
Satoshi MIKAWA, Gen-ichi YOSHIMOKA, Hideyuki AOKI, Yoshiaki YAMANO,* Hiroshi SAKAI, and Tohru KOMANO**

Laboratory of Biochemistry, Department of Agricultural Chemistry, Kyoto University, Kyoto 606-01, Japan
Received August 10, 1994

Goat IGF-I (gIGF-I) cDNA was cloned using the reverse transcriptase-polymerase chain reaction (RT-PCR). The cDNA, which was homologous to rat class 1 IGF-I cDNA, was 969 bp long. From the open reading frame found in it, we predicted a 154 amino acid protein consisting of a 49 amino acid signal peptide, a 70 amino acid mature IGF-I peptide, and a 35 amino acid E domain (the COOH-terminal peptide). From the gIGF-I gene, we isolated and sequenced some segments containing four exons that encompassed the entire gIGF-I cDNA sequence. Goat liver RNA was analyzed by RT-PCR, and the nucleotides of the RT-PCR products were sequenced and checked with the nucleotide sequences of the segments from the gIGF-I gene. The gene had three leader exons (1W, 1, and 2, from upstream to downstream) and was transcribed into three kinds of mRNAs (classes 1W, 1, and 2). Another RNA species was detected by RT-PCR analysis of exon 1W. We sequenced it and found that in this transcript, the 3′-portion of exon 1 was inserted between exons 1W and 3, resulting in class 1W-1 del. mRNA. That is to say, the gIGF-I gene had three leader exons and four kinds of mature mRNA.

Insulin-like growth factors IGF-I and II are polypeptide hormones structurally homologous to proinsulin. Their precursors consist of a signal peptide, a mature protein, and an E domain, which is removed post-translationally. The two growth factors share the functions of regulation of development and somatic growth, but IGF-I acts postnatally and IGF-II acts prenatally. IGF-I is synthesized in a variety of tissues and the transcription of its gene is regulated by growth hormone. The biological actions of IGF-I begin by interaction with its cell-surface receptor, which is a ligand-activated tyrosine-specific protein kinase with structural and functional similarities to the insulin receptor. IGF-I can bind the insulin receptor also, but the affinity between them is about one-hundredth than between IGF-I and the IGF-I receptor.

cDNA and genes coding for IGF-I of some animals have been isolated (Fig. 1). The rat IGF-I gene consists of six exons and five introns, and is more than 80 kb long. Many kinds of mature mRNA are synthesized from the IGF-I gene through multiple polyadenylation sites and alternative splicing. In rats and humans because of two leader exons (1 and 2), there are two kinds of IGF-I mRNA: classes 1 and 2. There is a third kind of mRNA in the rat: class 1 del., which lacks the central region of exon 1 from class 1 mRNA. There are two kinds of E domains. In rat, the Ea is encoded by exons 4 and 6, and the Eb is encoded by exons 4, 5, and 6. In human, the Ea is encoded by exons 3 and 4, and the Eb is encoded by exons 3 and 5. In sheep, another leader exon (exon 1W) was found, resulting in class 1W mRNA, but exon 5 was not found. The IGF-I precursors encoded by the multiple mRNAs differ in their signal peptide and E domain but are processed to become the same mature IGF-I. The biological significance of the diversity in the signal peptides and E domains is unknown, but perhaps the secretion of IGF-I is regulated in different ways specified by the different signal peptides or by differences in the processing of the IGF-I precursor.

We report here the structure and mode of expression of the goat IGF-I gene and the structure of its cDNA. The gIGF-I gene contained at least six exons including three leader exons (1W, 1, and 2). By the differential use of the three leader exons, three kinds of gIGF-I mRNA are produced. We found a fourth kind of mRNA that contained both exon 1W and the 3′-portion of exon 1.

Materials and Methods

Animals. We isolated the liver and spleen from a female Shiba goat (Capra hircus), 3 years old.

Isolation of the goat IGF-I cDNA by RT-PCR. For isolation of gIGF-I

† The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL, and GenBank Nucleotide Sequence Databases with the accession numbers S12578, D26116, D26117, D26118, and D26119.
* Present address: Faculty of Agriculture, Tottori University, Tottori 680, Japan.
** Corresponding author.

