Characterization of rat gastric inhibitory peptide cDNA

S. K. Sharma, C. Austin*, A. Howard*, G. Lo*, C. G. Nicholl* and S. Legon

Departments of Chemical Pathology and *Medicine, Royal Postgraduate Medical School, Hammersmith Hospital, Ducane Road, London W12 ONN, U.K.

(S. K. Sharma is currently at 417 Gali Chandi Wali, Pahar Ganj, New Delhi 110 055, India)

(Requests for offprints should be addressed to S. Legon)

Received 3 June 1992

ABSTRACT

Gastric inhibitory peptide (GIP) is a 42 amino acid gastrointestinal peptide which inhibits gastric acid secretion and stimulates pancreatic insulin secretion in the presence of glucose. Here we report the sequence of the cDNA encoding the rat GIP precursor. PreproGIP was 144 amino acids in length and comprised the GIP peptide itself, N- and C-terminal flanking peptides of 22 and 59 amino acids respectively and a typical hydrophobic signal peptide. The sequence indicated that GIP is released from its precursor by cleavage at single arginine residues. The C-terminal flanking peptide may have an important function since it was well conserved and contained a region of 16 amino acids with only a single, conservative replacement. Rat GIP mRNA was found in the duodenum and jejunum. Levels of GIP mRNA in the duodenum were increased twofold after a period of 2 days of starvation. There was no detectable expression of the GIP gene in other parts of the gastrointestinal tract or in other endocrine tissues. However, in pancreatic mRNA preparations, a larger mRNA was detected after low stringency hybridization. This could represent a further member of this gene family.

Journal of Molecular Endocrinology (1992) 9, 265–272

INTRODUCTION

Gastric inhibitory peptide (GIP) is a 42 amino acid regulatory peptide produced by specific endocrine cells (K cells) in the proximal small intestine (Buffa et al. 1975). The amino acid sequences of porcine (Jornvall et al. 1981; Moody et al. 1981), bovine (Carlquist et al. 1984) and human (Moody et al. 1984) GIP are known; each is 42 amino acids in length and the sequences are highly conserved. The sequence of GIP indicates that it is a member of a family of related peptides which includes secretin, glucagon, vasoactive intestinal polypeptide, growth hormone-releasing factor, peptide histidine isoleucine and pituitary adenyl cyclase-activating protein (Bell, 1986; Miyata et al. 1989). Despite the fact that GIP was first isolated on the basis of its ability to inhibit histamine-stimulated gastric acid secretion (Brown et al. 1969), the physiological role of GIP remains unclear, as this inhibitory effect is limited (Maxwell et al. 1980). GIP is also known as glucose-dependent insulino tropic polypeptide, on the basis of its ability to enhance insulin release in the presence of elevated levels of plasma glucose (Dupre et al. 1973; Pederson et al. 1975; Pederson & Brown, 1976; Andersen et al. 1978). This effect is seen at physiological levels of GIP and it has been suggested that GIP forms a component of the incretin effect. It has also been suggested that GIP may play an important role in regulating glucagon secretion from pancreatic alpha cells (Opara & Go, 1991). GIP may therefore influence glucose homeostasis at a number of levels.

Little is known about the regulation of GIP biosynthesis. The structure of the precursor protein is known in man (Takeda et al. 1987) but not in other species. Much of the work on the localization and activity of GIP has been performed in the rat. We now report the cloning and sequencing of cDNA encoding rat preproGIP and present Northern blot analysis of the expression of the mRNA in the rat.

MATERIALS AND METHODS

Amplification of GIP-coding sequence

Oligonucleotides were made on an Applied Biosystems 380B oligonucleotide synthesizer (Applied
Biosystems Inc., Foster City, CA, U.S.A.). Primers for amplification of the coding sequence were based on the human and pig amino acid sequences close to the ends of the GIP peptide (see Fig. 1c). They covered all possible codon choices. The 5' primer was a mixture of 256 species (5'M: GCNGARGGNACNTTYYAT), the 3' primer was a mixture of 16 species (3'M: ATRTTTRTGYTTCCARTC) and both included a further nine bases to provide an EcoRI restriction site. The cDNA for amplification was generated by random-primed reverse transcription (Super RT; Anglian Biotechnology, Colchester, Essex, U.K.) of 1 µg duodenal mRNA in a reaction of 20 µl, and 4 µl of this reaction were boiled in 30 µl water for 5 min. The polymerase chain reaction (PCR) was then performed in a volume of 100 µl containing 20 µl boiled cDNA, 12 µg 5'M primer, 3 µg 3'M primer, 200 µM of each dNTP, 2·5 units Taq DNA polymerase (Cetus Corp., Emeryville, CA, U.S.A.) and the buffer supplied with the enzyme. PCR conditions were as follows: 30 cycles at 94 ºC for 0·5 min, 55 ºC for 3·5 min and 72 ºC for 0·5 min. Reaction products were fractionated on a 1·4% agarose gel, spin-eluted (Heery et al. 1990) and ethanol-precipitated. After digestion with EcoRI, the DNA was cloned into M13 mp18 and recombinants were prepared for sequencing using standard techniques (Sambrook

