Vol. 216, No. 1, 1995
November 2, 1995
BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 34-41

MOLECULAR CLONING AND SEQUENCING OF THE MURINE GASTRIN GENE

Theodore J. Koh and Timothy C. Wang*

Gastrointestinal Unit, Massachusetts General Hospital, Boston, Massachusetts 02114

Received September 26, 1995

Summary: Gastrin is a peptide hormone important in acid regulation, growth of enterochromaffin cells of the oxyntic mucosa, and smooth muscle contractility. We isolated a genomic clone of gastrin from a SV-129 murine genomic library. The murine gastrin gene contains an open reading frame of 101 amino acids. The deduced amino acid sequence has 91.1% homology with the rat and 67.3% homology with the human analogue. Southern blot analysis of the murine gastrin clone gave the same restriction pattern as that seen from mouse kidney genomic DNA, consistent with this being a single copy of the gastrin gene in the mouse genome. Transfection studies demonstrated that the murine gastrin gene expression is stimulated by epidermal growth factor to the same extent as the human gastrin gene.

Gastrin is a peptide hormone important in the regulation of gastric acid secretion, mainly through the regulation of histamine [1]. It also stimulates growth of the enterochromaffin-like (ECL) cells of the oxyntic mucosa [2] and to a lesser extent parietal cells of the fundus [3], stimulates contraction of gastric smooth muscle cells [4] and may play a role in gastric motility. Recent evidence has suggested a role for gastrin in the development of the gastrointestinal tract, especially in the pancreas. The major site of fetal gastrin expression is within the immature pancreatic islets with peak expression occurring on fetal day 18, when differentiation of ductular precursor cells into exocrine and endocrine cells occurs, followed by the rapid proliferation of islet cells [5]. Infusion of gastrin into adult rats has resulted in increased [^{3}H]thymidine uptake and pancreatic weight, suggesting that it may act as a pancreatic growth factor [6]. A transgenic mouse construct developed in this laboratory containing a rat insulin-human gastrin (INS-GAS) chimeric transgene showed no increase in pancreatic islet cell mass. However, when crossed with a transgenic mouse which overexpresses TGF-α in the pancreas, the double transgenic INS-GAS/TGF-α mice demonstrated a two-fold increase in pancreatic islet mass [7].

*Corresponding author. Fax:(617)726-3673.

0006-291X/95 $12.00

The biological effects of gastrin are mediated to a large extent through the CCK-B/gastrin receptor. Preliminary reports of a targeted gene disruption of the CCK-B/gastrin receptor suggest that the CCK-B/gastrin receptor is not essential for the development of the gastrointestinal tract, pancreas or brain [8]. However recent studies indicate that some of the growth effects of gastrin may be mediated by incompletely processed forms, such as glycine-extended gastrin. Glycine-extended gastrin can stimulate the growth of a pancreatic cell line AR4-2J, and this effect appears to be mediated through a novel receptor [9] and a different signaling pathway [10] than amidated gastrin. Thus, a growth effect due to glycine-extended gastrin would not be revealed through knockout of the CCK-B receptor.

In addition, studies examining the expression of the gastrin promoter in transgenic mice have shown important species-specific differences [11]. In particular, the rat gastrin promoter was more effective in targeting antral G cells of the stomach compared to the human gastrin promoter. These studies also raised the possibility that intragenic sequences may be required [11]. Thus the question arises as to whether the mouse gastrin promoter would be even more effective in achieving tissue specific expression in transgenic mice.

Therefore, to better study gastrin's role in development, as well as its regulation, we have isolated a murine genomic clone of gastrin. This clone will serve as a useful reagent for future studies utilizing the techniques of targeted gene disruption and the creation of chimeric transgenes.

MATERIALS AND METHODS

Cloning of Murine Gastrin: A murine genomic library SV-129 (Stratagene Lambda FIX II Library) was screened using a [α-^{32}P]dCTP-labeled random-primed probe of rat gastrin exon II obtained by isolating the Bgl II-Sac I fragment from the rat gastrin cDNA [12]. Hybridization was carried out at 65°C in Rapid-hyb buffer (Amersham Life Science) for two hours. Positive clones were plaque purified. The DNA from the positive clones was harvested using the plate lysis method [13] and then purified by using a cesium chloride gradient [14]. The DNA was then subcloned into the Not I site of pBluescript (KS+).

