

**Biochemistry.** In the article “Characterization of residual structure in the thermally denatured state of barnase by simulation and experiment: Description of the folding pathway” by Chris J. Bond, Kam-Bo Wong, Jane Clarke, Alan R. Fersht, and Valerie Daggett, which appeared in number 25, December 9, 1997, of *Proc. Natl. Acad. Sci. USA* (**94**, 13409–13413), one of the authors regrets that she inadvertently omitted references to the computer program and protein potential function that the authors used for their simulations of barnase cited above. The following sentence should have been the first sentence of the *Methods* section: Molecular dynamics simulations were performed with the program ENCAD (44) and the potential energy function of Levitt *et al.* (45).

- 44. Levitt, M. (1990) ENCAD, Energy Calculations and Dynamics. (Molecular Applications Group, Palo Alto, CA and Yeda, Rehovot, Israel).
- 45. Levitt, M., Hirshberg, M., Sharon, R. & Daggett, V. (1995) *Comp. Phys. Commun.* **91**, 215–231.

**Biochemistry.** In the article “Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin” by Kim C. Quon, Bing Yang, Ibrahim J. Domian, Lucy Shapiro, and Gregory T. Marczyński, which appeared in number 1, January 6, 1998, of *Proc. Natl. Acad. Sci. USA* (**95**, 120–125), the authors wish to note that the institutional affiliations in the author line were incorrectly attributed. The correct affiliations are as follows. Bing Yang and Gregory T. Marczyński are at McGill University in Montreal, and Kim C. Quon is now at the Netherlands Cancer Institute.

**Biochemistry.** In the article “The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecBCD enzyme from *Escherichia coli*” by Misook Yu, Jehanne Souaya, and Douglas A. Julin, which appeared in number 3, February 3, 1998, of *Proc. Natl. Acad. Sci. USA* (**95**, 981–986), the following correction should be noted. The symbols in the graph (Fig. 3C) were identified incorrectly in the manuscript. The corrected legend and graph with accompanying symbols are printed below.

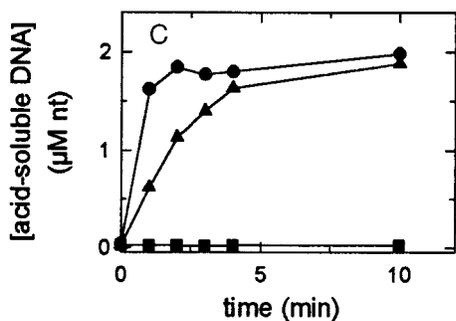


FIG. 3. Nuclease assays. All reaction mixtures contained buffer A and 10 mM MgCl<sub>2</sub>. (A) ssDNA exonuclease assay with 100 nM single-stranded linear 5'-<sup>32</sup>P-labeled 25-mer as the substrate. (B) ssDNA endonuclease assay using 6 nM single-stranded circular M13 phage DNA as the substrate. (C) dsDNA exonuclease assay using *Ava*I-digested [<sup>3</sup>H]pTZ19R. Each reaction contained 2.3 µM (nucleotides) double-stranded [<sup>3</sup>H]pTZ19R, 250 µM ATP, and 9 nM RecB<sub>1-929</sub>CD (■), 0.2 nM RecBCD (▲), or 9 nM RecBCdD (●).

**Biochemistry.** In the article “*Escherichia coli* RNA polymerase terminates transcription efficiently at rho-independent terminators on single-stranded DNA templates” by Susan M. Uptain and Michael J. Chamberlin, which appeared in number 25, December 9, 1997, of *Proc. Natl. Acad. Sci. USA* (**94**, 13548–13553), the authors request that the following correction be noted. It is critical that the bands in lanes 6 and 8 of Fig. 3 indicated by the T7Te arrow be visible. The existence of these terminated bands is a major point on which the conclusions of the paper depend. Therefore, to enhance their visibility, Fig. 3 and its accompanying legend are reprinted below with greater contrast.

