the products were analyzed by 1% agarose or 7% polyacrylamide gel electrophoresis.

**DNA Sequence Analysis.** Restriction endonuclease-digested DNA fragments were either 5′-end labeled by using polynucleotide kinase and [$\gamma$-$^{32}$P]ATP or 3′-end labeled by using terminal transferase and cordycepin 5′-[$\alpha$-$^{32}$P]triphosphate (15). Labeled DNA strands were separated by 5% polyacrylamide gel as described (16). The nucleotide sequence was determined by using the chemical cleavage procedure of Maxam and Gilbert (17).

## RESULTS

**Synthesis of ds cDNA.** By using 5′-$^{32}$P-labeled $(dT)_{15}$ primer, near full-length ss cDNA was synthesized from poly(A)-RNA. Because the next step of synthesis (Fig. 1) involves extension of ss cDNA at its 3′ end by terminal transferase, excess $(dT)_{15}$ and small cDNA products (shorter than 100 nucleotides) were separated from the cDNA·RNA hybrids by Sephadex G-200 column chromatography. The yield of the cDNA·RNA hybrid was 2-fold higher (2 pmol of 5′ ends of cDNA) than that of ss cDNA, which suggests that ss cDNA binds nonspecifically to Sephadex but the double-stranded cDNA·RNA hybrid does not. The ds cDNA synthesized as described in *Materials and Methods* may contain 3′- and 5′-protruding ends due to variable length of the poly(dA) tail and incomplete second strand cDNA synthesis. Blunt-ended ds cDNA was synthesized by using both 3′-exonucleolytic and polymerizing activities of the DNA polymerase I (large fragment). All of the 5′-$^{32}$P label was present with cDNA, demonstrating that nucleotides from this end were not hydrolyzed during this step of synthesis. That the cDNA thus synthesized was blunt ended was suggested by the subsequent terminal transferase reaction, in which the rate of addition of dC tails to the 3′ ends of this cDNA was slower than the rate of addition of the dA tail to the ss cDNA.

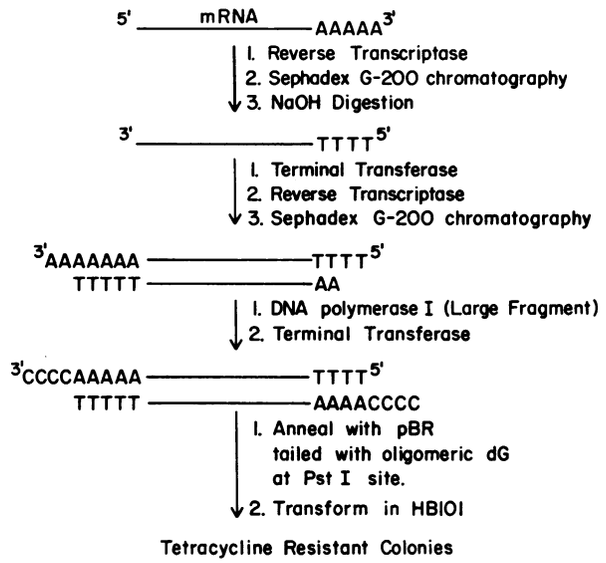


FIG. 1. Outline of protocol used for cloning of porcine antral poly(A)-RNA.

**Construction of cDNA Clones.** Plasmid pBR322 DNA was linearized with *Pst* I and tailed with approximately 20 dG residues. The tailed plasmid (0.06 pmol) and tailed ds cDNA (0.06 pmol) were annealed and used to transform *E. coli* HB101. From 0.06 pmol of ds cDNA, 4032 tetracycline-resistant colonies were obtained (a transformation efficiency of $6.7 \times 10^4$ colonies per pmol of ds cDNA), 90% of which showed ampicillin sensitivity.