Nucleotide Sequence Analysis: The dideoxynucleotide method [15] (Sequenase II- U.S. Biochemicals) was used. Synthetic oligonucleotide primers were made initially from conserved sequences within the gastrin genome. After murine gastrin sequences were obtained, all subsequent primers were synthesized from the mouse sequence.

Human Genomic Southern Blot Analysis: 15 μg of murine kidney genomic DNA (BALB/c, Clontech) were digested with either Bgl II, Nhe I, Not I, or Xba I. The DNA was then separated on a 1% agarose gel and transferred to a Magna nylon membrane (Micron Separations Inc.) using standard procedures [12]. Hybridization with a rat gastrin exon II [α-^{32}P]dCTP-labeled random-primed probe was carried out at 65°C in Rapid-hyb

buffer (Amersham Life Sciences) for two hours, then washed at 65°C in 0.2 X SSC, 0.1% SDS for 40 minutes. Identical studies were carried out with 1 μg plasmid DNA containing our murine gastrin clone .

Construction of Gastrin Promoter-Luciferase Reporter Construct: A gastrin promoter-luciferase reporter was created by mutating the start site on exon II by polymerase chain reaction (GeneAmp, Perkin Elmer) and inserting the firefly luciferase gene (pGL2-Basic, Promega) into the Hind III site. Synthetic oligonucleotide probes were made from a Kpn I site at -1043 and from the start of exon II, substituting the translational start site with two stop codons and a Hind III site. The polymerase chain reaction product was then inserted into the corresponding Kpn I- Hind III fragment in the murine gastrin plasmid. The luciferase gene was then inserted into the Hind III site. This resulted in a mouse gastrin-luciferase minigene containing 1043 base pairs of 5' flanking sequence, 3 kb of 3' flanking sequence, both introns, and all three exons.

Transfection Studies: A human gastric carcinoma cell line AGS-B (ATCC) grown in serum-free media (Ultraculture, BioWhittaker) were transfected with the reporter construct using the calcium phosphate transfection method (5 Prime→3 Prime). pGL2 basic was transfected in some cells as a control. The cells were washed at 24 hours and half of the cells were exposed to serum-free media containing 10^{-9} M epidermal growth factor (EGF), the other half to serum-free media alone. Cells were harvested 24 hours later and luciferase activity measured in a monolight luminometer (Analytical Luminescence Laboratory) [16].

RESULTS AND DISCUSSION

A murine SV-129 genomic library was screened with a random-primed probe to rat gastrin exon II. 100,000 to 200,000 individual clones were screened. Two positive clones were obtained. After performing phage preps, a 20 kb insert containing all three coding exons and 15 kb of 5' flanking sequence was subcloned into pBluescript KS+. Restriction mapping was performed (Figure 1). A genomic Southern (Figure 2) revealed that the murine gastrin clone had the same restriction pattern as murine kidney genomic DNA, providing further proof that the clone obtained is indeed that of murine gastrin.

The sequence of the murine gastrin gene is shown (Figure 3). There is 91.0% homology with the rat gastrin gene at the nucleotide level in the coding exons, compared to 78.8% homology with the non-coding exon I. The gene is 2.6 kb long; comparable in

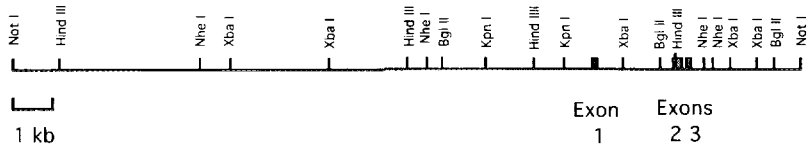

Figure 1. A restriction endonuclease map of the mouse gastrin gene isolated from a Not I insert of a mouse genomic DNA library (SV-129). Lines represent introns and flanking DNA sequences, boxes represent exons.