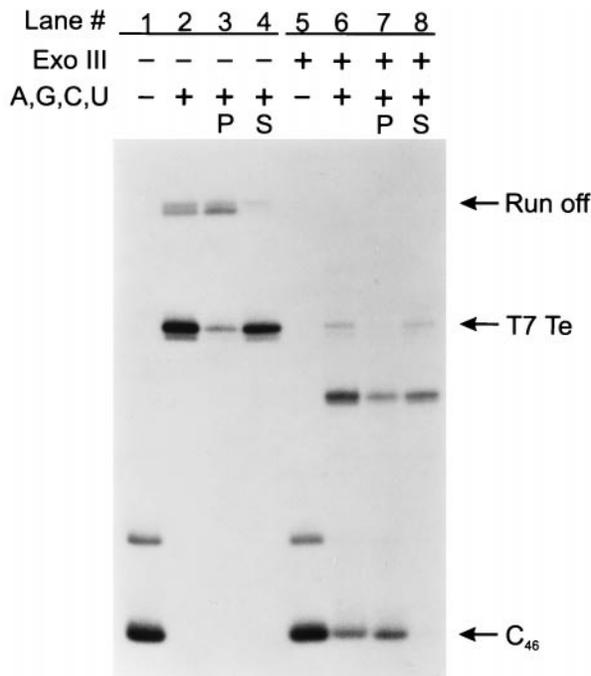


FIG. 3. Assaying for intrinsic transcript termination at T<sub>7</sub>Te on ssDNA. C<sub>46</sub> complexes bound to Ni<sup>2+</sup>-NTA agarose in TGK-B<sub>40</sub>M<sub>4</sub> were chased with 500 µM ATP, 500 µM GTP, 500 µM CTP, and 500 µM UTP for 10 min at 37°C in the presence of rifampicin at 20 µg/ml and yeast *Torula* RNA at 0.8 mg/ml. The lanes are assigned as for Fig. 2A. The T<sub>7</sub>Te terminator is at positions +95 and +96, whereas transcription to the end of the DNA template generates a run-off RNA of +146 nucleotides. C<sub>46</sub> complexes in lanes 5–8 were digested with Exo III at 5,000 units/ml for 5 min at 37°C. Unlike C<sub>47</sub>, some of the C<sub>46</sub> complexes failed to resume elongation after treatment with Exo III (see lanes 6 and 7). The mechanism of this inactivation is unknown but similar observations have been made by others (29, 30).

**Cell Biology.** In the article “Scavenger receptor class B, type I (SR-BI) is the major route for the delivery of high density lipoprotein cholesterol to the steroidogenic pathway in cultured mouse adrenocortical cells” by Ryan E. Temel, Bernardo Trigatti, Ronald B. DeMattos, Salman Azhar, Monty Krieger, and David L. Williams, which appeared in number 25, December 9, 1997, of *Proc. Natl. Acad. Sci. USA* (**94**, 13600–13605), the following correction should be noted. The equation on page 13601 should be as follows:

$$P_{\text{total}} = \frac{[P_{\text{max}}][\text{HDL}]}{K_{\text{HA}} + [\text{HDL}]} + C[\text{HDL}]$$

**Cell Biology.** In the article "Identification of a family of low-affinity insulin-like growth factor binding proteins (IGFBPs): Characterization of connective tissue growth factor as a member of the IGFBP superfamily" by Ho-Seong Kim, Srinivasa R. Nagalla, Youngman Oh, Elizabeth Wilson, Charles T. Roberts, Jr., and Ron G. Rosenfeld, which appeared in number 24, November 25, 1997, of *Proc. Natl. Acad. Sci. USA* (94, 12981–12986), the authors request that the following corrections be noted. In Fig. 8, the units on the scale should indicate the number of substitution events, rather than "million years." The lengths of the branches represent the relative distance between the sequences of mammalian IGFBPs compared in this figure. In the *Discussion*, the statement "The dendrogram depicted in Fig. 8 indicates that, based upon structural similarities, all ten members of the superfamily can be traced back to an ancestor gene 60 million years ago" should read: "The dendrogram depicted in Fig. 8 indicates that all ten members of the superfamily share a common ancestral gene based upon their sequence similarities."

