MOLECULAR CLONING OF A NEW HUMAN INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN

Michael C. Kiefer 1 , Richard S. Ioh 1 , Diane M. Bauer 1 and Jürgen Zapf 2*

 Chiron Corporation, Emeryville, CA 94608
 Metabolic Unit, Department of Medicine, University Hospital, CH-8091 Zürich, Switzerland

Received February 14, 1991

We have cloned a new insulin-like growth factors binding protein (IGFBP) from a human osteosarcoma cDNA library. Two conserved regions in the COOH-terminal third of the five known human IGFBPs were used to design primers and to perform polymerase chain reaction (PCR) with osteosarcoma cDNA as a template. One of the eight PCR products encoded a unique IGFBP sequence. The DNA sequence was used to synthesize probes to screen an osteosarcoma cDNA library and isolate full length cDNA clones. The amino acid sequence was deduced from one of them. It contains two possible signal peptidase cleavage sites yielding a mature molecule of 257 or 252 amino acids, and 18 cysteines in identical positions to the other IGFBPs. The most pronounced homology exists with human IGFBP-3 (50 % in the NH2- and 45 % in the COOH-terminal region).

© 1991 Academic Press, Inc.

Insulin-like growth factors (IGF) occur in serum and other biological fluids complexed with specific high-affinity binding proteins (BPs) (1,2). The biological role of these proteins is still undetermined. They have been found to inhibit (3,4,5,6,7) as well as to stimulate (6,8,9) IGF actions. The most abundant IGFBP of human serum, termed IGFBP-3 (10), has been purified by Martin and Baxter (11) and its amino acid sequence has been derived from a cDNA cloned from human liver (12). In addition to BP-3 and a COOH-terminal truncated BP-3 we have recently purified four other BPs from human serum and obtained their NH2-terminal amino acid sequences (13,14). Two of these proteins have identical NH2-termini and differ only by the presence of an N-glycosyl residue. The complete amino acid sequences of these serum IGFBPs have been derived from their cDNAs which have been cloned from HepG2 hepatoma and from osteosarcoma cDNA libraries (13,14). The amino acid sequences of all known human IGFBPs show a high degree of homology within their NH2- and COOH-terminal thirds (14). Two conserved regions in the COOH-termini of the

Abbreviations: IGF, insulin-like growth factor; BP, binding protein; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate.

^{*} To whom correspondence should be sent.

known IGFBPs were used to construct primers, perform PCR with osteosarcoma cDNA as a template, and to screen the osteosarcoma cDNA library with a probe synthesized according to a unique sequence of one of the PCR products. Thereby, a new, sixth IGFBP was discovered.

MATERIALS AND METHODS

<u>Tissues</u> - Human osteosarcoma tissue was obtained from Dr. Marshall Urist (University of California, Los Angeles).

RNA isolation - RNA was isolated by the guanidinium thiocyanate method (15) with modifications (16). Poly(A)+RNA was purified by a single fractionation over oligo (dT)-cellulose (17).

Oligonucleotide synthesis - Oligonucleotide adaptors, probes, and primers were synthesized by the phosphoramidite method with an Applied Biosystems model 380A synthesizer, purified by polyacrylamide gel electrophoresis (PAGE) and desalted on Sep-Pak C18 cartridges (Waters, Milford, MA).

A 14-mer oligonucleotide (5' CCTGTAGATCTCCG 3') and a 18-mer oligonucleotide (5' AATTCGGAGATCTACAGG 3') were synthesized and used as the EcoRI adaptors for the cDNA library constructed in λ ZAP II. The 14-mer was phosphorylated (18), then immediately heated to 95° C for 15 min to inactivate the polynucleotide kinase. The adaptors also contain an internal Bql II site.