Fig. 1. Structure of the Human, Rat, Ovine IGF-I Genes.

Introns are represented by lines and exons by boxes. Solid boxes correspond to the open reading frames encoding the IGF-I precursors. The hatched area of the rat IGF-I gene indicates the sequence that is spliced out, giving the class 1 del. mRNA.
cDNA, we prepared primer 5' (ATGGCAGATTTGCTCTACA), which
was in the 5'-flanking region conserved among human22 and sheep 23
IGF-I cDNAs, and primer 3' (GGCTAGAAGAAGATCGCAAAAT),
which was complementary to the 3'-terminal part of gIGF-I cDNA cloned
from a goat liver cDNA library (Fig. 2). Total RNA was prepared from
goose liver by acid-guanidino-thiocyanate-phenol-chloroform (AGPC)
methods. 24 One microgram of the total RNA was transcribed into cDNA
with 200 units of reverse transcriptase from murine leukemia virus (Gibco
BRL) in 20 µl of a reaction mixture containing 50 mM KCl, 20 mM Tris-Cl
(pH 8.4), 2.5 mM MgCl2, 0.1 mg/ml BSA, 1 mM each dNTP, 5 mM random
hexamer, and 20 units of ribonuclease inhibitor RNasin (Promega). The
cDNA synthesized during 1 h of incubation at 24°C was used as a template
for PCR in a reaction mixture containing 5 units of Taq DNA polymerase,
50 mM KCl, 20 mM Tris-Cl (pH 8.4), 2.5 mM MgCl2, 0.1 mg/ml BSA, 0.2 mM
each dNTP, and 0.2 mM each of the two primers. After 25 cycles (1 min
at 95°C, 1 min at 55°C, and 2 min at 72°C) of the PCR, the mixture was
electrophoresed through a 1% agarose gel, and the DNA fragments were
isolated and subcloned into pBluescriptII. Sequence analysis was done for
twelve clones from three independent RT-PCR.
Cloning of the goat IGF-I gene. A genome DNA library was constructed
from goat spleen DNA that had been partially digested by Sau3AI, and
the phage vector λEMB3L. Clones having segments of the gIGF-I gene
were isolated by plaque hybridization with gIGF-I cDNA as a probe. The
dNA of the positive clones were digested with restriction endonucleases,
and Southern hybridization analysis was done using the gIGF-I cDNA as
the probe.

Detection of additional leader exons. For detection of exons, we prepared
primer 1W (TTTACCCCAAGTGGTTTGGAG), primer 1 (AGACCTTG-
CACCCAGAAGCA), and primer 2 (TGCACGCTGCTAAGAGTTCT)
with sequences found in the regions of the IGF-I gene corresponding to
exons 1W, 1, and 2 of the sheep IGF-I gene, respectively, and we prepared
primer 6 (GTGACAGGGAGCGAAGGACTT), which was complementary
to the 20-nt sequence in exon 6 of the gIGF-I gene. RT-PCR
was done using goat liver RNA as described above.

Primer extension. We prepared primer pW (AAGAGCGGCTGCG-
GATGACGACGTCGCCTGGTTTGGAG) complementary to a 36-nt se-
quence in exon 1W, primer p1 (GCAAGCTAGCTATCCTGCTGGT)
complementary to a 20-nt sequence in exon 1, and primer p2 (TCACCTGTG-
TAGGGTAAAC) complementary to a 20-nt sequence in exon 2. Ten
picoatoms of each synthetic primer were 5'-end labeled with [γ-32P]ATP.
One picomole of a labeled primer and 6 µg of polyadenylated RNA from
goose liver were hybridized to each other for 12 h at 30°C in 40 mM PIPES
(pH 6.4), 1 mM EDTA, 0.2 mM NaCl, and 40% formamide. After pre-
cipitation of the nucleic acids with ethanol, chain extension from the 3'-
ends of the primers was done with 200 units of reverse transcriptase of
murine leukemia virus (Gibco BRL) for 2 h at 37°C in a mixture of 50 mM
Tris-Cl (pH 7.6), 60 mM KCl, 10 mM MgCl2, 1 mM each dNTP, 1 mM DTT,
and 50 mg/ml actinomycin D. After treatment with ribonuclease A, then
with phenol-chloroform (1:1, v/v), the samples were analyzed on a
denaturing polyacrylamide gel.