**Neurobiology.** In the article "Hair cell-specific splicing of mRNA for the  $\alpha_{1D}$  subunit of voltage-gated  $Ca^{2+}$  channels in the chicken's cochlea" by Richard Kollmar, John Fak, Lisa G. Montgomery, and A. J. Hudspeth, which appeared in number 26, December 23, 1997, of *Proc. Natl. Acad. Sci. USA* (94, 14889–14893), the authors wish to note that the quality of reproduction of Fig. 1 was below standard. In all three panels,

**Immunology.** In the article "Parasite-mediated nuclear factor  $\kappa$ B regulation in lymphoproliferation caused by *Theileria parva* infection" by Guy H. Palmer, Joel Machado, Jr., Paula Fernandez, Volker Heussler, Therese Perinat, and Dirk A. E. Dobbelaere, which appeared in number 23, November 11, 1997, of *Proc. Natl. Acad. Sci. USA* (94, 12527–12532), the following correction should be noted. The concentration of *N*-acetylcysteine used in the experiments was 30 mM, not 25  $\mu$ g/ml as erroneously reported on page 12528, lines 14 and 15 of the paragraph entitled "Cell Lines and Cultures" in the *Materials and Methods* section.

the middle parts were affected. Specifically, the reverse (white-on-black) type denoting exons 9a, 22a, and 30a was illegible; parts of the arrows that represented primers such as F<sub>9</sub> were missing; and the outlines of several of the boxes that depicted exons such as 9, 10, and 20 were defective. The figure and its legend are reproduced below.