The two consensus PCR primers used to identify BP-6 were a sense primer consisting of a mixture of 32 24-mers (5' AGATCTGAATTCGCCXAA $^{C}/_{T}$ TG $^{C}/_{T}$ A $_{G}$ A $_{S}$ 3') and an antisense primer consisting of a mixture of 16 25-mers (5' AGATCTAAGCTTCXACA/GCACCAA/GCA 3') where X denotes all four deoxynucleotides. EcoRI and Hind III sites were included in the sense and antisense primers, respectively, to allow for subcloning of the PCR products into M13 sequencing vectors. The BP-6 probes used to screen the cDNA library were two 19-mers, (5' GCAAAGGATTCTACAAGAG 3') and (5' CAAACCTTCCCGTGGCCGC 3'). PCR Amplification - The PCR reactions were performed with the PCR kit (Perkin Elmer Cetus) according to the instructions of the supplier using the PCR primers described above at a final concentration of 8 µM. The template cDNA was synthesized from 2.5 μg of human osteosarcoma (Ost2) poly(A) † RNA. The conditions of cDNA synthesis were identical to those for first strand cDNA synthesis (see construction of cDNA library). The cDNA was fractionated on Biogel A-15m, recovered by ethanol precipitation and resuspended in 100 µl of sterile water. 1 µl of cDNA template was used for the PCR reaction. 35 cycles of PCR were performed in a Perkin Elmer Cetus DNA thermal cycler. The first 10 cycles consisted of a 94° C, 1 min denaturation step, a 33° C, 1 min annealing step and a 33° C, 1 min extension step. The next 25 cycles consisted of a 94° C, 1 min denaturation step, a 55° C, 1 min annealing step and a 72° C, 1 min extension step. The final extension step of the last cycle was 7 min. The sample was extracted once with phenol/chloroform/isoamylalcohol (1:1:0.04), once with chloroform/isoamylalcohol (24:1), and recovered by ethanol precipitation. The PCR DNA product was them incubated for 20 min at 37° C with 10 units of DNA polymerase I, Klenow fragment in 10 mM Tris-HCl ph 7.5, 10 mM MgCl2, 50 mM NaCl, 1 mM dethiothreitol and 40 μ M each of dATP, dGTP, dTTP and dCTP. The sample was extracted as above, recovered by ethanol precipitation, digested with EcoRI and Hind III, and fractionated by electrophoresis on a 7 % acrylamide, 1xTBE (Tris/borate/EDTA) gel. DNA migrating between 80-100 base pairs was excised from the gel, passively eluted for 16 h with gentle shaking in 10 mM Tris-HCl ph 7.5, 1 mM EDTA, purified by passage over an elutip-D column as described by the supplier (Schleicher and Schuell), ligated to an EcoRI- and Hind III-cut M13 sequencing vector (mp18) and introduced into E. coli strain DH5 α F'

<u>Construction of the cDNA library</u> - A λ ZAP II human osteosarcoma cDNA library was constructed from human osteosarcoma poly(A)[†]RNA as described (13). A library of 1.75x10⁷ independent recombinant clones was obtained.

Screening of the cDNA library - Approximately 300'000 recombinant phages from the Ost4 cDNA library were plated (50'000 phages/137 mm diameter plate) in E. coli BB4 and grown for 5-6 hours at 37° C. The phages were transferred onto nitrocellulose filters (Millipore, HATF 137), processed (19) and screened with the two BP-6 probes. The probes were labeled with T₄ polynucleotide kinase and $[\delta^{32}P]$ ATP (18) to a specific activity of 1-2x10⁸ cpm/µg. The filters were prehybridized for 1-2 h at 37° C as described (13). Labeled probe was added to a concentration of 10^6 cpm/ml and hybridization was continued overnight at 37° C with gentle shaking. The filters were washed in 2 x SCC (1 x SSC = 0.15 M sodium chloride/0.015 M sodium citrate, pH 7), 0.1 % SDS at 50° C and exposed overnight at -80° C to Kodak XAR-2 films with a DuPont Lightning Plus intensifying screen. Areas of plaques giving duplicate signals were picked, replicated and rescreened until pure plagues were obtained. Plasmid isolation, subcloning and sequencing - Bluescript SK(-) plasmids containing BP-6 cDNA inserts were released from λ ZAP by the M13 rescue/excision protocol described by the supplier (Stratagene). Plasmid DNA was isolated by the alkaline lysis method (18). The inserts were excised from the Bluescript SK(-) vector by Bgl II digestion and fractionated by agarose gel electrophoresis. Inserts were cut out from the gel and eluted for 12 h with gentle shaking in 10 mM Tris-HCl pH 7.5, 1 mM EDTA (TE), purified over an elutip-D column (see above) and subcloned into a M13 sequencing vector (20). DNA sequencing was performed by the dideoxy chain termination method (21) using M13 primers as well as specific internal primers. Ambiguous regions were resolved using 7-deaza-2-deoxyquanosine- triphosphate (22) and Sequenase (US Biochemicals). Northern blot analysis - Poly(A) +RNA was fractionated on 1.4 % agarose gel in the presence of formaldehyde (23), transferred directly to nitrocellulose, and processed as described (24). Filters were hybridized with the purified cDNA insert of BP-6.1 as described above (screening of cDNA library). The filters were washed twice for 20 min in 0.1 x SSC, 0.1 % SDS at 65° C. The cDNA probes were labeled as described (24) to a specific activity of 2×10^9 cpm/mg.