**Identification of the Gastrin Clone.** Transformants that contained the gastrin cDNA insert were identified by the *in situ* hybridization procedure of Grünstein and Hogness (12). The purified cDNA insert from clone pPG 2.1 that contained the 5′ half of the gastrin mRNA sequence was used as the hybridization probe. One strongly positive and three weakly positive clones were obtained. As shown in Fig. 2*a*, colony 3282 exhibits substantial hybridization to the cDNA probe compared with colony 3275 (hybridization of the remaining two colonies was similar to 3275 and is not shown). Restriction endonuclease analysis revealed that colony 3282 was a strong candidate for the gastrin clone and the remaining three clones, although con-

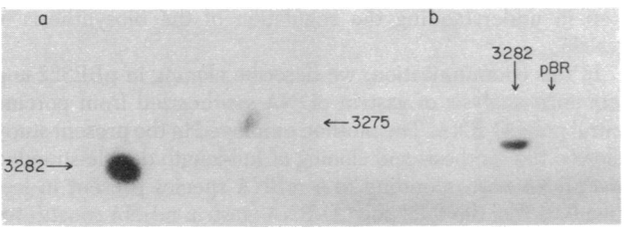


FIG. 2. Colony hybridization of transformed clones and Southern hybridization of plasmid pPG 3282 DNA with cDNA. (*a*) Transformants were grown and their plasmid DNA was immobilized to nitrocellulose filters and hybridized with the $^{32}$P-labeled gastrin cDNA insert isolated from plasmid pPG 2.1 DNA. An autoradiograph of the colony hybridization of the strongly positive (clone 3282) and one of the weakly positive (clone 3275) clones is shown. (*b*) Plasmid DNAs pPG 3282 and pBR322 were digested with *Eco*RI, fractionated in 1% agarose gel, immobilized on nitrocellulose, and hybridized with the $^{32}$P-labeled gastrin cDNA insert isolated from plasmid pPG 2.1 DNA.

