Vol. 193, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
May 28, 1993 Pages 182-190

ISOLATION AND CHARACTERIZATION OF THE GENE ENCODING RAT GLUCOSE-DEPENDENT INSULINOTROPIC PEPTIDE

Yoshifumi Higashimoto and Rodger A. Liddle

Department of Medicine, Duke University Medical Center and Department of Veterans Affairs, Durham, North Carolina

Received April 8, 1993

The rat glucose-dependent insulinotropic peptide (GIP) gene has been isolated and characterized . The gene spans approximately 8.2 kilobase pairs (kb) and the GIP mRNA (0.8 kb) is encoded by six exons. The 42 amino acid hormone is encoded by exons 3 and 4. The exon-intron organization of the rat GIP gene revealed that the splice acceptor site for intron 2 is 24 nucleotides downstream compared to the comparable splice acceptor site in the human gene. This intron sliding results in an 8 amino acid deletion in the amino terminal extension of the prepropeptide. Primer extention analysis and RNase protection assay demonstrated the existence of multiple closely spaced sites for transcriptional initiation. Both the 5'-flanking region and intron 1 contain TATA and CCAAT boxes consistent with initiation of gene transcription, although a TATA box in intron 1 is functionally inactive in adult rats in spite of its reasonable location.
© 1993 Academic Press, Inc.

Glucose-dependent insulinotropic peptide (GIP) was first isolated from porcine small intestine and originally called gastric inhibitory peptide because of its ability to inhibit gastric acid secretion (1,2). Subsequent physiological studies indicated that this effect was observed only at supraphysiologic levels (3). However, GIP was found at physiological concentrations to stimulate insulin release in the presence of elevated glucose concentrations (4) and it is believed to be the most important incretin regulator of the gastrointestinal tract.

GIP consists of 42 amino acid residues in porcine (5), human (6), and bovine (7) intestine and the amino acid sequence is well conserved among species. Recently, cDNA's encoding human (8) and rat (9) GIP were isolated and characterized, indicating that both are derived by proteolytic processing of precursors consisting of 153 and 144 amino acids, respectively. The deduced amino acid sequence of rat GIP differs from the human sequence by two amino acid substitutions: arginine for histidine at position 18 and leucine for isoleucine at position 40.

0006-291X/93 $4.00
Copyright © 1993 by Academic Press, Inc.
All rights of reproduction in any form reserved.

Secretion of GIP from small intestine is regulated by dietary factors. Both ingestion of a mixed meal and oral glucose administration stimulate GIP release (10,11). Interestingly, starvation also has been associated with an increase in intestinal concentrations of GIP in rats (12) and plasma GIP levels in humans (13). Furthermore, we recently suggested that starvation might also increase intestinal GIP mRNA levels in rats (14). Although many physiological studies of GIP have been conducted in the rat, the gene encoding rat GIP has not been cloned. In order to futher investigation of GIP gene regulation, we have isolated and characterized the rat GIP gene and its promoter region.

Materials and Methods

Isolation and Sequencing of the Rat GIP Gene

Rat GIP genomic clones were isolated from a rat genomic library constructed in lambda dash II vector (Stratagene) by plague hybridization using a ^{32}P-labeled rat GIP cDNA probe (9). In order to obtain a restriction map of the gene, the cloned DNAs were digested with different restriction enzymes and Southern blot analyses were performed. Appropriate DNA restriction fragments were then subcloned into the pBluescript II SK$^+$ vector (Stratagene), and sequenced by the chain-termination method using modified T7 polymerase (USB). Since our previously isolated cDNA (9) was known to contain only 13 bases of exon 1, after sequencing exon 2, a synthetic antisense oligonucleotide of exon 1 (5'-CCAACAGCTCTT-3') was used for determining the location of exon 1 in the cloned DNAs.