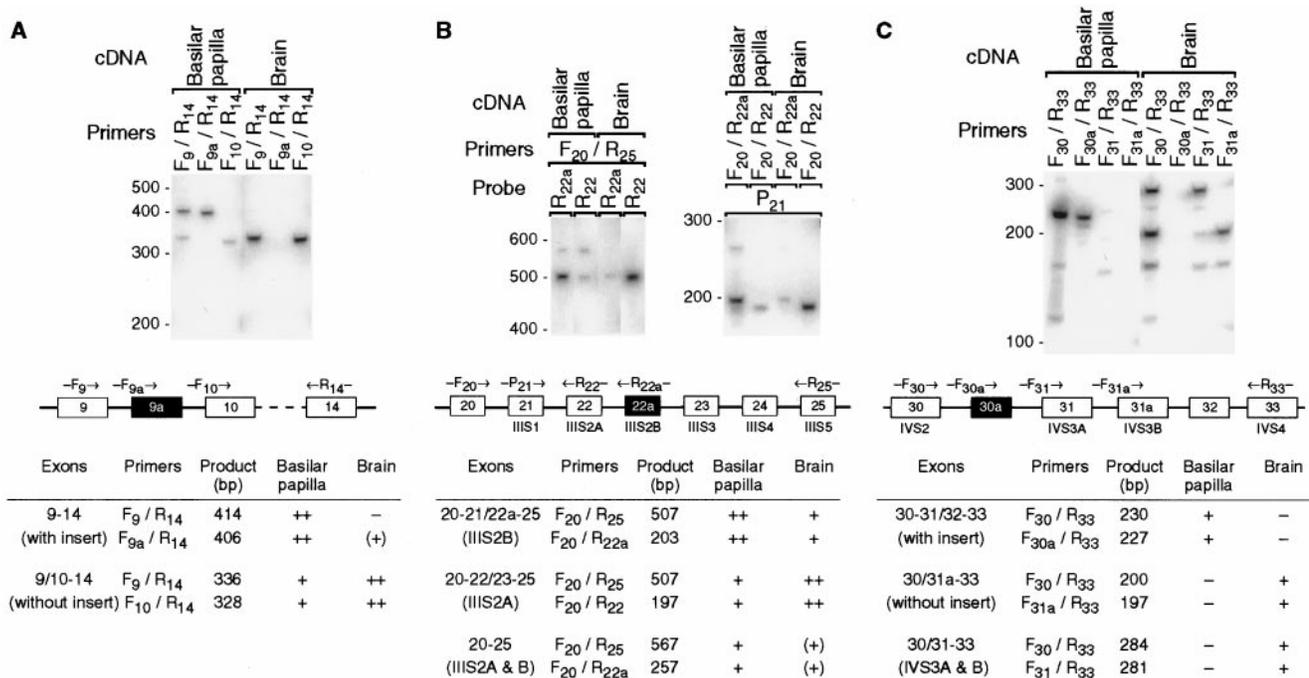


FIG. 1. Alternative splicing of the  $\alpha_{1D}$  mRNA in the basilar papilla and the brain. (A) Southern blot of PCR products amplified with primers flanking the insert in the I-II loop (exon 9a). Marker sizes in base pairs are indicated on the left. The diagram below of the putative genomic structure (not drawn to scale) depicts exons as rectangles, introns as horizontal lines, and PCR primers as arrows. To amplify all isoforms together, we used primers F<sub>9</sub> and R<sub>14</sub>. To amplify rare isoforms without interference from more abundant ones, we used exon-specific primers: primer F<sub>9a</sub> binds across the splice junction of exons 9 and 9a, and primer F<sub>10</sub> binds across that of exons 9 and 10. The table at the bottom lists product size and occurrence for each splice variant and primer pair. ++, abundant; +, detectable; (+), barely so; -, not detectable. (B) Same as A, but for the alternative IIIS2 segment (exon 22a). Note the abundance in the basilar papilla of mRNAs with exons for both IIIS2 segments. (C) Same as A, but for the insert in the IVS2–3 loop (exon 30a). Primer F<sub>30a</sub> binds across the splice junction of exons 30 and 30a, primer F<sub>31</sub> binds across that of exons 30 and 31, and primer F<sub>31a</sub> binds across that of exons 30 and 31a. For the basilar papilla, the lengths of even the minor products were consistent only with splice isoforms containing exon 30a; for the brain, they were consistent only with isoforms lacking exon 30a. Note the abundance in the brain of mRNAs with exons for both IVS3 segments.

# Identification of a family of low-affinity insulin-like growth factor binding proteins (IGFBPs): Characterization of connective tissue growth factor as a member of the IGFBP superfamily

(peptide hormones/serum inducible factors/transforming growth factor  $\beta$ )

HO-SEONG KIM, SRINIVASA R. NAGALLA, YOUNGMAN OH, ELIZABETH WILSON, CHARLES T. ROBERTS, JR., AND RON G. ROSENFELD\*

Department of Pediatrics, School of Medicine, Oregon Health Sciences University, Portland, OR 97201

Communicated by Melvin M. Grumbach, University of California, San Francisco, CA, September 15, 1997 (received for review April 15, 1997)

**ABSTRACT** The insulin-like growth factor (IGF) binding proteins (IGFBPs) modulate the actions of the insulin-like growth factors in endocrine, paracrine, and autocrine settings. Additionally, some IGFBPs appear to exhibit biological effects that are IGF independent. The six high-affinity IGFBPs that have been characterized to date exhibit 40–60% amino acid sequence identity overall, with the most conserved sequences in their NH<sub>2</sub> and COOH termini. We have recently demonstrated that the product of the *mac25/IGFBP-7* gene, which shows significant conservation in the NH<sub>2</sub> terminus, including an “IGFBP motif” (GCGCCXXC), exhibits low-affinity IGF binding. The closely related mammalian genes connective tissue growth factor (CTGF) gene, *nov*, and *cyr61* encode secreted proteins that also contain the conserved sequences and IGFBP motifs in their NH<sub>2</sub> termini. To ascertain if these genes, along with *mac25/IGFBP-7*, encode a family of low-affinity IGFBPs, we assessed the IGF binding characteristics of recombinant human CTGF (rhCTGF). The ability of baculovirus-synthesized rhCTGF to bind IGFs was demonstrated by Western ligand blotting, affinity cross-linking, and competitive affinity binding assays using <sup>125</sup>I-labeled IGF-I or IGF-II and unlabeled IGFs. CTGF, like *mac25/IGFBP-7*, specifically binds IGFs, although with relatively low affinity. On the basis of these data, we propose that CTGF represents another member of the IGFBP family (IGFBP-8) and that the CTGF gene, *mac25/IGFBP-7*, *nov*, and *cyr61* are members of a family of low-affinity IGFBP genes. These genes, along with those encoding the high-affinity IGFBPs 1–6, together constitute an IGFBP superfamily whose products function in IGF-dependent or IGF-independent modes to regulate normal and neoplastic cell growth.