![Diagram](image)

**Figure 1.** Amplification and sequencing of gastric inhibitory peptide (GIP) cDNA. (a) Strategy for amplifying the GIP-coding sequence and the 3' and 5' ends of the cDNA. The boxed section of the mRNA represents the region coding for the GIP peptide. Mixed-sequence primers 5'M and 3'M were used to amplify the GIP peptide region. The nested specific primers 5'S and 3'S were used in conjunction with the linker-dT primer (LdT) to amplify the 3' and 5' ends of the cDNA respectively. (b) Nucleotide sequence of preproGIP cDNA. The coding sequence starts at the ATG triplet at base 108 and terminates at the TGA triplet at base 540. (c) Comparison of the predicted amino acid sequence of rat preproGIP with that of man and the porcine and bovine GIP sequences. The GIP region is boxed. The rat sequence has been hyphenated to maximize its homology with the human sequence.

Journal of Molecular Endocrinology (1992) 9, 265–272
Sequencing was performed using the extended Klenow chain-termination method (Stambaugh & Blakesley, 1988).

**Amplification of cDNA ends**

The rest of the GIP cDNA was amplified using an anchor PCR technique, essentially as described by Frohman et al. (1988). The strategy and primers are indicated in Fig. 1.

For the 3′ end, mRNA was reverse-transcribed and amplified as above, but using oligo(dT) to prime reverse transcription. Primers for the first round of 3′ amplification were 5′G (120 μg/ml) and linker-dT primer (LdT: GTGGAATTCTCGAGTCGACT TTTTTTTTTTT, 2.5 μg/ml). Amplification conditions were as above except that annealing was for 3 min. A tenfold dilution of this reaction (20 μl) was then amplified using the 5′ specific primer (5′S, 7.5 μg/ml) and LdT (2.5 μg/ml) with reaction conditions as above. For the 5′ end, random-primed cDNA was ethanol-precipitated twice and tagged with dATP. A reaction of 20 μl contained cDNA from 1 μg mRNA, 60 μM dATP, 1.5 mM CoCl2, 30 units terminal transferase (Boehringier Mannheim, Mannheim, Germany) and the buffer provided by the supplier, and was incubated for 15 min at 37 °C. After ethanol-precipitation, the tagged cDNA was amplified as for the 3′ end. Primers for the first round of 5′ amplification were 3′M (30 μg/ml) and LdT (2.5 μg/ml), and for the second round 3′ specific primer (3′S, 7.5 μg/ml) and LdT (2.5 μg/ml).

**Cloning and sequencing of cDNA ends**

PCR reaction products (10 μl) were fractionated on a 1% agarose gel, blotted and hybridized with the GIP-coding-sequence probe. Regions corresponding to the hybridizing region were then eluted from a gel containing the remainder of the reaction, digested with EcoRI and cloned into M13 as for the GIP-coding region. Plaques containing the GIP cDNA ends were identified by plaque-lift hybridization using the coding-region probe. Sequencing was as above for the 3′ end, but was performed following a protocol using Taq DNA polymerase for the 5′ end (Brow, 1990). Computer analysis was performed using the IBI/Pustell DNA sequence analysis programs (IBI Ltd, New Haven, CT, U.S.A.).