Primer Extension Analysis and RNase Protection Assay

The transcriptional initiation sites of the GIP gene were determined by primer extension analysis (15) and RNase protection assay (16). A 27-mer primer, 5'-CCAACAGCTCTTCTTAGCTCTTCCAGG-3', complementary to nucleotides 52-78 of exon 1, was synthesized and end-labeled with [γ-^{32}P] ATP and T4 polynucleotide kinase. The labeled primer ($1x10^5$cpm) was annealed to total RNA (50 µg) isolated from rat small intestine (17) for 3 h at 50°C in 10 µl of 300 mM NaCl/10 mM TrisCl, pH 7.5/2 mM EDTA. The annealed RNA/primer mixture was added to 40 µl of reverse transcriptase buffer (1.25 mM dNTP's, 12.5 mM DTT, 12.5 mM TrisCl, pH 7.5, 7.4 mM $MgCl_2$, and 75 µg/ml actinomycin D) and 1 U of RNase Block II (Stratagene), and then extended for 80 min at 37°C using 100 U of M-MLV reverse transcriptase (GIBCO, BRL). The primer-extended cDNA products were analysed by electrophoresis on an 8% urea-polyacrylamide gel in parallel with a sequencing ladder generated with the same primer on the appropriate template.

A Pst I/Sma I restriction fragment of the GIP gene (Fig. 1) was subcloned into the Bluescript vector and linearized by Eco R1 digestion. An RNA probe was generated using [γ-^{32}P] UTP and T3 RNA polymerase (Stratagene). The probe consisted of a sequence complementary to a 288-nucleotide Pst I/Sma I restriction fragment of the rat GIP gene plus 61 nucleotides of the vector. Total RNA samples (30µg) of rat small intestine were hybridized overnight at 46°C with the antisense RNA probe ($5x10^5$ cpm). The hybrids were digested by RNase

A (40 μg/ml) and RNase T1 (2 μg/ml) at 30°C for 45 min and the sizes of the protected fragments were determined in a similar manner to the primer extension analysis.

In order to examine whether a TATA box in intron 1 was functionally active or not, another RNase protection assay was performed. An RNA probe was used containing a sequence complementary to a 303-nucleotide Sma I restriction fragment of rat GIP cDNA including 13 bp of exon 1, 108 bp of exon 2, 147 bp of exon 3, and 35 bp of exon 4, as well as 56 nucleotides of the vector. The hydridization of total RNA samples (30 μg) from adult rat small intestine with a ^{32}P-labeled cRNA probe and digestion by RNases A and T1 were performed in the same manner as described above. The resulting product was analysed on a 6% urea-polyacrylamide gel along with a α-^{35}S labeled 1 kb DNA ladder (GIBCO, BRL).

Results and Discussion

The rat GIP cDNA (9) was used to screen 8 x 10^5 plaques of a rat genomic library. Two plaques hybridized with the probe yielding separate clones with inserts ~18 kb in length. These were isolated and partially sequenced after subcloning into the plasmid vector (Fig. 1, 2) revealing clones with overlapping sequences (λrGIP5-1, λrGIP5-2) (Fig. 1). The rat GIP gene has six exons and spans approximately 8.2 kb (Fig. 1). It is ~1.8 kb shorter than the human gene (18). The nucleotide sequences of the rat GIP cDNA (9) were identical to those of the GIP exons (Fig. 2). The first intron interrupts the gene region encoding the 5'-untranslated region of the mRNA. The second intron interrupts the region

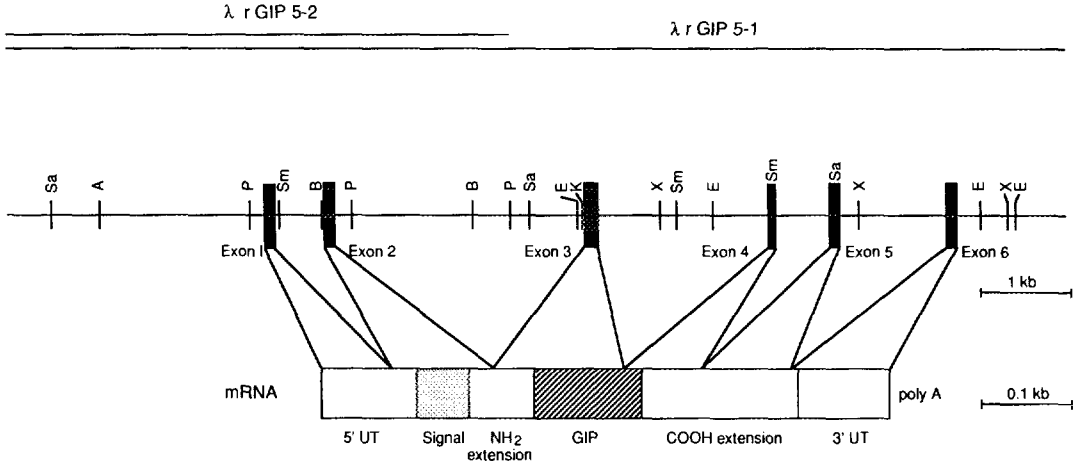


Figure 1. Structure of the rat GIP gene. Inserts within the genomic clones are indicated on the upper portion. The middle and lower portions represent the relationship of each exon to the mature mRNA. The restriction sites are: A, Apa I; B, Bam H1; E, EcoR1; K, Kpn I; P, Pst I; Sa, Sac I; Sm, Sma I; and X, Xba I.