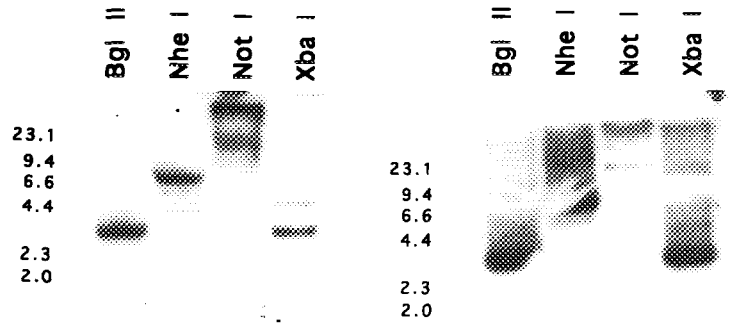

Figure 2. (Left) Southern blot of mouse genomic DNA (15 μg per lane) digested with either Bgl II, Nhe I, Not I, or Xba I and probed with a rat gastrin exon II random-primed probe. (Right) Southern blot of plasmid DNA from the mouse gastrin gene clone (1 μg per lane) cut with the same enzymes and hybridized with the same probe.

size to rat gastrin gene but considerably shorter than the human gastrin gene where the first intron is much larger.

The deduced amino acid sequence is shown in Figure 4. The open reading frame encodes for a peptide of 101 amino acids, of which the first 21 residues encode a typical signal peptide. Overall, the mouse preprogastrin amino acid sequence has 86.1% homology with the rat, 67.3% homology with the human, and 67.3% homology with the cow. Of note, there is 94.1% homology of murine gastrin 34 with rat gastrin 34 (differing in only two amino acids, Leu^{62} and Glu^{76}). There is also conservation of the key endoproteolytic cleavage sites, including the site of signal peptidase, Ala^{21}-Ser^{22}, and the dibasic cleavage sites Arg^{57}-Arg^{58}, Lys^{74}-Lys^{75}, and Arg^{94}-Arg^{95}. In addition, Tyr^{87}, which is a target of sulfotransferase activity, and Ser^{96}, which undergoes phosphorylation [17] were conserved. The conservation of these sites between species confirms the idea that these sites are important in gastrin processing. The mouse C-terminal extension peptide is only six amino acids long, as is the case with the human gastrin gene, whereas in the rat, pig, and cow it is nine amino acids long, suggesting that the terminal three amino acids are not important in processing.

The mouse gastrin promoter region has a TATA-like element (-33 TTTATTAA -26) which is more similar to the human gastrin gene (TTTATAA) than that of the rat gastrin gene (TATATAA), differing from the human gastrin gene by the addition of a single T. In other respects, the mouse gastrin promoter resembles the rat gastrin promoter which we have previously reported [11] very closely. In the mouse gastrin promoter, the β-interferon-like negative element (-97 ATTTCTCT -90) and the E-box element (-69 CACCTGG -63) were identical to those in the rat [18]. In addition, the mouse gastrin