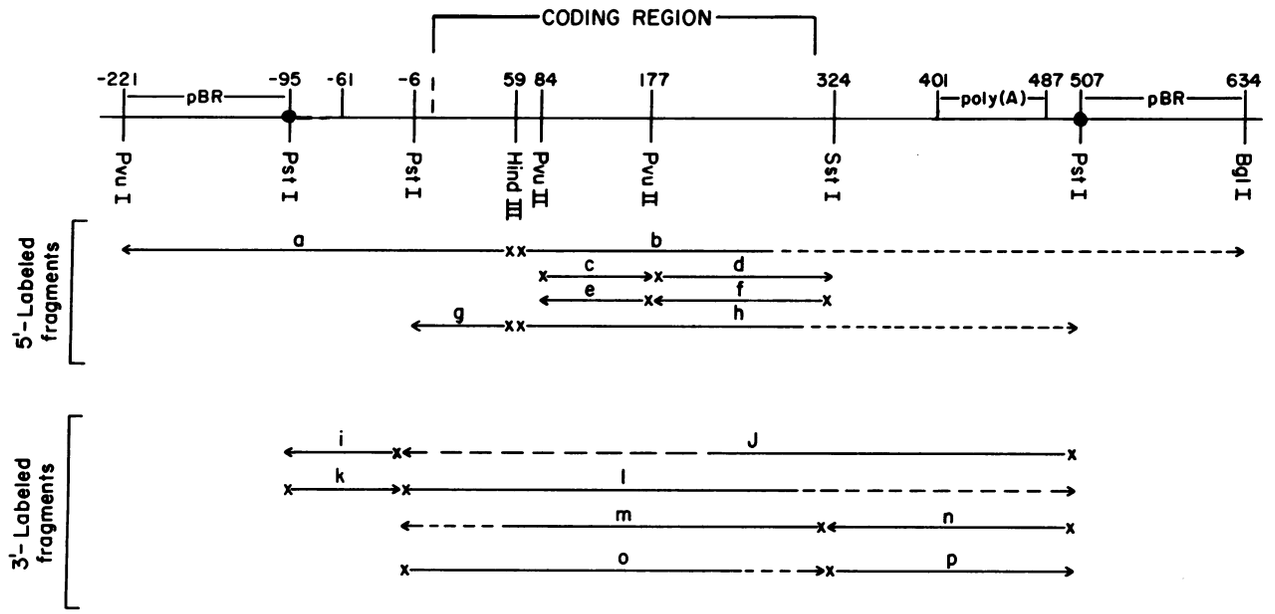


FIG. 3. Restriction map and sequencing strategy for the cDNA insert in the plasmid pPG 3282. Nucleotide residues are numbered in the direction from 5′ and 3′ in the message strand, beginning with the first residue in the coding region for gastrin. The negative numbers indicate nucleotides on the 5′ side of residue 1. The solid circle indicates the position of the cDNA insert, which includes poly(dG) and poly(dC) tails of 20 and 18 base pairs on the left and right sides, respectively, and a poly(dA) tail of 16 nucleotides on the left. The coding region for porcine gastrin is indicated by the bracket at the top of the figure. The nucleotide position 401 to 487 represents the poly(A) tail. Only the restriction sites used for sequence analysis are shown and identified by nucleotide numbers. Lines below the map indicate the restriction fragments used in sequence determinations. The fragments were either 5′-end labeled or 3′-end labeled as indicated. The labeled end of the fragment is shown by the x. Solid parts of each line indicate the clearly readable sequence. Strands of the fragments c, d, e, f, i, j, k, l, m, n, o, and p were separated on a 5% polyacrylamide strand-separating gel. The position of each nucleotide of the insert was determined at least twice.

taining sequences related to gastrin, were not gastrin specific. Further evidence that clone pPG 3282 is gastrin specific was obtained by Southern hybridization (13) of *Eco*RI-digested DNA with the nick-translated, and *Hin*dIII-cleaved pPG 2.1 fragment. As shown in Fig. 2*b*, clone pPG 3282 hybridizes effectively with the gastrin-specific probe, whereas pBR322 DNA does not.

**Characterization of the Gastrin cDNA Clone.** Plasmid DNAs from pPG 3282 and pPG 3275 were prepared (14) and their cDNA inserts were excised by digestion with *Pst* I. Analysis of the products by 7% polyacrylamide gel electrophoresis showed that the cDNA inserts in the plasmids were 602 and 300 nucleotides, respectively. Because plasmid pPG 3275 contained a short cDNA insert and showed a different restriction digestion pattern than the plasmid pPG 3282, it was not studied further. The restriction map of the pPG 3282 cDNA insert is shown in Fig. 3. The analysis showed that the insert contains single *Pst* I, *Hin*dIII, *Sst* I, and *Hin*fI sites, two *Pvu* II sites, and several *Eco*RII and *Hae* III sites (*Eco*RII and *Hae* III sites are omitted from the figure because the fragments generated by these enzymes were not used in the subsequent nucleotide sequence analysis).

For nucleotide sequence analysis, plasmid pPG 3282 was cleaved with restriction enzymes *Pst* I, *Hin*dIII, *Pvu* II, *Sst* I, *Pvu* I, and *Bgl* I. In each case, the resulting fragments were either 5′-labeled with [$\gamma$-$^{32}$P]ATP and polynucleotide kinase or 3′-labeled with cordycepin 5′-[$\alpha$-$^{32}$P]triphosphate and terminal transferase. Fragments labeled on both ends were either digested with another restriction endonuclease or subjected to strand separation (16) prior to sequence analysis. The nucleotide sequence of the end-labeled fragments was determined by the chemical cleavage procedure of Maxam and Gilbert (17). During the sequence analysis, it was observed that nucleotides at the beginnings and ends of the restriction fragments were usually difficult to resolve, and therefore we used a number of different restriction enzymes to generate overlapping fragments. As shown in Fig. 3, the sequences of both strands were partially or fully determined and in each case the sequence was determined at least twice to avoid ambiguities. The nucleotide sequence of the cDNA insert is shown in Fig. 4. The ds cDNA insert contains homopolymeric tails of 16 dA·dT followed by 18 dC·dG residues on the left and 20 dC·dG residues on the right. The mRNA sequence deduced from the cDNA is composed of 548 nucleotides, including the poly(A)-tail of 86 nucleotides. The coding region of preprogastrin is 312 nucleotides long. There are 61 nucleotides in the 5′-untranslated region, which includes AUG (initiator) and UGA (terminator) codons. The 3′-untranslated region of 86 nucleotides contains a UGA termination codon that is 54 nucleotides downstream from the first UAG termination codon. The sequence A-A-U-A-A-A, 14 nucleotides from the poly(A) in the 3′ region, is similar to sequences found in other eukaryotic mRNAs (18) and is presumed to be involved in the addition of the 3′ poly(A) sequence after transcription.