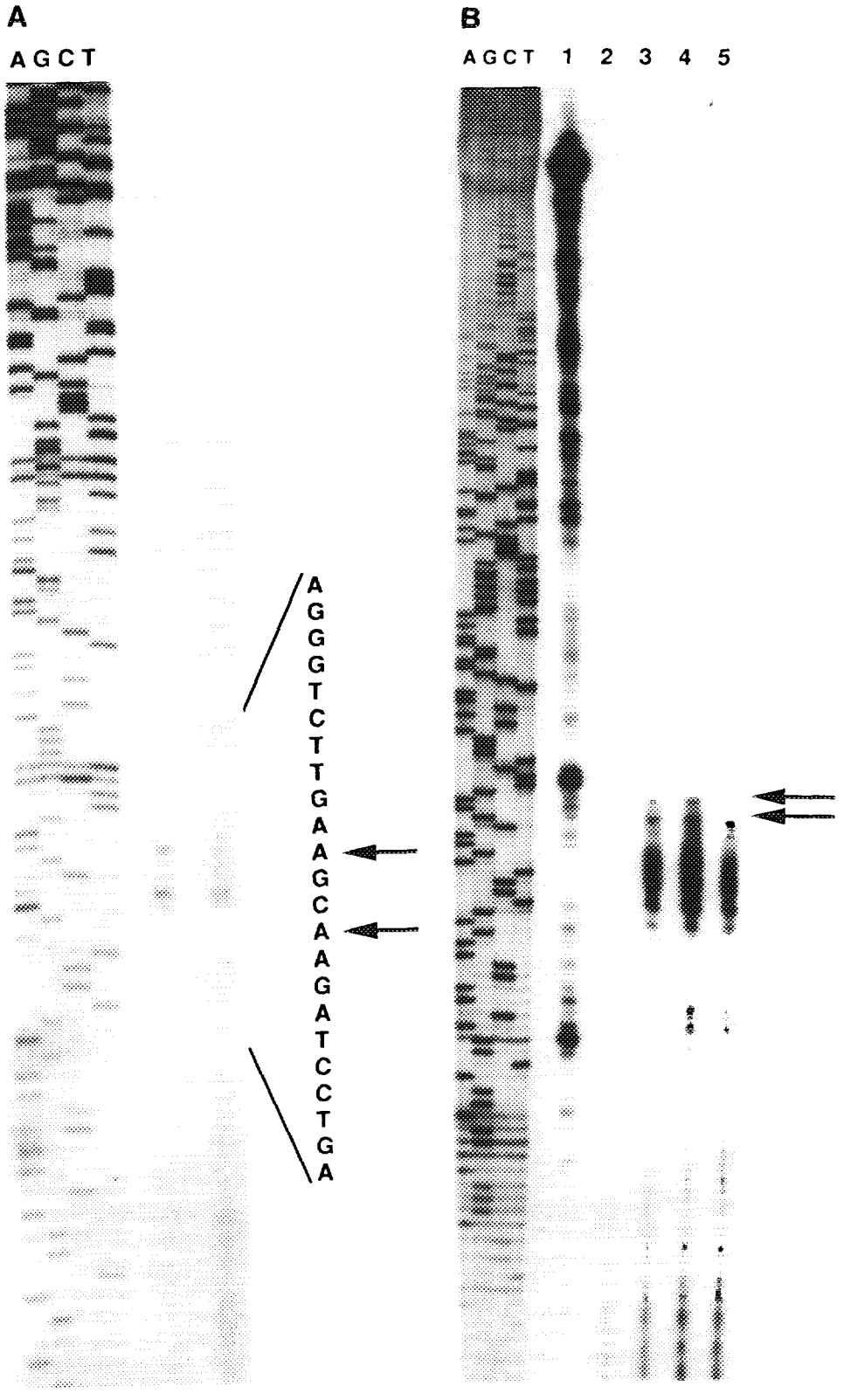
Figure 2. Nucleotide sequence of the rat GIP gene and its flanking regions. Exons are shown in capital letters; introns and flanking sequences are shown in lower case letters. Nucleotides are numbered relative to the transcriptional start site. TATA box and CCAAT box motifs are printed in bold type. Potential binding sites are underlined with the appropriate designation underneath. AT-rich regions in the 5'-flanking region and a possible YGTGTTYY consensus in the 3'-flanking region are underlined. The polyadenylation signal, AATAAA, is in underlined, bold type. A nucleotide sequence complementary to that indicated by an arrow was used as a primer in primer extension analysis.