**RNA analysis**

Tissues from adult male Wistar rats were snap-frozen in liquid nitrogen and stored at −70 °C before extraction. For cloning, RNA was prepared from fed animals. RNA was also prepared from two groups of six rats (weighing approximately 100 g), one group having been starved for 2 days. Total RNA was prepared using the guanidinium isothiocyanate/acid phenol method (Chomczynski & Sacchi, 1987) and then subjected to chromatography on oligo(dT) celluloose to prepare polyadenylated RNAs. RNA was fractionated on formaldehyde/agarose gels, blotted onto Hybond-N nylon membranes (Amersham International plc, Amersham, Bucks, U.K.) and hybridized with cDNA probes as previously described (Howard et al. 1992). The rat glucagon and neuromedin U probes were labelled by random-primed synthesis (Feinberg & Vogelstein, 1983); the probes are described by Heinrich et al. (1984) and Lo et al. (1992), Genbank accession numbers K02808 and M94555 respectively. GIP probes were similarly labelled but using an oligonucleotide primer specific for each probe. To quantitate the results of the starvation experiment, after probing with the 5′ GIP cDNA probe, the filter was counted in an Autograph using Autosoft software (Oxford Positron Systems, Oxford, Oxon, U.K.). The blot was stripped at 80 °C in a buffer containing 10 mM Tris–HCl, pH 7.6, 100 μM EDTA and 0.2% (w/v) sodium dodecyl sulphate (SDS), and reprobed with a probe representing the 3′ end of the cDNA. It was then stripped again and reprobed with a 30 base oligonucleotide complementary to the 3′ end of 18S rRNA. The blot was recounted and the figures for 5′ GIP hybridization were adjusted for slight differences in loading by dividing by the rRNA figures. For the low stringency experiment, the hybridization temperature was reduced from 60 °C to 42 °C and washing was performed at 50 °C in a buffer containing 0.25 M sodium phosphate, pH 7.2, 1 mM EDTA and 2% SDS.

**RESULTS**

In the absence of any information about rat GIP, we made the assumption that it would be similar to the human and porcine peptides. Two mixed-sequence PCR primers were made, based on conserved sequences of six amino acids at the N and C termini of human and porcine GIP (Fig. 1c; the bovine sequence was not known to us at that time). These were then used to amplify cDNA prepared from rat duodenal mRNA. When analysed on an agarose gel, the reaction products gave a single major band which was cloned into M13 for sequencing. The sequence between the primers coded for a peptide having only a single difference from the corresponding region of human or porcine GIP (Fig. 1c).

The scheme used to amplify the remainder of the cDNA sequence was based on the anchor PCR strategy of Frohman et al. (1988) illustrated in Fig. 1a.
The sequence of the GIP-coding sequence between the mixed primers was used to provide two further primers for the amplification of the 5' and 3'-ends, as the mixed primers did not provide sufficient specificity for this method. For the 3'-end amplification, rat duodenal cDNA was prepared using the LdT primer to prime from the poly(A) sequence. This primer was then used in combination with the original 5'M primer to amplify from the GIP-coding sequence to the 3' end. This did not provide sufficient specificity to give a discrete band on an agarose gel. However, when this material was reamplified using the 5'S and LdT primers, a major band was visible. Aliquots of both reactions were analysed on an agarose gel and a Southern blot was probed with the cloned GIP-coding sequence obtained previously. Although the first round of PCR did not provide sufficient specificity to give a visible band, hybridization to a region of approximately 400 bases was seen. The second round of PCR gave a distinct band which hybridized strongly with the probe. Prolonged exposure did not reveal any further bands which might represent the use of alternative polyadenylation sites. The remainder of the second-round reaction was then run on an agarose gel, eluted, cloned and sequenced (Fig. 1b). Four independent clones were analysed, but these differed only in orientation and in the number of adenosine residues in the poly(A) tail.

For the 5'-end amplification, duodenal mRNA was reverse-transcribed with random primers. After removal of excess primers and nucleotides, a poly(A) tail was added using terminal transferase. The tailed cDNA was then amplified using the same strategy as had been used for the 3' end (Fig. 1a). On Southern blotting, GIP-specific products were seen at about 350 bp, with some hybridization to a diffuse region representing shorter products. The major species was eluted, cloned and sequenced as for the 3' end. Five clones were sequenced and showed length heterogeneity at the 5' end, indicating a problem in the generation of full-length cDNAs. The first 12 bases were derived from one clone only and may therefore be subject to PCR errors. The cDNA sequence of 654 bases is shown in Fig. 1b. There was an ATG triplet at position 108 followed by an open reading frame of 144 amino acids which included the sequence of GIP itself (amino acids 44–85).