## DISCUSSION

We have previously demonstrated that an oligodeoxynucleotide can be used as an effective probe in the detection and characterization of gastrin mRNA (5). Furthermore, in our studies with insulin mRNA, oligonucleotide-generated cDNA has been used as a specific hybridization probe in the selection of recombinant DNA molecules (19). However, the relatively low level of gastrin mRNA (0.006%) in the total poly(A)-RNA limits the amount of oligonucleotide-initiated cDNA that can be synthesized and used in the selection of recombinant DNA molecules by colony hybridizations. To overcome this difficulty, a recombinant plasmid containing oligodeoxynucleotide-generated ds cDNA was constructed and introduced into HB101 by transformation. A

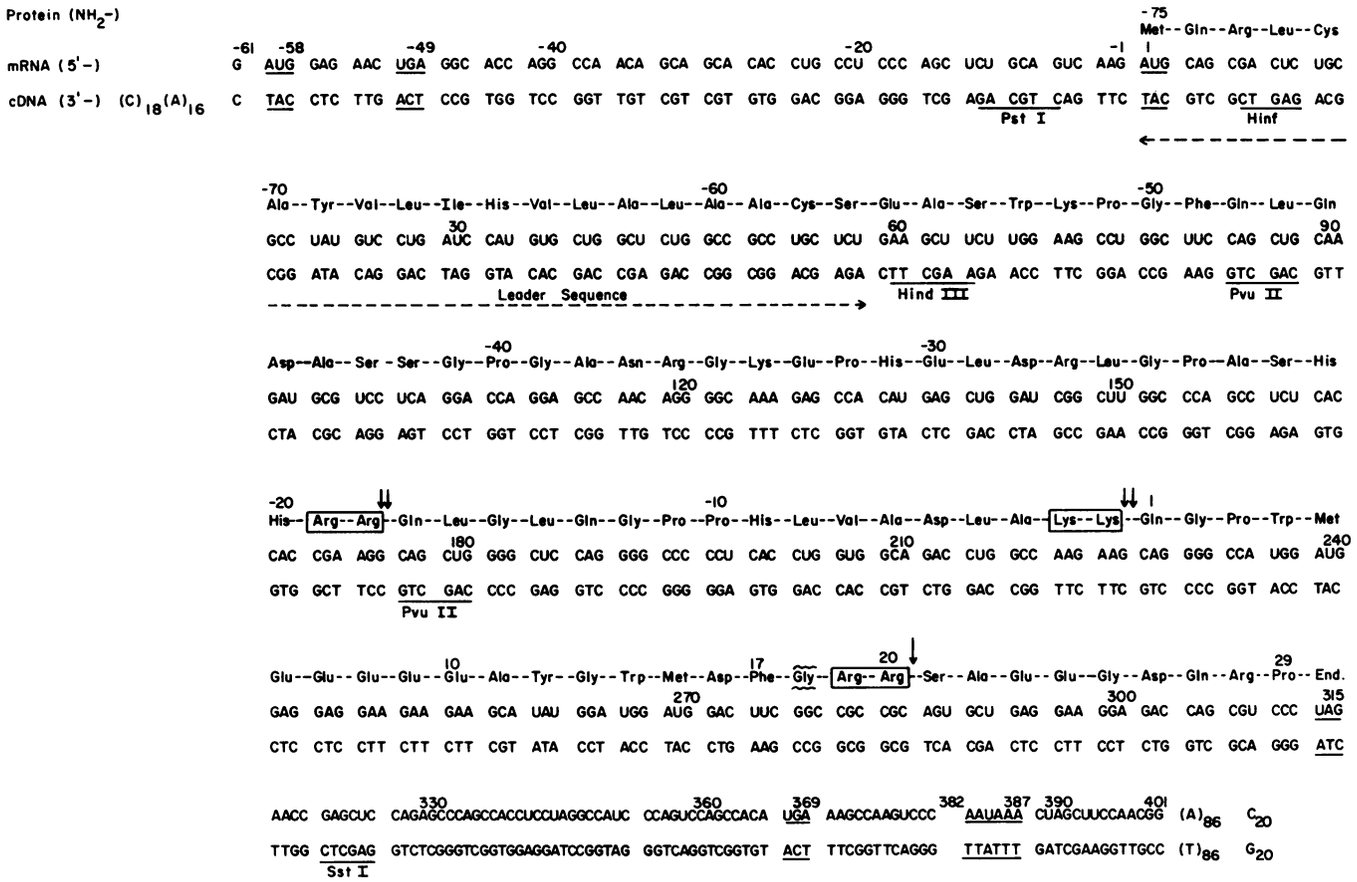


FIG. 4. Primary structure of porcine preprogastrin mRNA. The entire nucleotide sequence of the mRNA was deduced from that of the 602-base-pair cDNA insert in pPG 3282 (Fig. 3). The nucleotide numbers are indicated above the mRNA sequence, beginning with the initiation codon. The nucleotides in the untranslated region are indicated by negative numbers. Major restriction sites are underlined below the cDNA sequence. The initiation codon (AUG), termination codon (UGA, UAG), and the polyadenylylation site (A-A-U-A-A-A) are underlined. Orientation of the cDNA insert is shown in Fig. 3. The amino acid sequence derived from the mRNA is depicted in the upper line and begins with Met at position −75 and ends with Pro at position 29. The gastrin sequence begins at Gln at position 1 and ends with Phe at position 17. Double arrows indicate the proposed trypsin-like cleavages at the pair of basic residues (enclosed in the boxes). The single arrow indicates presumed processing leading to enzymatic conversion of Phe-Gly to Phe-$NH_2$. The presumed leader or signal sequence beginning at residue −75 and continuing to residue −57 is shown by broken lines.

clone, pPG 2.1, that contained the 5′ half of the gastrin messenger sequence was subsequently isolated and employed as a hybridization probe in the isolation of clones containing the entire mRNA sequence.