encoding the N-terminal extension peptide . The third intron interrupts the region encoding the mature GIP, most of which is encoded in exon 3. The fourth and last introns are located in the region coding for the C-terminal extension peptide.

The transcriptional initiation site was determined by primer extension analysis (Fig. 3A). The major extension products were 78 and 75 nucleotides long and were located at 27 and 30 nucleotides downstream, respectively, from the first T of a consensus TATA box. RNase protection assay confirmed this result and, in addition, demonstrated multiple initiation sites including others located 6-13 nucleotides further downstream (Fig. 3B).

Potential regulatory elements were identified in the 5'-flanking region and intron 1 (Fig. 2). Putative TATA boxes [consensus TATA($^{A}/_{T}$)A($^{A}/_{T}$)] (19) are located at positions -115, -27, and +755; CCAAT-like boxes [consensus GG($^{C}/_{T}$)CAATCT] (19) at positions -171, -158, and +599 relative to the initiation start site. Potential AP-1 [consensus TGA($^{G}/_{C}$) TCA(G)] (20,21) and AP-2 [consensus ($^{T}/_{C}$)C($^{C}/_{G}$)CC($^{A}/_{C}$) N($^{G}/_{C}$)($^{C}/_{G}$)($^{G}/_{C}$)] (22) sites were found at five sites (positions -842, -801, -411, -53, and +624) and five sites (positions -813, -751, +153, +688 and +745), respectively. Putative cAMP response elements (CRE, consensus TGACGTCA) (23) are present at positions -379, +35, and +226. The sequence CGTCA, found at position +226 in intron 1, is essential for biological activitiy of the cAMP-regulating enhancer in the gene encoding vasoactive intestinal peptide (24) which together with GIP is a member of the glucagon superfamily of gastrointestinal peptides (25). In addition, there are sequences homologous with those of the enhancer core [consensus (G)TGG($^{A}/_{T}$)($^{A}/_{T}$)($^{A}/_{T}$)(G)] (26) and a common cellular transcription factor, ATF [consensus GTGACGT($^{C}/_{A}$)($^{G}/_{A}$)] (27), in the 5' flanking region (at site -133) and intron 1 (at site +509).

Figure 3. Determination of the transcriptional initiation site of the rat GIP gene. (A) Primer extension analysis. A ^{32}P-end-labeled 27-base synthetic antisense oligonucleotide primer corresponding to nucleotides +52 to +78 (Fig. 2) was annealed to total RNA samples (50 μg) of rat small intestine and then extended with M-MLV reverse transcriptase as described in materials and methods. The extended products are shown on the right by bold arrows. The sequence ladder obtained using the same oligonucleotide as a sequencing primer is shown in adjacent lanes. (B) RNase protection assay. An RNA probe ($5x10^5$ cpm) containing 288 nucleotides complementary to the gene sequence including the 3'-end of exon 1 was hybridized with total RNA samples (30 μg) from 3 rat small intestines (lanes 3-5) and 30 μg Yeast tRNA (lane 2) before being digested with RNases A and T1. The products were analysed on an 8% polyacrylamide-urea gel. The same sequence ladder as that for the primer extension analysis is shown in adjacent lanes. Two longer protected products indicated by arrows are apparently longer by one nucleotide than two major extended products obtained in primer extension analysis. Since RNA has a slightly lower mobility than DNA of the same length, the two protected products are considered to be the same in size as the extended products. Furthermore, many shorter protected fragments are observed. Lane 1, the undigested RNA probe.


A
A G C T
B
A G C T 1 2 3 4 5
A
G
G
G
T
C
T
T
G
A
A
G
C
A
A
G
A
T
C
C
T
G
A

Interestingly, in the 5'-flanking region, there are six AT-rich regions (11-22 bp) at positions -904, -886, -869, -668, -645, and -629. AT-rich sequences are thought to play a role in melting double-stranded DNA (28). Towards the 3' end of the gene a consensus sequence YGTGTTYY(Y; pyrimidine), an important requirement for efficient formation of mRNA 3' termini (29), is found 28 bp downstream from the AATAAA signal.