The distribution of rat GIP mRNA was studied by using the cloned 3'-end cDNA as a probe for Northern blots of rat tissues. Within the gastrointestinal tract, expression was detected at high levels in the duodenum and jejunum, with no signals in other regions (Fig. 2a). The blot was rehybridized with probes for rat glucagon (1084 bases plus poly(A); Heinrich et al. 1984) and neuromedin U (707 bases plus poly(A); Lo et al. 1992). The rat GIP band had the same mobility as neuromedin U, indicating that the GIP cDNA sequence (654 bases plus poly(A)) was close to full length. On further blots, mRNA from rat whole brain, hypothalamus, pancreas, pituitary, spinal cord and thyroid was analysed in a similar fashion, with negative results (not shown). However, when these mRNAs were hybridized and washed under conditions of low stringency, a band was seen in the mRNA from pancreatic tissue. In further experiments, we demonstrated that this mRNA migrated between glucagon mRNA and 18S rRNA (Fig. 2b), and that it could be detected with the probes for the GIP-coding sequence and the 5'-end sequence described above (not shown). Finally, as we intend to use the GIP probe in rat models of diabetes, we wished to demonstrate that the expression of the GIP gene responded to physiological changes. It is known that rat duodenal GIP peptide levels are elevated after starvation (Shulkes et al. 1983). We measured GIP mRNA levels in the duodenum after a period of 2 days of starvation and found double the level of GIP mRNA in the starved group using the 5'-end probe (Fig. 3). The figure for the control group (mean ± S.D.) was 5.22 ± 0.98 and that for the starved group was 10.42 ± 1.83 (P = 0.0039; Mann–Whitney U Wilcoxon rank sum test, two-tailed). The blot was stripped and reprobed with the 3'-end probe, to check that the full cDNA sequence was represented in the band, and identical results were obtained (Fig. 3).

DISCUSSION

We have characterized cDNA encoding the precursor of rat GIP using the combination of mixed primer and anchor PCR developed in this laboratory during the cloning of neuromedin U (Lo et al. 1992). This approach is particularly appropriate for the study of regulatory peptide cDNAs. Conservation between species of short, biologically active regions means that mixed-sequence PCR primers can be used to amplify cDNAs from a variety of species. Having determined the nucleotide sequence between the mixed primers, this sequence can then be used to design primers for anchor PCR, enabling a longer sequence to be isolated. However, this approach has the disadvantage that it is necessary to sequence a number of independent clones to guard against the possibility of PCR errors. In current studies we are using direct sequencing of PCR products to bypass this problem and to shorten the procedure further. With long cDNAs, the alternative of isolating a conventional cDNA clone might be preferable.
FIGURE 2. Expression of the rat gastric inhibitory peptide (GIP) gene. (a) Upper panel: polyadenylated RNA (10 μg) from oesophagus (lane 1); stomach (lane 2); duodenum (lane 3); jejunum (lane 4); ileum (lane 5); caecum (lane 6); colon (lane 7) and rectum (lane 8) was analysed by Northern blotting with the 3' GIP cDNA probe. Lane 9 contained 0.5 μg pancreatic mRNA. O = origin, G = glucagon mRNA marker, N = neuromedin U mRNA marker. Lower panel: same blot reprobed for glucagon, origin region not shown. (b) Lane 1, pancreatic mRNA (10 μg) probed at low stringency with 3' GIP cDNA; lane 2, same blot, stripped and reprobed for glucagon mRNA. O = origin; the position of the 18S rRNA marker is indicated.

The nucleotide sequence of rat preproGIP and the corresponding deduced amino acid sequence are shown in Fig. 1b and c. The GIP cDNA had a single open reading frame encoding a 144 amino acid protein. The coding region was flanked by a 5' untranslated region of 107 nucleotides and a 3' non-coding region of 121 bases, with a typical polyadenylation signal 15 bases from the start of the poly(A) tail. The sequence is likely to be almost complete, as the overall size of 654 bases corresponds with that estimated by Northern blot analysis (approximately 700 bases plus poly(A); Fig. 2a). However, when following the effects of starvation on GIP gene expression, we noted that the increased expression level was accompanied by a noticeable broadening of the band (Fig. 3). This might have indicated the induction of a novel type of GIP-related mRNA, as has been seen, for instance, in the calcitonin/calcitonin gene-related peptide gene (Amara et al. 1982). However, we reprobed the blot with the 3' GIP cDNA probe and observed the same effect, ruling out the possibility of a second mRNA sharing only N-terminal coding sequences with GIP mRNA. We have not investigated this phenomenon further; it may be caused by heterogeneity in the length of the poly(A) tail, although other possibilities clearly exist. Comparison of the amino acid sequence with that of human preproGIP shows a very similar structure (Fig. 1c). As in man, the rat precursor protein has a hydrophobic signal peptide with a probable cleavage site at glycine-21. The remaining proGIP consists of GIP itself, with N- and C-terminal flanking peptides of 22 and 59 amino acids respectively. The sequence indicates that rat GIP, like its human equivalent (Takeda et al. 1987), is released from its precursor by cleavage at single arginine residues. The conservation of the sequence around these potential monobasic cleavage sites in both human
and rat precursors suggests that the actual signal may be more extensive.