The procedures commonly employed in the synthesis of ds cDNA invariably yield cDNAs that are shorter than their parent mRNAs (20, 21). Furthermore, use of some of these procedures has resulted in the introduction of an artifactual inverted sequence at the 5′ terminus of the mRNA strand (19, 22). Both of these undesirable events have been shown to occur during the synthesis of ds cDNA from ss cDNA by the use of DNA polymerase I and S1 nuclease. However, full-length ds cDNA, free of inverted sequences, can be synthesized by using the approach described here.† As shown in Fig. 1, the 3′ end of the ss cDNA was extended by poly(dA) residues and used as a template in the synthesis of ds cDNA using reverse transcriptase and a $(dT)_{15}$ primer. Because the ds cDNA synthesized in this manner contained sticky ends due to the heterogeneity in the length of the poly(A)-tail or incomplete double strand synthesis, subsequent treatment with terminal transferase would have extended the 3′ ends of this DNA unevenly. For this reason, blunt-ended ds cDNA was generated by treatment with DNA polymerase I (large fragment) in the presence of four deoxynucleoside triphosphates prior to addition of poly(dC) tails by terminal transferase. Because the ds cDNA–pBR322 hybrid was effective in transforming *E. coli* with high efficiency, it indicates that the majority of the cDNA molecules are double stranded and tailed at their 3′ ends. Comparison of the 5′-mRNA sequence from nucleotide 246 to −61 (307 nucleotides, Fig. 4) with the size of the specific cDNA synthesized from the gastrin oligodeoxynucleotide primer beginning at nucleotide 246 (304 nucleotides, ref. 9) strongly indicates that the ds cDNA thus synthesized contains the entire 5′ mRNA sequence. However, we cannot rule out the absence of a few nucleotides from the 5′ end. Isolation of only one gastrin clone out of 4032 recombinants shows that low-abundance mRNA sequences can be cloned. In addition, these results demonstrate that this procedure for ds cDNA synthesis is efficient and generates recombinants with full-length cDNA inserts.