RESULTS AND DISCUSSION

Identification and Cloning of BP-6 - An amino acid sequence comparison of the five known human IGFBPs (14) reveals a high degree of homology in the NH2- and the COOH-terminal regions. The longest stretch of identical amino acids in all five BPs resides in two areas of the COOH-terminal region and consists of three amino acids Pro-Asn-Cys and four amino acids Cys-Trp-Cys-Val. These conserved amino acids fall within a region of the BPs that has been shown to be homologous to 10 repeats within the NH2-terminal two thirds of the human thyroglobulin molecule (14).

In an attempt to identify new BPs, degenerate primers were designed based on these sequences and PCR was performed using human osteosarcoma cDNA as a template. DNA sequence analysis of the eight PCR products yielded one sequence identical to BP-2, three sequences identical to BP-4, three sequences identical to BP-5 and one unique sequence, tentatively termed BP-6. The BP-6 PCR product shows a 60 % DNA sequence identity and a 76 % amino acid sequence identity to BP-3. Based on the PCR DNA sequence of BP-6, two unique BP-6 DNA probes were synthesized and used to screen a λ ZAP II human osteosarcoma cDNA library. From the 300'000 recombinant clones screened, twelve clones were identified which hybridized to both probes. Five clones were further purified

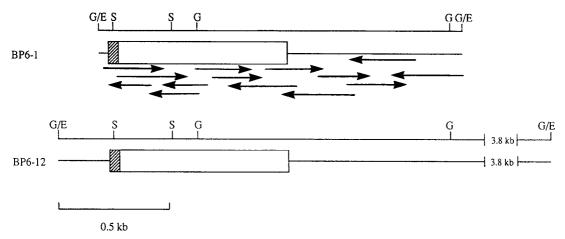


Fig. 1. Restriction endocuclease map of clones BP-6.1 and BP-6.12 Sequences were obtained by using M13 primers and specific internal primers. Arrows indicate the direction and extent of individual sequencing runs. The NH2-terminal signal sequence is represented by a hatched box and the mature protein by an open rectangular box. Restriction endonucleases are abbreviated as follows: S, Sac II; G, Bg1 II; E, EcoRI. The 3'end of BP-6.12 between G/E and G has not been sequenced.

and the cDNA inserts were analyzed by Bg1 II and EcoRI restriction enzyme digestion and agarose gel electrophoresis. The cDNAs fell into two size classes of ~1.7 kb and ~6 kb, which are exemplified by clones 1 and 12, respectively (fig. 1).

Expression of BP-6 mRNA - Northern blot analysis of several different tissues using 32 P-labeled clone 1 cDNA, confirmed that these two size classes of BP-6 mRNA exist (results not shown) and that osteosarcoma tissue synthesizes BP-6 mRNA. Other tissues analyzed (liver, kidney, brain) produced BP-6 mRNA transcripts, but at lower levels.

Sequence analysis of BP-6 - BP-6 clone 1 (BP-6.1) cDNA has been sequenced and is shown in fig. 2 with the deduced amino acid sequence. The NH₂-terminal hydrophobic region of the encoded BP-6 molecule has two acceptable signal peptidase cleavage sites (ψ) (25) following amino acids 15 and 20, respectively, and yielding mature molecules of 257 or 252 amino acids with a Mr of 29'017 or 28'577. There are no N-glycosylation sites on the encoded BP-6 protein. BP-6 contains 18 cysteine residues all of which coincide with cysteine residues in BP-1-5. There is a high degree of amino acid homology between BP-6 and the other five BPs, which is most pronounced in the NH₂- and the COOH-terminal regions of the molecules.