Results
Cloning of gIGF-I cDNA
We isolated goat IGF-I cDNA from goat liver RNA by RT-PCR. Sequence analysis showed that this cDNA was
homologous to rat class 1 IGF-I cDNA. This cDNA was

![Fig. 2. Nucleotide Sequence of Goat IGF-I cDNA and the Deduced Amino Acid Sequence.](image_url)

Goat IGF-I cDNA was isolated by RT-PCR as described in Materials and Methods, and its nucleotides were sequenced with an ABI 373A DNA sequencer (Applied Biosystems). The underlined amino acid sequence is the mature IGF-I protein, which consists of the B, C, A, and D domains. The E domain corresponds to a COOH-terminal peptide, which is removed post-translationally. The nucleotide sequences with double underlining are the locations of the primers for RT-PCR. The shorter goat IGF-I cDNA, 966 bp long, lacks the italic nucleotide (C<sub>440</sub>A<sub>143</sub>G<sub>143</sub>).

b. Diagram of the gIGF-I precursor. Boxes, from left to right, are the signals peptide, B, C, A, D, and E domains. Mature IGF-I consists of the B, C, A, and D domains.
969 nucleotides long, and from the open reading frame found in it, we predicted a 154-amino acid protein consisting of a 49-amino acid signal peptide, a 70-amino acid mature IGF-I peptide, and a 35 amino acid E domain (Fig. 2). Another kind of cDNA was isolated. It was 966 nucleotides long and lacked the trinucleotide C148A142G143, so that it encoded an IGF-I precursor that lacked one amino acid, Gln 22, in its signal peptide.

In the nucleotide sequence of the coding region, gIGF-I cDNA was highly homologous (93%) to human IGF-I cDNA.16-18 Only one amino acid (Thr 67) of the mature IGF-I protein was different from the residues in the same position in human and bovine23 IGFI (Ala 67).

**Cloning of the gIGF-I gene**

Upon screening of $2 \times 10^5$ clones by plaque hybridization, four positive clones (2gIGF I, 3, 5, and 6) were isolated. Southern hybridization and sequence analysis showed that 2gIGF 3 contained exons 1 and 3 and region corresponding to exons 1W and 2 of the sheep IGF-I gene, that 2gIGF 5 contained exon 4, and that 2gIGF 1 and 6 contained exon 6 (Fig. 3). Exons 1 and 3 encoded the signal peptide, and exons 3 and 4 encoded mature IGF-I (B, C, A, and D domains). The E domain was encoded by exons 4 and 6. The trinucleotide C148A142G143 missing in the shorter cDNA was found to be at the 5' splicing site of exon 3. Each exon was numbered as in the human IGF-I gene.

**Detection of exon 1W and 2**

It has been reported that some mammalian IGF-I genes have multiple leader exons. For example, the human IGF-I gene has two leader exons (1 and 2),22 and the sheep IGF-I gene has three leader exons (1W, 1, and 2).23 The nucleotide sequence of exon 1 was mapped in the gIGF-I gene, and was homologous to the sequences of the human and sheep IGF-I genes. We searched for additional leader exons in the gIGF-I gene by RT-PCR using goat liver RNA as described in Materials and Methods (Fig. 4). When we used primers 1W and 6 designed to detect exon 1W, two DNA fragments were amplified. One was 600 bp long and its nucleotide sequence showed that it contained exons 1W, 3, 4, and 6. The other was longer than expected, 680 bp. Sequence analysis showed that this DNA fragment contained the 3'-portion of exon inserted between exons 1W and 3. With primers 1 and 6 designed to detect exon 1, a 545-bp DNA fragment was amplified that contained exons 1, 3, 4, and 6. When primers 2 and 6 were used, a 510-bp