The primary sequence of the gastrin mRNA was deduced from the cDNA insert and is presented in Fig. 4. The gastrin mRNA contains 462 nucleotides, excluding the 3′ poly(A) tail of 86 nucleotides.‡ Out of 462 nucleotides, 312 nucleotides code

† Since the completion of this work, an approach similar to but not identical with the approach described here has been reported (23).

‡ We have previously determined that the gastrin mRNA is 620 nucleotides long (24) and therefore an average length of 160 nucleotides for the poly(A) sequence is predicted.

for the 104 amino acids of porcine preprogastrin. The discovery of AUG (initiator) and UGA (terminator) codons in the 5′ untranslated region is similar to the observation made for preproparathyroid hormone mRNA (25). The notion that most eukaryotic mRNAs are monocistronic would suggest that these codons are ineffective in the translation of mRNA. The significance of the presence of these codons in the 5′ untranslated region is unknown.

The amino acid sequence of preprogastrin was deduced from its mRNA sequence by using the reading frame for amino acid codons corresponding to the amino acid sequence of gastrin. As shown in Fig. 4, the mRNA codes for a 104-amino-acid polypeptide in which the gastrin sequence (amino acids 1 to 17) is located near the carboxyl terminus. Inspection of this sequence revealed that the formation of gastrin involves trypsin-like cleavages at both the amino and the carboxyl terminus of gastrin. Processing at the amino terminus of gastrin appears to involve cleavages at the carboxyl side of the pairs of basic residues (boxed, Fig. 4) by trypsin-like enzyme(s) as indicated by double arrows. This step in the processing generates amino-terminal glutamine residues at positions 1 and -17 that are subsequently cyclized to pyroglutamic acids (26). The amino-terminal processing of progastrin is similar to the processing for all proproteins that contain pairs of basic residues at cleavage sites (27).

The conversion of preprogastrin to gastrin also requires cleavage and modification at the carboxyl terminus. The amino acid sequence of preprogastrin reveals the presence of Gly and an Arg-Arg pair adjacent to the carboxyl-terminal Phe. Because gastrin contains a carboxyl-terminal Phe-$NH_2$, it appears likely that Gly at position 18 is enzymatically converted to an amino group, forming an amide. After the cleavage at the Arg-Arg pair, the synthesis of the amide can be envisioned to occur in one of two ways: (*i*) The Arg-Arg pair is removed by a carboxypeptidase-B-like activity and then Gly is converted into the amino group of the amide. (*ii*) The Gly-Arg-Arg sequence serves as a recognition site for the enzymatic synthesis of the amide. The second possibility seems more reasonable because the sequence Gly-Lys-Lys has been found at the site of amidation in preprocalcitonin (28). Moreover, preliminary studies indicate that treatment of the synthetic heptapeptide Tyr-Gly-Trp-Met-Asp-Phe-Gly (this sequence corresponds to the carboxyl terminus of gastrin) with dog antral G cell extract does not produce the hexapeptide amide (unpublished observations). It therefore appears likely that this unique tripeptide sequence of glycine-basic amino acid-basic amino acid is necessary for this amidation to occur.