The insulin-like growth factor (IGF) binding proteins (IGFBPs) are a family of homologous proteins that regulate the biological activities of the IGFs and may also be capable of IGF-independent actions. Six distinct IGFBPs that bind IGFs with high affinity have been described (1–7). They share an overall protein sequence identity of 50% and contain 16–18 conserved cysteines in the NH<sub>2</sub>- and COOH-terminal regions (8). We have recently demonstrated that the protein product of the human *mac25* cDNA, which is structurally related and which contains the “IGFBP motif” (GCGCCXXC) in its NH<sub>2</sub> terminus, specifically binds IGFs, although with relatively low affinity, and constitutes another member of the IGFBP family, IGFBP-7 (9). This finding suggests that a family of low-affinity IGFBPs, distinct from the high-affinity members, may exist, and together these may constitute an IGFBP superfamily.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/9412981-6\$2.00/0  
PNAS is available online at <http://www.pnas.org>.

A closely related family of genes encoding connective tissue growth factor (CTGF) (10), the *nov* oncogene (11), and *cyr61* (12) has been identified; the predicted proteins are products of “immediate-early genes” expressed after induction by serum, growth factors, or certain oncogenes (10–13). These proteins show an overall identity of 30–38% compared with IGFBPs 1–6. Although the similarity of the COOH-terminal sequences is low (<20%), the NH<sub>2</sub>-terminal region is well conserved among these new members and the IGFBPs. Moreover, these proteins also contain the conserved “IGFBP motif” (GCGC-CXXC) in their NH<sub>2</sub> terminus, and as many as 17 of the 18 cysteines are conserved in IGFBPs 1–6, suggesting that the CTGF/*nov* oncogene/*cyr61* family shares significant structural homology with IGFBPs (14) and may potentially bind IGFs.

CTGF has been identified as a major chemotactic and mitogenic factor for connective tissue cells (10). It has platelet-derived growth factor (PDGF)-related biological and immunological activities, and it competes with PDGF for a cell-surface receptor (10). The CTGF gene, residing on chromosome 6q23.1, proximal to *c-myc*, was originally cloned from human umbilical vein endothelial cells (10, 11). CTGF gene expression is increased in human foreskin fibroblasts after activation with transforming growth factor  $\beta$  (TGF- $\beta$ ), but not other growth factors, including PDGF, epidermal growth factor, and basic fibroblast growth factor (15). The CTGF gene encodes a 38-kDa prepeptide of 349 amino acids (10).

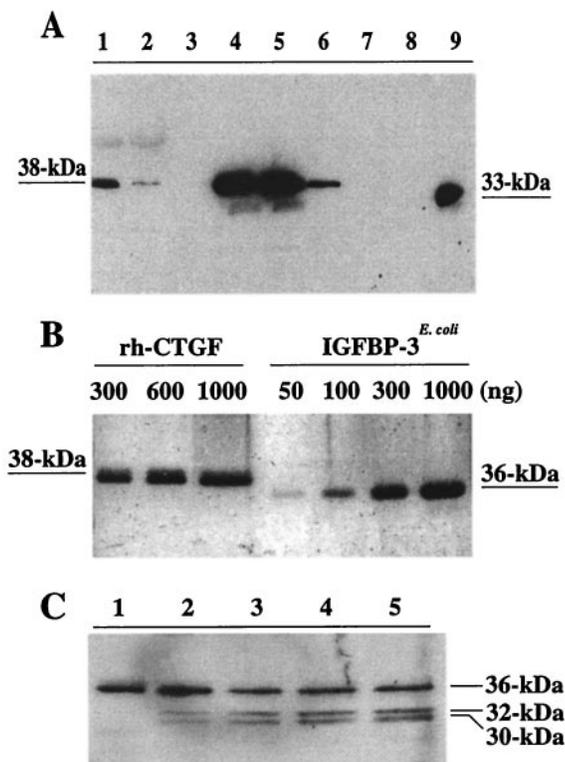
In the present study, we report the expression of recombinant human CTGF (rhCTGF) in a baculovirus system, and we demonstrate that the 36-kDa CTGF protein specifically binds IGFs. CTGF thus meets criteria that define it as another member of the IGFBP superfamily, IGFBP-8. We further propose that the six high-affinity IGFBPs and the four potential low-affinity IGFBPs constitute a superfamily of proteins that regulate cell growth through both IGF-dependent and IGF-independent actions.