Mouse Gastrin Sequence

5' Flanking Sequence					
-1043	ggtaccactg	ccettatata	cacaagaccc	aagggacatc	aaagctcctc
-993	aagacaggaa	agggggccca	atgacgcata	catgatagca	aatgtctctc
-943	atctcagtta	tggaaaaaga	agtctagaga	ttgttataac	cttgcaaacc
-893	acattcaggt	ggaagtttat	caaataaacc	aagaaatgta	tcatttttcc
-843	ttgcccggta	catctttgaa	aaaagccttc	gttcactaac	cactttaaga
-793	tgtaattttg	ttgcaatgtt	atgaattctt	taagcaaatc	acaacctacc
-743	cctgaaagca	acatggaaga	gtcttgaaac	acaaggtcct	caaggaactg
-693	cctcccagta	gggcctggct	agaggtctga	ttaggtgcaa	accaaggagt
-643	ccctccaaga	gtgttgggca	gatcccgccc	caacatagag	ggcatagaca
-593	agagatgctg	agggttttcc	atggtaacag	gagcataaaa	ggtttccaag
-543	gcaaccaaca	tgagaatctc	aacagaaaga	gcaagcccat	catcccatgg
-493	ccatttctcc	tctgataact	atgacaatga	tatcctgttt	gcaaaagccc
-443	aatggagggg	acagggatca	ataagtcagc	ctacatgcat	acactaccig
-393	ttcatcaaac	atcaacatat	gcgcacattg	gtcatttgat	tctcatttat
-343	tctctcccti	cccictcctc	tcgagccccg	ccccatggga	tgtagtccaa
-293	tagaagacta	ggaagccgat	ggcccaaggc	atacagtaga	atgagctgct
-243	ctggagccca	gggaatctgt	ctcctgtctg	tctcctccct	gttcagccc
-193	ttcccctttc	ccatattctg	aataaataac	taattgattg	acaacctctg
-143	tgacagggtg	gggcagagtg	acaggttgat	cctacattcc	aggcccattt
-93	ctcttgctgt	ggggagtctg	gcctcacctg	gaaggagggt	ttgggaggaa
-43	cctcgagggc	tttattaa gg	caagacctgg	agcagcgcac	aag
Exon 1					
1	agcagagctg	acccagcgcc	acaacagcca	actatttccc	agctctgtgg
Intron 1					
51	gtgagaaatt	gggaagggaa	aagagggaaa	gggaggggag	ggaagggaca
101	gcaaggatga	gtattgacaa	gcccccttcc	tttagaggct	ccaggccatg
	.....	.....	.....	.....	.....
2040	ctagcgaggc	agattagagc	tcatctggtc	agcccttiac	ctctgggact
2090	atggaatagc	cccatcttga	agcccagttc	ttctcccttc	tcag
Exon 2					
2134	acaagATGCCT	CGACTGTGTG	TGTACATGCT	GGTCTTAGTG	CTGGCTCTAG
2184	CTACCTTCTC	GGAAGCTTCT	TGGAAACCCC	GCTCCCAGCT	ACAGGATGCA
2234	TCATCTGGAC	CAGGGACCAA	TGAGGACCTG	GAACAGCGCC	AGTTCAACAA
2284	GCTGGGCTCA	GCCTCTCACC	ATCGAAGGCA	GCTGGGGCTC	CAGGGTCCTC
2334	AACACTTCAT	AGCA			
Intron 2					
2348	ggtagtaagt	cctagctgag	ccaggtttgg	ccatggttcc	tccatactga
2398	cccccacagtt	ctttgagact	tggccttctt	ctctccgttc	tttacctcct
2448	gttacctctt	ca			
Exon 3					
2460	GACCTGTCCA	AGAAAGAGAG	GCCACGAATG	GAGGAAGAAG	AGGAAGCATA
2510	CGGATGGATG	GACTTTGGCC	GCCGCAGTGC	TGAGGAAGAC	CAGTAGgact
2560	agcaacactc	ttccagagcc	cagccatctc	cagccaccce	tccccccagct
2610	ccgtccttac	aaaaacatat	taaaaataac	gctagcttcc	aatgt
3' Flanking Sequence					
2655	atcccctgagt	catgtcatgc	ttgactggag	agggggttgag	gcaggaggga
2705	gctgagagct	gagtagaggt	tagaggcaga	cccaagagta	gcactaatca

Figure 3. The sequence of the mouse gastrin gene. The nucleotides are numbered from the initiation site. The coding sequence is capitalized. The TATA box is underlined. A portion of the sequence of the first intron is not included but is available in Gene Bank.

gene also contained an upstream AT-rich element (-164 CTAATTGA -157) which was identical to that in the rat promoter but also similar to that in the human (CTAAATGA) [19]. However the putative CACC box in the mouse gastrin promoter (-105 CCAGGCC -99) was somewhat divergent from the homologous elements in the rat (CCACACC) and the human (CCCCACCCC) gene promoters [20] and resembled more closely the consensus sequence for AP2 (CCCAGGC) [21].

```
          1                   10                  20
MOUSE:    M P R L C V Y M L V L V L A L A T F S E A S W K P R S Q
RAT:                C
HUMAN:      Q           V   I F A         A
COW:        Q         A H V   I           A   C                 H

          30                  40                  50
MOUSE:    L Q D A S S G P G T N E D L G Q R Q F N K L G S A S H H
RAT:                    R T   G A   E   H   L E       P
HUMAN:    Q P       P L   T   A   R     E L P W L E Q Q   P
COW:                P V A   R A   R G   E P L R M   R       P       N P

                60                  70                  80
MOUSE:    R R Q L G L Q G P Q H F I A D L S K K E R P P M E E E E
RAT:              P                                   Q
HUMAN:            P         P   L V       P         Q G   W L
COW:                    D   P   M V                 Q G   W

                90                  100
MOUSE:    E A Y G W M D F G R R S A E E D Q
RAT:                                  E D Q Y N
HUMAN:                              D E N
COW:      A                           G D Q H P
```

Figure 4. The deduced amino acid sequence of the mouse gastrin. The amino acids are numbered beginning at the translation initiation site. Below the mouse gastrin sequence is shown the amino acid sequence of the rat gastrin, human gastrin, and bovine gastrin genes. Only amino acids which differ from the mouse gastrin sequence are shown.