Like its porcine, bovine and human equivalents, rat GIP is 42 amino acids in length. Rat GIP differs from the others in just two (human and porcine) or three (bovine) amino acids (Fig. 1c). There is 63% homology between the rat and human N-terminal flanking peptides, though optimal alignment requires the deletion of eight amino acids from the human sequence. The rat C-terminal flanking peptide is 59 amino acids in length compared with 60 in man, and there is 71% homology between the two, including a region of 16 amino acids with only a single conservative replacement. This level of conservation suggests a significant function for this region, perhaps in the processing of the precursor, or possibly as an independently secreted peptide.

In the case of the pro-opiomelanocortin gene, different peptides are produced by different cell types (Hinman & Herbert, 1980; Rosa et al. 1980). If proGIP were similarly processed, then the GIP gene might be expressed in cells or tissues containing only small amounts of the GIP peptide itself. We therefore performed Northern blot analysis to investigate this possibility. GIP mRNA was readily detectable in the duodenum and jejunum, with no expression elsewhere in the gastrointestinal tract (Fig. 2a). Although duodenum and jejunum are the regions containing most immunoreactive GIP (Polak et al. 1973; Buffa et al. 1975), smaller amounts have been reported in the stomach, ileum and colon (O'Dorisio et al. 1977; Alumets et al. 1978). However, in our studies, GIP mRNA was not detected in these tissues (Fig. 2).

There have also been reports of GIP-like immunoreactivity in the pancreas (Smith et al. 1977; Alumets et al. 1978; Erlandsen, 1980; Ahren et al. 1981). However, Sarson (1982) reported that no GIP peptide could be extracted from the pancreas, and other workers were unable to detect pancreatic GIP-like immunoreactivity (Larsson & Moody, 1980; Buchan et al. 1982). The situation was clarified by the work of Sjolund et al. (1983) who used a panel of six anti-GIP antibodies. Three of the antibodies were specific to cells in the small intestine whilst the other three also stained pancreatic A cells. Control experiments indicated that the GIP-like immunoreactivity in the pancreas was not due to either GIP or glucagon. They concluded that the glucagon-producing cells of the pancreas also synthesize a protein which cross-reacts with some of their anti-GIP antisera. Our data indicate that there may be a GIP-like mRNA in the pancreas. When we screened pancreatic and other endocrine tissues, GIP mRNA was not detected under normal hybridization conditions. However, when pancreatic tissue was screened for GIP mRNA at low stringency, a cross-hybridizing RNA, slightly larger than glucagon, was detected (Fig. 2b). Of the known members of the GIP/glucagon gene family, GIP mRNA is closest in sequence to glucagon mRNA itself, with 64% homology over an 87 base region. However, the cross-hybridizing RNA seen in the pancreas is clearly not glucagon mRNA (Fig. 2b). Further studies will be required to elucidate the nature of this species, which may represent a further member of the glucagon gene family.

ACKNOWLEDGEMENTS

This work was supported by the British Diabetic Association. S. K. S. was in receipt of a fellowship from the British Council and was a participant in the Royal Postgraduate Medical School Molecular Biology Training Scheme. We thank Prof. J. Messing, University of Minnesota, MN, U.S.A. and Dr P. Lund, University of North Carolina, NC, U.S.A. for their gifts of M13 mp18/19 and glucagon cDNA clones respectively. We also thank M. Millbourn, M. Thorpe and Dr D. Bretherton-Watt for their help with the animal studies. D. J. Liyanage for the oligonucleotide synthesers and Prof. S. R. Bloom for his helpful comments on this manuscript. This sequence has been submitted to the E.M.B.L. Data
REFERENCES