As a result of this processing (shown by single and double arrows in Fig. 4), formation of peptides of 55, 15, and 9 amino acid residues is predicted. The peptide of 55 amino acid residues contains a leader sequence of 19 hydrophobic amino acids, which are presumably removed as the nascent peptide chain enters the endoplasmic reticulum (29). Some or all of these released peptides may be secreted along with gastrin and may have an as yet unrecognized biological activity or some structural role that is required for the proper folding of the precursor into a conformation readily recognized by the processing enzymes.

We recognize the contributions of the late Professor Morton I. Grossman in the gastrointestinal hormone field. We thank Drs. W. Epstein, C. Cladaras, M. Mevarech, and C. Ney for helpful discussions and critically reading the manuscript. This work was supported in part by U.S. Public Health Service Grant AM-21901. O.J.Y. was supported by Training Grant GM-07183 to the University of Chicago. K.L.A was supported by U.S. Public Health Service Career Development Award GM-00224.

1. Walsh, J. H. & Grossman, M. I. (1975) *N. Engl. J. Med.* **292**, 1377–1384.
2. Clark, J. L. & Steiner, D. F. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1964–1968.
3. Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A. C. Y., Cohen, S. N. & Numa, S. (1979) *Nature (London)* **278**, 423–427.
4. Panet, A., van de Sande, J., Loewen, P., Khorana, H., Raae, A., Lillehaug, J. & Kleppe, K. (1973) *Biochemistry* **12**, 5045–5050.
5. Noyes, B., Mevarech, M., Stein, R. & Agarwal, K. L. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1770–1774.
6. Glisin, V., Crkvenjakov, R. & Byus, C. (1974) *Biochemistry* **13**, 2633–2637.
7. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
8. Clewell, D. B. & Helinski, D. R. (1970) *Biochemistry* **9**, 4428–4440.
9. Agarwal, K. L., Brunstedt, J. & Noyes, B. E. (1981) *J. Biol. Chem.* **256**, 1023–1028.
10. Goodman, H. M. & MacDonald, R. J. (1979) *Methods Enzymol.* **68**, 79–90.
11. Nelson, T. & Brutlag, D. (1979) *Methods Enzymol.* **68**, 41–50.
12. Grünstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3965.
13. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
14. Klein, R. I., Selsing, E. & Wells, R. D. (1980) *Plasmid* **3**, 88–91.
15. Tu, C.-P. D. & Cohen, S. N. (1980) *Gene* **10**, 177–183.
16. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
17. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
18. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263**, 211–214.
19. Chan, S. J., Noyes, B. E., Agarwal, K. L. & Steiner, D. F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5036–5040.
20. Maniatis, T., Kee, S. G., Efstratiadis, A. & Kafatos, F. C. (1976) *Cell* **8**, 163–182.
21. Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) *J. Biol. Chem.* **253**, 2483–2495.
22. Weaver, C. A., Gordon, D. F. & Kemper, B. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 4073–4077.
23. Laud, H., Grez, M., Hauser, H., Lindenmaier, W. & Schütz, G. (1981) *Nucleic Acids Res.* **9**, 2251–2266.
24. Mevarech, M., Noyes, B. E. & Agarwal, K. L. (1979) *J. Biol. Chem.* **254**, 7472–7475.
25. Kronenberg, H. M., McDevitt, B. E., Majzoub, J. A., Nathans, J., Sharp, P. A., Potts, J. T., Jr., & Rich, A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 4981–4985.
26. Gregory, R. A. & Tracy, H. J. (1975) in *Gastrointestinal Hormones*, ed. Thompson, J. (Univ. Texas Press, Austin, TX), pp. 13–24.
27. Steiner, D. F., Kemmler, W., Tager, H. S., Rubenstein, A. H., Lernmark, A. & Zühlke, H. (1975) in *Proteases and Biological Control*, eds. Reich, E., Rifkin, D. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 531–549.
28. Jacobs, J. W., Goodman, R. H., Chin, W. W., Dee, P. C., Habener, J. F., Bell, N. H. & Potts, J. R., Jr. (1981) *Science* **213**, 457–459.
29. Shields, D. & Blobel, G. (1978) *J. Biol. Chem.* **253**, 3753–3756.