The mouse gastrin gene, like the rat gastrin gene, also lacks the consensus EGF response element (GGGGCGGGGTGGGGGG) present in the human gastrin gene [22]. Nevertheless, the mouse gastrin promoter was responsive to EGF. A mouse gastrin promoter-luciferase construct was made which contained the entire mouse gastrin gene, including flanking regions, and the luciferase coding region inserted at the Hind III site of exon 2 (Figure 5A). The construct contained 1043 base pairs of 5' flanking DNA. In addition, the initiator methionine codon of mouse gastrin was altered by PCR mutagenesis, as described in Materials and Methods. This minigene construct was created, as opposed to a simpler promoter construct, based on previous transgenic studies which suggested that cis-acting elements present in both the 5'flanking region and intragenic sequences are necessary to achieve tissue-specific expression [11,23]. Transfection studies with the mouse gastrin gene-luciferase construct in the human gastric carcinoma cell line AGS showed a two-fold response to stimulation with epidermal growth factor (Figure 5B), similar to that seen with the human gastrin promoter [22], indicating that the mouse gastrin gene expression can be regulated by EGF despite the presence of a clearly defined EGF response element.

In summary, we have obtained a 20 kb clone of the murine gastrin gene containing 15 kb of the 5' promoter. The gene is very similar to the other mammalian gastrin gene's in terms of structure, with conservation of important processing sites. It can also be regulated by EGF in a similar manner to the human gastrin gene although it lacks the previously described EGF response element. The cloning of the murine gastrin gene now

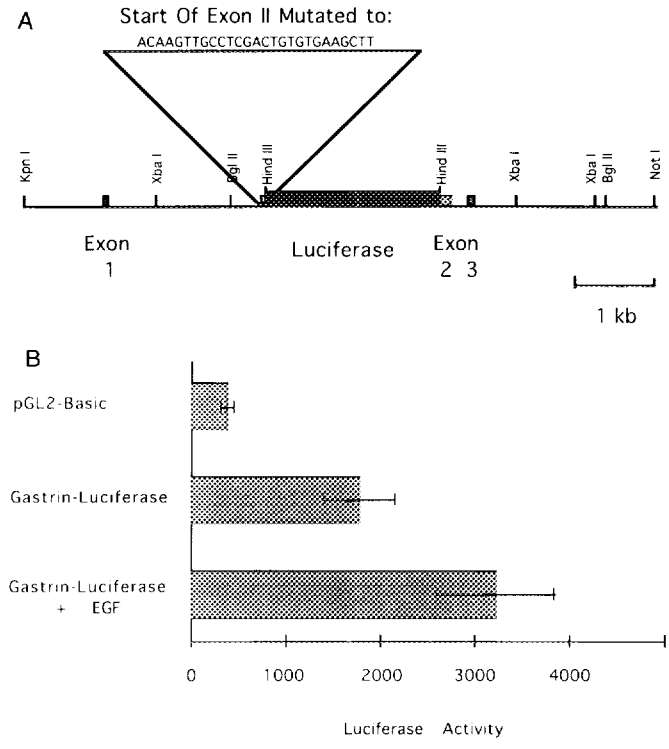

Figure 5.(A) A diagram of the mouse gastrin gene-luciferase contruct. The luciferase coding regions (shown as a black box) were inserted into the Hind III site in exon II. The initiator methionine codon of mouse gastrin was altered by PCR mutagenesis (new sequence is shown). This minigene construct contains all three exons (shown as gray boxes), two introns (flanking DNA and introns are represented by lines), 1043 bp of 5' flanking sequence and 3 kb of 3' flanking sequence. (B) Effect of EGF on a mouse gastrin-luciferase response.The minigene was transfected into a human gastric carcinoma cell line (AGS-B), pGL2-Basic (Promega) containing a promoterless luciferase reporter was used as a control. Twelve hours later, the cells were placed in serum free media and half the plated cells were exposed to EGF (10^{-9} M) for 24 hours. The cells were then harvested and luminescence measured. EGF stimulation resulted in a two-fold increase in luciferase activity.

allows the use of powerful techniques of targeted gene disruption and the creation of chimeric genes in mice to allow further study on gastrin's role in development.