Because intron 1 possesses some potential cis-acting elements and a TATA-like box is located at a reasonable position, 27 nucleotides upstream from the beginning of exon 2 another RNase protection assay (Fig. 4) was performed

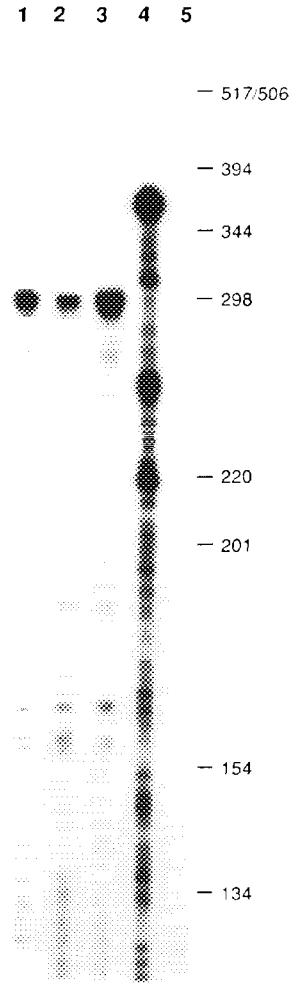


Figure 4. RNase protection assay. An RNA probe ($5x10^5$ cpm) including a sequence complementary to the 303-nucleotide Sma I fragment of rat GIP cDNA (9) containing 13 nucleotides of exon 1 was hybridized with 30 µg total RNA from adult rat small intestine (lanes 1-3) and yeast tRNA (lane 5). The procedure was performed in the same manner as the RNase protection assay shown in Fig. 3. One major protected fragment (~300 nucleotides) was obtained, but there was no protected fragment around 290 nucleotides, indicating no transcript generated downstream to exon 1 in adult rats. Lane 4, shows the undigested RNA probe.

```
        Exon 2 (Splice Donor Site)

                              (Ar
        GluLysLysGluGlyHisPheSe
Human   GAGAAGAAAGAGGGTCACTTCAGgtaggagtgg
        •• ••  •••••• • • ••• •••••   ••
Rat     GAAAAAGAAGAGGTTGAGTTCCGgtagagttgc
        GluLysGluGluValGluPheAr
                              (Ar


        Exon 3 (Splice Acceptor Site)

                                          g)
                    rAlaLeuProSerLeuProValGlySerHisAla
Human   ttgtttccagCGCTCTCCCCTCCCTGCCTGTTGGATCTCATGCT
             • •   • • ••     •   ••• •  •••• ••••••
Rat     actggtaccctggtatctgtctctcacctctcagATCCCATGCT
                                           gSerHisAla
                    gGlyIleCysLeuSerProLeuArg)
```

Figure 5. Comparison of intron arrangement within intron 2 of the human and rat GIP genes. The nucleotides within intron 2 are shown in lower case letters, and the amino acid sequences in the N-terminal extension peptides encoded by exons 2 and 3 are indicated. The predicted sequences of the rat and human N-terminal extension peptides are shown in brackets if the same acceptor splice sites were used. A dot indicates nucleotide identity.

to examine whether the TATA box in intron 1 is functionally active. Examination of rat small intestinal RNA revealed that the major protected fragment is approximately 300 nucleotides long, the size expected for a full-length transcript (303 nucleotide). However, there specifically was not a smaller protected fragment 13 nucleotides shorter than the 303 nucleotide fragment. This finding indicates that no transcript is generated from the last five exons of the gene excluding exon 1 in adult rats even though an intact peptide could be produced by transcription starting at exon 2.

We previously pointed out that the region encoding the rat N-terminal extension peptide lacks 24 consecutive nucleotides coding for 8 amino acids that are present in the human preprohormone (9). Interestingly, the exon-intron organization in the rat GIP gene revealed that the splice acceptor site for intron 2 has moved 24 nucleotides downstream compared with the human gene, although both genes use the same splice donor sites for intron 2 (Fig. 5). A similar relation of intron sliding has been observed in the human (30) and rat (31) genes encoding preprogrowth hormone-releasing factor which is also a member of the glucagon superfamily of peptides (25).