---

**Fig. 4. Detection of Leader Exons.**

(a) RT-PCR was done with total RNA from goat liver as described in Materials and Methods. When primers 1W and 6 were used, 680-bp and 600-bp DNA fragments were amplified (lane 1). When primers 1 and 6 were used, 545-bp DNA fragments were amplified (lane 2). When primers 2 and 6 were used, 510-bp DNA fragments were amplified (lane 3). b) Structure of the goat IGF-I gene. Sequence analysis showed that the goat IGF-I gene had three leader exons (1W, 1, and 2), which were transcribed to three kinds of mRNAs (classes 1W, 1, and 2). The 600-bp DNA fragment, a larger fragment than we expected, contained the 3'-portion of exon 1 inserted between exons 1W and 3, resulting in a fourth kind of goat IGF-I mRNA (class 1W-1del). Mature mRNAs are represented by bold lines connected by thin lines. Arrows above exons 1W, 1, 2, and 6 show the position of primers 1W, 1, 2, and 6, respectively.

**Fig. 5. Comparison of the Splice Recipient Sites in Exon 1 of the Rat22,23 and Goat IGF-I Genes.**

Exons are shown by boxes; solid areas are translational regions and hatched areas show the regions spliced out, giving the rat class 1 del. and the goat 1W-1 del. mRNAs. The nucleotides in the rat class 1 del. mRNA and the goat class 1W-1 del. mRNA are shown in upper-case letters. The nucleotides that are removed upon splicing are shown in lower-case letters. The nucleotide sequences of these two splice recipient sites indicated by bold letters are conserved.
DNA fragment was amplified that contained exons 2, 3, 4, and 6. The results showed that the IGF-I gene had three leader exons (1W, 1, and 2, from upstream), and that the gene was transcribed and processed into three kinds of mature mRNAs for the leader exon (class 1W, 1, and 2). The gIGF-I gene was transcribed into a fourth kind of mRNA, which had both exon 1W and the 3'-portion of exon 1 (class 1W-1 del.; Fig. 4).

Exon 1 of the rat IGF-I gene and exon 1 of the gIGF-I gene were homologous to each other. Exon 1 of the rat IGF-I gene is transcribed into two kinds of mature mRNAs, one of which contains intact exon 1 (class 1), and the other of which lacks the central region of exon 1 (class 1 del.). Sequence analysis around the splicing recipient sites in exon 1 of goat class 1W-1 del. cDNA and rat class 1 del. cDNA showed that these splicing recipient sites were in identical regions and were conserved between rat and goat (Fig. 5).

**Mapping of transcriptional initiation sites**

To map transcriptional initiation sites for each first exon, primer extension analysis was done (Fig. 6). With primer p1W, a transcriptional initiation site was mapped 274 bp upstream from the 3'-end of exon 1W (panel a; lane W). With primer p1, major initiation sites were mapped 247, 251, and 253 bp upstream from the 3'-end of exon 1 and there were some minor sites around there (panel b; lane 1). We could not exclude the possibility that there are other signals because of smearing. With primer p2, initiation sites were mapped 68 and 69 bp upstream from the 3'-end of exon 2 (panel c; lane 2). Upstream from exons 1W, 1, and 2 of the gIGF-I gene, CAAT and TATA boxes were not found. Upstream from exon 2, two cyclic AMP response element-like sequences were found. In exons 1W and 1, multiple translation initiation codons (ATG) were found, but were followed immediately by stop codons. We found that each of exons 1, 2, and 3 had an ATG codon organizing an open reading frame encoding the IGF-I precursor, but exon 1W did not have it (Fig. 7).

**Discussion**

We isolated two kinds of class 1 IGF-I cDNA from goat liver RNA, which differ from each other only in the trinucleotide C_{141}A_{142}G_{143}. This trinucleotide is not present in IGF-I cDNA from humans,\(^{16-18}\) rats,\(^{14,21}\) mice,\(^{26}\) and swine.\(^{27}\) Sequence analysis of the gIGF-I gene showed that the trinucleotide was at the 5'-splicing site of exon 3. The nucleotide sequence including the splicing site was CAGCAGGTGAAG, and the two 5'-splicing sites of exon 3 were cacCAGGTGAAG and cacagGTGAAG. In sheep, the nucleotide sequence of this region is conserved as reported by Wong et al.\(^ {28}\) and Ohlsen et al.\(^ {29}\) Sheep class 1 and 2 cDNAs have this trinucleotide, but sheep class 1W cDNA does not. We also detected two kinds of goat class 2 mRNAs, in which one had this trinucleotide, but the other did not, in goat liver RNA (data not shown). However it is not known whether splicing of exon 3 is regulated differently for tissues or for each class of cDNA.