For amino acid sequence comparison, the BPs can be divided into three domains (fig. 3). Domain I contains the NH2-terminal region of the BP molecules (the first 75-95 amino acids) and is defined by 12 conserved cysteine residues (10 in the case of BP-4). In this region BP-6 is most similar to BP-3 (50 % identity). Domain II spans the middle region of the BP

CTCTCCTGCCCCACCCCGAGGTAAAGGGGGCGACTAAGAGAAG

1

1 Met Val Leu Leu Thr Ala Val Leu Leu Leu Leu Ala Ala Tyr Ala $^{\Psi}$ Gly Pro Ala 44 ATG GTG TTG CTC ACC GCG GTC CTC CTG CTG GCC GCC TAT GCG GGG CCG GCC 19 Gln Ser Leu Gly Ser Phe Val His Cys Glu Pro Cys Asp Glu Lys Ala Leu Ser 98 CAG AGC CTG GGC TCC TTC GTG CAC TGC GAG CCC TGC GAG AAA GCC CTC TCC 37 Met Cys Pro Pro Ser Pro Leu Gly Cys Glu Leu Val Lys Glu Pro Gly Cys Gly 152 ATG TGC CCC CCC AGC CCC CTG GGC TGC GAG CTG GTC AAG GAG CCG GGC TGC GGC 55 Cys Cys Met Thr Cys Ala Leu Ala Glu Gly Gln Ser Cys Gly Val Tyr Thr Glu 206 TGC TGC ATG ACC TGC GCC CTG GCC GAG GGG CAG TCG TGC GGC GTC TAC ACC GAG 73 Arg Cys Ala Gln Gly Leu Arg Cys Leu Pro Arg Gln Asp Glu Glu Lys Pro Leu 260 CGC TGC CCC CAG GGG CTG CGC TGC CCC CGG CAG GAC GAG GAG AAG CCG CTG 91 His Ala Leu Leu His Gly Arg Gly Val Cys Leu Asn Glu Lys Ser Tyr Arg Glu 314 CAC GCC CTG CTG CAC GGC CGC GGG GTT TGC CTC AAC GAA AAG AGC TAC CGC GAG 109 Gln Val Lys Ile Glu Arg Asp Ser Arg Glu His Glu Glu Pro Thr Thr Ser Glu 368 CAA GTC AAG ATC GAG AGA GAC TCC CGT GAG CAC GAG GAG CCC ACC ACC TCT GAG 127 Met Ala Glu Glu Thr Tyr Ser Pro Lys Ile Phe Arg Pro Lys His Thr Arg Ile 422 ATG GCC GAG GAG ACC TAC TCC CCC AAG ATC TTC CGG CCC AAA CAC ACC CGC ATC 145 Ser Glu Leu Lys Ala Glu Ala Val Lys Lys Asp Arg Arg Lys Lys Leu Thr Gln 476 TCC GAG CTG AAG GCT GAA GCA GTG AAG AAG GAC CGC AGA AAG AAG CTG ACC CAG 163 Ser Lys Phe Val Gly Gly Ala Glu Asn Thr Ala His Pro Arg Ile Ile Ser Ala 530 TCC AAG TTT GTC GGG GGA GCC GAG AAC ACT GCC CAC CCC CGG ATC ATC TCT GCA 181 Pro Glu Met Arg Gln Glu Ser Glu Gln Gly Pro Cys Arg Arg His Met Glu Ala 584 CCT GAG ATG AGA CAG GAG TCT GAG CAG GGC CCC TGC CGC AGA CAC ATG GAG GCT 199 Ser Leu Gln Glu Leu Lys Ala Ser Pro Arg Met Val Pro Arg Ala Val Tyr Leu 638 TCC CTG CAG GAG CTC AAA GCC AGC CCA CGC ATG GTG CCC CGT GCT GTG TAC CTG 217 Pro Asn Cys Asp Arg Lys Gly Phe Tyr Lys Arg Lys Gln Cys Lys Pro Ser Arg 692 CCC AAT TGT GAC CGC AAA GGA TTC TAC AAG AGA AAG CAG TGC AAA CCT TCC CGT 235 Gly Arg Lys Arg Gly Ile Cys Trp Cys Val Asp Lys Tyr Gly Met Lys Leu Pro 746 GGC CGC AAG CGT GGC ATC TGC TGC TGC GTG GAC AAG TAC GGG ATG AAG CTG CCA 253 Gly Met Glu Tyr Val Asp Gly Asp Phe Gln Cys His Thr Phe Asp Ser Ser Asn 800 GGC ATG GAG TAC GTT GAC GGG GAC TTT CAG TGC CAC ACC TTC GAC AGC AGC AAC 271 Val Glu OP 993 GGAAACTGAGGACCTCGGAATCTCTAGCAAGGGCTCAACTTCGAAAATGGCAACAACAGAGATGCAAAAAG 1064 CTAAAAAGACACCCCCCCCTTTAAATGGTTTTCTTTTTGAGGCAAGTTGGATGAACAGAGAAGGGAAGAG 1135 AGGAAGAACGAGGGAAGAGGGAAGGGAAGTGTTTGTGTAGAAGAGAGAGAAAGACGAATAGAGTTAGG 1348 TGGGGTTTCCCAGGATGGATTGGTCAGGGGGAGAAAGGAAAAGGCAAAACACTCCAGGACCTCTCCCGGAT 1490 TGTTGTCTCTCCCCAAATTTATAAAAACTAAAATGCATTCCATTCCTCTGAAAGCAAAACAAATTCATAAT 1561 TGAGTGATATTAAATAGAGAGGTTTTCGGAAGCAGATCTGTGAATATGAAAT