References

1. Brown, J.C., Pederson, R.A., Jorpes, J.E., and Mutt, V. (1969) Can. J. Physiol. Pharmacol. 47, 113-114.

2. Brown, J.C., Mutt, V., and Pederson, R.A. (1970) J. Physiol. 209, 57-64.
3. Maxwell, V., Shulkes, A., Brown, J.C., Solomon, T.E., Walsh, J.H., and Grossman, M.I. (1980) Dig. Dis. Sci. 25, 113-116.
4. Dupre, J., Ross, S.A., Watson, D., and Brown, J.C. (1973) J. Clin. Endo. Metab. 37, 826-828.
5. Jornvall, H., Carlquist, M., Kwauk, S., Otte, S.C., McIntosh, C.H.S., Brown, J.C., and Mutt, V. (1981) FEBS Lett. 123, 205-210.
6. Moody, A.J., Thim, L., and Valverde, I. (1984) FEBS Letters172, 142-148.
7. Carlquist, M., Maletti, M., Jornvall, H., and Mutt, V. (1984) Eur. J. Biochem.145, 573-577.
8. Takeda, J., Seino, Y., Tanaka, K.-I., Fukumoto, H., Kayano, T., Takahashi, H., Mitani, T., Kurono, M., Suzuki, T., Tobe, T., and Imura, H. (1987) Proc. Natl. Acad. Sci. USA 84, 7005-7008.
9. Higashimoto, Y., Simchock, J., and Liddle, R.A. (1992) Biochim. Biophys. Acta 1132, 72-74.
10. Pederson, R.A., Schubert, H.E., and Brown, J.C. (1975) Diabetes 24, 1050-1056.
11. Falko, J.M., Crockett, S.E., Cataland, S., O'Dorisio, T.M., Kramer, W., and Mazzaferri, E.L. (1980) Clin. Endocrinol. 13, 587-593.
12. Shulkes, A., Caussignac, Y., Lamers, C.B., Solomon, T.E., Yamada, T., and Walsh, J.H. (1983) Aust. J. Exp. Biol. Med. Sci. 61, 581-587.
13. Willms, B., Ebert, R., and Creutzfeldt, W. (1978) Diabetologia14, 379-387.
14. Higashimoto, Y., Simchock, J., and Liddle, R.A. (1992) Gastroenterology102, A736.(Abstract)
15. Boorstein, W.R. and Craig, E.A. (1989) Meth. Enzymol. 180, 347-369.
16. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
17. Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156-159.
18. Inagaki, N., Seino, Y., Takeda, J., Yano, H., Yamada, Y., Bell, G.I., Eddy, R.L., Fukushima, Y., Byers, M.G., Shows, T.B., and Imura, H. (1989) Mol. Endocrinol. 3, 1014-1021.
19. Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349-383.
20. Lee, W.P., Mitchell, P., and Tijan, R. (1987) Cell 49, 741-752.
21. Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R.J., Rahmsdorf, H.J., Jonat, C., Herrlich, P., and Karin, M. (1987) Cell 49, 729-739.
22. Imagawa, M., Chiu, R., and Karin, M. (1987) Cell 51, 251-260.
23. Montminy, M.R., Sevarino, K., Wagner, J., Mandel, G., and Goodman, R.H. (1986) Proc. Natl. Acad. Sci. USA 83, 6682-6686.
24. Fink, J.S., Verhave, M., Kasper, S., Tsukada, T., Mandel, G., and Goodman, R.H. (1988) Proc. Natl. Acad. Sci. USA 85, 6662-6666.
25. Bell, G.I. (1986) Peptides 7, 27-36.
26. Weiher, H., Konig, M., and Gruss, P. (1983) Science 219, 626-631.
27. Lin, Y.S. and Green, M.R. (1988) Proc. Natl. Acad. Sci. USA 85, 3396-3400.
28. Klein, R.D. and Wells, R.D. (1982) J. Biol. Chem. 257, 12954-12961.
29. McLauchlan, J., Gaffney, D., Whitton, J.L., and Clements, J.B. (1985) Nucleic Acids Res. 13, 1347-1368.
30. Mayo, K.E., Cerelli, G.M., Lebo, R.V., Bruce, B.O., Rosenfeld, M.G., and Evans, R.M. (1985) Proc. Natl. Acad. Sci. USA 82, 63-67.
31. Mayo, K.E., Cerelli, G.M., Rosenfeld, M.G., and Evans, R.M. (1985) Nature 314, 464-467.