Results of RT-PCR using goat liver RNA showed that the gIGF-I gene had three leader exons (1W, 1A, and 2), and that they resulted in four kinds of mature mRNAs (class 1W, 1W-1 del., 1, and 2) in liver. A preliminary study with RT-PCR showed that class 1 mRNA was expressed in various tissues, and that class 1W, 1W-1 del., and 2 mRNAs could be detected only in liver, uterus, and ovary (data not shown). In exons 1W and 1, we found multiple translation initiation codons (ATG) followed immediately by stop
codons. These regions may help to regulate translation. The translational initiation sites of the four kinds of goat matureIGF-I mRNAs were not located, but we found that each of exons 1, 2, and 3 had an ATG codon organizing an open reading frame encoding the IGF-I precursor (Fig. 7). A translational initiation codon was not found in exon 1W, so translation of the class 1W mRNA may start at an ATG in exon 3. In this case, the putative signal peptide consists of 25 amino acid residues. It is possible that the class 1W mRNA is not translated at all. In the class 1W.1 del. and 1 cDNAs, the putative translational initiation codons, found in exon 1, were identical. The signal peptides of the IGF-I precursors encoded by the class 1W.1 del. and 1 mRNAs were 49 amino acid residues long. In exon 2 of the class 2 cDNA, we found an ATG leading to a signal peptide 33 amino acid residues long. If all the four kinds of mRNAs are translated, there are three kinds of IGF-I precursors, which differ from one another only in their signal peptides. Activity of promoters upstream of each leader exon may determine the abundance of each type of the signal peptide in cell. Moreover, for transcripts containing exon 1W, post-transcriptional processing may participate in selection of types of the signal peptides.

The amino acid sequences were highly homologous among human, bovine, and goat mature IGF-I. Only the 67th amino acid in the mature IGF-I was different (threonine...
in goat IGF-I and alanine in human and cattle). The
conserved structure of the mature IGF-I may be functionally
advantageous for interaction with various factors (IGF
binding proteins and receptors). The signal peptide of each
class of IGF-I precursor was also conserved among human,
bovine, and goat. It is not known whether the difference in
signal peptides has any special biological significance, but
it is possible that the IGF-I is secreted in a different way
under the direction of the different signal peptides, resulting
in paracrine, autocrine, and endocrine actions.

Four kinds of gIGF-I mRNAs encoded the same mature
protein and the E domain. For rats, two kinds of E domains
have been identified, one of which is encoded by the exons
4 and 6 (Ea), and the other by exons 4, 5, and 6 (Eb). None
of the DNA products of the RT-PCR contained a sequence
inserted between exons 4 and 6, so the goat IGF-I gene
may have no exon corresponding to exon 5 of the rat IGF-I
gene. In the human IGF-I precursor, two kinds of E domains
have been found, encoded by exons 4 and 6 (Ea) or by
exons 4 and 5 (Eb). We prepared mix primers with sequences
encoding the part of human Eb domain, and RT-PCR was
done using goat liver RNA. But no DNA fragments was
amplified (data not shown). To find whether the gIGF-I
gene has exon 5, we have to isolate the region of the goat
IGF-I gene corresponding to exon 5 of the human IGF-I
gene and analyzed the transcription of this region.