Fig. 2. cDNA and derived amino acid sequence of human IGFBP-6. The amino acid sequence is given in the three letter code above the nucleotide sequence. The possible NH2-terminal signal peptidase cleavage sites are shown by arrows.

molecules (the following 56-93 amino acids) and is characterized by its lack of cysteine residues (except BP-5, which contains 2). This region displays little homology between all of the BPs including BP-6. Domain III consists of the COOH-terminal region of the BP molecules (80-101 amino acids) and is also homologous to 10 repeats found in the NH2-terminus of the thyroglobulin molecule, to a gastrointestinal tumor-associated antigen and to the invariant chain of the class II histocompatibility (MHC) antigen (14). This region is characterized by its 6 invariant cysteine residues and displays considerable

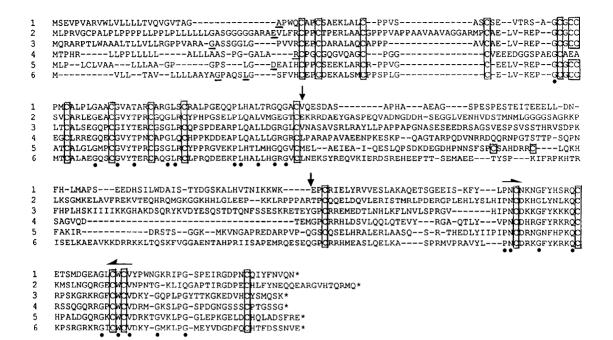


Fig. 3. Comparison of the amino acid sequences of the 6 human IGFBPs.
Sequences have been aligned such as to give maximal homology. Cysteine residues are boxed. The first amino acid of the mature proteins is underlined. For comparison of homologies the sequences are divided into three domains (see text) as indicated by vertical arrows. Dots indicate identical amino acid positions shared by all six IGFBPs. Horizontal arrows indicate two consensus regions that were used to construct PCR primers. The amino acid sequences for BP-1 to BP-5 are taken from the following references: BP-1 from ref. 28, BP-2 from ref. 13, BP-3 from ref. 12, BP-4 and -5 from ref. 14. A sequence identical to that of our BP-5 has recently been published by Shimasaki et al (29).

homology between the six BPs. BP-6 is most similar again to BP-3 in this domain (45 % sequence identity).

So far, all of the IGFBP bands identified by Western ligand blotting of human sera (26,27) have been assigned to one of the known IGFBPs (13,14). During HPLC purification of affinity-purified adult human serum we have not found IGFBP-1 which is a typical constituent of fetal human serum, nor have we detected a band that might correspond to BP-6 (13,14). We are presently trying to express BP-6 in yeast and to produce polyclonal antibodies. These will finally allow to screen normal and pathological human sera as well as other biological fluids and culture media from human cells for BP-6 and to see whether it is a normal member of the IGFBP family or a BP typical of tumor cells.

<u>ACKNOWLEDGMENTS</u>: We thank Ms. I. Buclin for devoted secretarial work. This study was supported by grant no. 3.046-0.87 of the Swiss National Science Foundation.