ACKNOWLEDGMENTS This work was supported by research grants from the NIH (R01 DK48077) and the American Diabetes Association to TCW. TJK was supported by a NIH Training Grant (NIDDK 5T3 DK07191 D21) and by the Merck/AGA Senior Fellowship Award. Oligonucleotides were provided by the Center for Studies of Inflammatory Bowel Disease, Massachusetts General Hospital.

REFERENCES

1. Walsh J.H., (1994) In Gut Peptides (J.H. Walsh and GJ Dockray, Ed.) pp. 75-121. Raven Press, New York, NY.
2. Tielemans, Y., Axelson, J., Sundler, F., Willems, G., and Hakanson, R. (1990) Gut 31, 274-278.
3. Crean, G.P., Marshall, and Rumsey R.D. (1969) Gastroenterology 57, 147-155.
4. Botella, A., Delvaux, M., Berry, P., Frexinos, J., Bueno, L. (1992) Gastroenterology 102, 779-786.
5. Brand, S.J., and Fuller, P.J. (1988) J. Biol. Chem. 263, 5341-5347.
6. Solomon, T.E., Morisset, J. Wood, J.G., and Bussjaeger, L.J. (1987) Gastroenterology 92, 329-435.
7. Wang, T.C., Bonner-Weir, S., Oates, P.S., Chulak, M. Simon, B., Merlino, G.T., Schmidt, E.V., and Brand, S.J. (1993) J. Clin. Invest. 92, 1349-1356.
8. Langhans, N.A., Beinborn, M., Quinn, S. M., Chiu, M.L., Ardman, B., and Kopin, A.S. (1995) Gastroenterology 108, A985.
9. Seva, C., Dickinson, C.J. and Yamada, T. (1994) Science 265, 410-412.
10. Todisco, A.Takeuchi Y. Diskinson, C.J., and Yamada, T., (1995) Gastroenterology 108, A1011.
11. Wang, T.C., Babyatsky, M.W., Oates, P.S., Zhang, Z., Tillotson, L., Chulak, M., Brand, S.J., Schmidt, E.V. (1995) Am. J. Physiol. 268, G1025-G1036.
12. Fuller, P.J., Stone. D.L., and Brand, S.J. (1987) Mol. Endocrinol. 1, 306-311.
13. Huynh, T.V., Young, R.A., and Davis, R.W. (1985) In DNA Cloning (D.M. Glover, Ed.), pp.49-78. IRL Press, Oxford, U.K.
14. Davis, R.W., Botstein, D., and Roth, J.R. (1980) Advanced Bacterial Genetics, p.80. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15. Sanger, F., Nicklen, S. and Coulsen A.R. (1977) Proc. Natl. Acad. Sci. U.S.A.74, 5463-5467.
16. Chen, C. and Okayama, H. (1987) Mol. Cell Biol. 7, 2745.
17. Varro, A., Nemeth, J., Bridson, J., Lee, C., Mooret, S., and Dockray,G.J. (1990) J. Biol. Chem. 265, 21476-21481.
18. Wang, T.C., and Brand, S.J. (1990) J. Biol. Chem. 265, 8908-8914.
19. Chung, D.C., Brand, S.J., and Tilltotson L.G. (1995) J. Biol. Chem. 270, 8829-8836.
20. Tillotson, L.G., Wang, T.C., and Brand, S.J. (1994) J. Biol. Chem. 269, 2234-2240.
21. Imagawa, M., Chiu, R., and Karin, M. (1987) Cell 51, 251-260.
22. Merchant, J.L., Demediuk, B., and Brand, S.J. (1991) Mol. Cell. Biol. 11, 1686-2696.
23. Montag, A.G., Oka, T., Kwang, H.B., Chel, S.C., Jay, G. and Agarwal, K. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 6696-6700.