Transcriptional initiation sites of exon 1 reported in rat
and human were mapped to two conserved regions, 30–32
which were about 250 and 350 bp upstream from the 3'end
of exon 1. In sheeps, initiation sites were mapped about 180
and 350 bp upstream from the 3'end of exon 1.29 In goats,
initiation sites were mapped at 247, 251, and 253 bp
upstream from the 3'end of exon 1 and conserved like that
of rat and human. Around 180 and 350 bp upstream of the
3'end of exon 1, no initiation site was detected. Transcriptional initiation sites of exon 2 were reported in
rat, human, and sheep. Major initiation sites were mapped
about 50 bp upstream from the 3'end of exon 2 in rat30
and sheep, 23 and about 90 bp upstream from 3'end of exon
2 in sheep.20 In goats, initiation sites were mapped at about
700 bp upstream from the 3'end of exon 2, and was
conformable to those of human. Additional initiation sites
which were mapped 700–850 bp further upstream in rat,
human, and sheep, were not detected in goat. Because the
transcripts from this region were small in rat, human, and
sheep, we cannot rule out the possibility that initiation
sites were not detected in this region because of the low
sensitivity. Upstream of exons 1W, 1, and 2, CAAT and
TATA boxes were not found. This may be a reason why
transcription initiates at multiple sites, and why initiation
sites were not defined unequivocally. Upstream of exon 2,
two cAMP response element-like sequences were also
found, suggesting that this element may be involved in
cAMP-dependent stimulation of the transcription of the
IGF-I gene as reported previously. 33,34

References
2) E. Schoene, J. Zapf, R. E. Humble, and E. R. Froesch, Nature,
4) A. Nilsson, J. Isgaard, A. Lindahl, A. Dahlstrom, A. Skottner, and
6) A. Kurtz, J. Zapf, K.-U. Eckardt, G. Clemens, and E. R. Froesch,
7) Y. Nakaniishi, M. L. Mulshine, P. G. Kasprzyk, R. B. Natale, R.
Maneckeck, I. Avish, A. M. Treston, A. F. Gazdar, J. D. Minna, and
8) R. Lajara, P. Rotwein, J. D. Bortz, V. A. Hansen, J. L. Sadow, C.
R. Betts, S. A. Rogers, and M. R. Hammerman, Am. J. Physiol.,
9) P. K. Lund, B. M. Moats-Scotta, M. A. Hynes, J. G. Simmons, M.
Jansen, A. J. D’Ercole, and J. J. Van Wyk, J. Biol. Chem., 261,
14539–14544 (1986).
11) A. Ullrich, A. Gray, A. W. Tam, T. Yang-Feng, M. Tsubokawa, C.
Collins, W. Henzel, T. Le Bon, S. Kathuria, E. Chen, S. Jacobs, U.
Francke, J. Ramachandran, and Y. Fujita-Yamaguchi, EMBO J.,
13) C. T. Roberts, Jr., S. R. Lackey, W. L. Lowe, Jr., and D. LeRoith,
14) S. J. Casella, E. P. Smith, J. J. Van Wyk, D. R. Joseph, M. A. Hynes,
15) C. Bucci, P. Mallucci, C. T. Roberts, R. Frunzio, and C. B. Brunni,
16) M. Jansen, F. M. A. Van Schaik, A. T. Rickler, B. Bullock, D.
Wood, K. H. Gabby, A. L. Nussbaum, I. S. Sussenbach, and J. L.
17) Y. L. Bouc, D. Dreyer, F. Jaeger, M. Binoux, and P. Sondemeyer,
20) H. L. Foyt, D. LeRoith, and C. T. Roberts, Jr., J. Biol. Chem.,
21) C. T. Roberts, Jr., S. R. Lacky, W. L. Lowe, Jr., W. T. Seaman,
26) G. I. Bell, M. M. Stempien, N. M. Fong, and L. B. Rall, Nucl. Acids
27) A. Takavakkol, F. A. Simmen, and R.C.M. Simmen, Mol. Endocrinol.,
28) E. A. Song, S. M. Ohlsen, J. A. Goddardson, D. M. Dean, and J.
29) S. M. Ohlsen, D. M. Dean, and E. A. Song, DNA Cell Biol., 12,
30) M. L. Adamo, H. Ben-Hur, D. LeRoith, and C. T. Roberts, Jr.,
31) E. Janssen, P. H. Stenenbergh, D. LeRoith, C. T. Roberts, Jr., and J.
34) T. L. McCarthy, M. Centrella, and E. Canalis, J. Biol. Chem.,