REFERENCES

1. Nissley, S.P., and Rechler, M.M. (1984) In: Hormonal Proteins and Peptides (C.H. Li), ed. p. 127-203. Academic Press, New York and London

- 2. Baxter, R.C., and Martin, J.L. (1989) Progress in Growth Factor Research 1.49-68
- 3. Zapf, J., Schoenle, E., Jagars, G., Sand, I., Grunwald, J., and Froesch, E.R. (1979) J. clin. Invest. 63,1077-1084
- 4. Meuli, C., Zapf, J., and Froesch, E.R. (1978) Diabetologia 14,255-259
- Knauer, D.J., and Smith G.L. (1980) Proc. Natl. Acad. Sci. USA 77,7252-7256
- De Mellow, J.S., and Baxter, R.C. (1988) Biochem. Biophys. Res. Commun. 156,199-204
- Mohan, S., Bautista, C.M., Wergedal, J., and Baylink, D.J. (1986) Proc. Natl. Acad. Sci. USA. 86,8338-8342
- Blum, W.F., Jenne, E.W., Reppin, F., Kietzmann, K., Ranke, M.B., and Bierich, J.R. (1989) Endocrinology 125,766-772
- 9. Elgin, R.G., Busby, W.H., and Clemmons, D.R. (1987) Proc. Natl. Acad. Sci. USA. 84,3254-3258
- 10. Ballard, J., Baxter, R.C., Binoux, M., Clemmons, D., Drop, S., Hall, K., Hintz, R.L., Rechler, M.M., Rutanen, E., and Schwander, J.J. (1989) Acta Endocrinol. (Kbh) 121,751-752
- 11. Martin, J.L., and Baxter, R.C. (1986) J. Biol. Chem. 261,8754-8760
- 12. Wood, W.I., Cachianes, G., Henzel, W.J., Winslow, G.A., Spencer, S.A., Hellmiss, R., Martin, J.L., and Baxter, R.C. (1988) Mol. Endocrinology 2,1176-1185
- 13. Zapf, J., Kiefer, M., Merryweather, J., Masiarz, F., Bauer, D., Born, W., Fischer, J.A., and Froesch, E.R. (1990) J. Biol. Chem. 265,14892-14898
- 14. Kiefer, M., Masiarz, F., Bauer, D., and Zapf, J. (1991) J. Biol. Chem., in press
- Chirgwin, J.M., Przbyla, A.E., MacDonald, R.J., and Rutter, W.J. (1979) Biochemistry 18,5294-5299
- Freeman, G.J., Clayberger, C., De Kruyff, R., Rosenblum, D.S., and Cantor, H. (1983) Proc. Natl. Acad. Sci. USA 80,4094-4098
- 17. Aviv, H., and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69,1408-1412
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Mol. Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 19. Benton, W.D., and Davis, R.W. (1977) Science 196,180-182
- 20. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene 33,103-119
- 21. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74,5463-5467
- 22. Barr, P.J., Thayer, R.M., Laybourn, P., Najarian, R.C., Seela, F., and Tolan, D. (1986) Biotechniques 4,428-432
- 23. Lehrach, H., Diamond, D., Wozney, J.M., and Boedtker, H. (1977) Biochemistry 16,4743-4751
- 24. Thomas, P. (1980) Proc. Natl. Acad. Sci. USA 77,5201-5205
- 25. Von Heyne (1986) Nucl. Acids Res. 11,4683-4690
- 26. Hardouin, S., Hossenlopp, P., Segovia, B., Seurin, D., Portolan, G., Lassarre, C., and Binoux, M. (1987) Eur. J. Biochem. 170,121-132
- 27. Zapf, J., Schmid, Ch., Guler, H.P., Waldvogel, M., Hauri, Ch., Futo, E., Hossenlopp, P., Binoux, M., and Froesch, E.R. (1990) J. Clin. Invest. 86,952-961
- 28. Lee, Y.L., Hintz, R.L., James, P.M., Lee, P.D.K., Shivley, J.E., and Powell, D.R. (1988) Mol. Endocrinol. 2,404-411
- 29. Shimasaki, S., Uchiyama. F., Shimonaka, M., and Ling N. (1990) Mol. Endocrinol. 4,1451-1458