## MATERIALS AND METHODS

**Peptides and Proteins.** Recombinant human IGF-I was obtained from Bachem, and recombinant human IGF-II was provided by Eli Lilly. [Gln<sup>6</sup>,Ala<sup>7</sup>,Tyr<sup>18</sup>,Leu<sup>19</sup>,Leu<sup>27</sup>]IGF-II ([QAYLL]IGF-II), a synthetic IGF-II analog, was synthesized as described previously (16). Recombinant human IGFBP-3<sup>*E. coli*</sup>, a nonglycosylated 29-kDa core protein ex-

Abbreviations: IGF, insulin-like growth factor; IGFBP, IGF binding protein; CTGF, connective tissue growth factor; rhCTGF, recombinant human CTGF; PDGF, platelet-derived growth factor; [QAYLL]-IGF-II, [Gln<sup>6</sup>,Ala<sup>7</sup>,Tyr<sup>18</sup>,Leu<sup>19</sup>,Leu<sup>27</sup>]IGF-II; Endo F, endoglycosidase F; WLB, Western ligand blot.

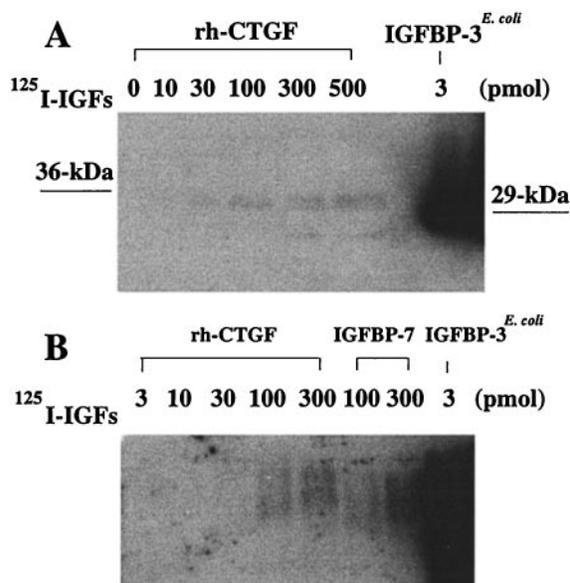
\*To whom reprint requests should be addressed at: Department of Pediatrics, School of Medicine, Oregon Health Sciences University, 3181 S.W. Sam Jackson Park Road, Portland, OR 97201. e-mail: rosenfer@ohsu.edu.



**FIG. 1.** Purification and characterization of FLAG epitope-tagged rhCTGF. (A) Purification and Western immunoblotting of FLAG-tagged rhCTGF using M2 monoclonal antibody under reducing conditions. Lanes represent 10  $\mu$ l of samples: medium from HI-5 infected cells (lane 1), flow-through from the M2 affinity column (lane 2), wash from the same immunoaffinity column (lane 3), serial fractions of rhCTGF eluted by FLAG peptides (lanes 4–8), and 500 ng of FLAG-tagged rh-IGFBP-7 as a positive control (lane 9). (B) Silver staining of purified rhCTGF (lane 1, 300 ng; lane 2, 600 ng; lane 3, 1000 ng) and nonglycosylated IGFBP-3<sup>E. coli</sup> (lane 4, 50 ng; lane 5, 100 ng; lane 6, 300 ng; lane 7, 1000 ng) under reducing conditions. (C) Western immunoblot of rhCTGF with M2 monoclonal antibody after treatment with various concentrations of Endo F at 37°C for 3 hr under nonreducing conditions. Lanes represent 500 ng of rhCTGF alone (lane 1) or after Endo F treatment (100 milliunits of Endo F, lane 2; 200 milliunits, lane 3; 400 milliunits, lane 4; 600 milliunits, lane 5).

pressed in *Escherichia coli*, was a generous gift from Celtrix Laboratories (Santa Clara, CA). Recombinant human IGFBP-7 was expressed in a baculovirus system as described previously (9). Iodination was performed by a modification of the chloramine-T technique, to a specific activity of 350–500  $\mu$ Ci/ $\mu$ g for IGF-I and IGF-II (1  $\mu$ Ci = 37 kBq) (17). Iodinated prolactin was a gift from Diagnostic Systems Laboratories (Webster, TX). Endoglycosidase F (Endo F) was purchased from Boehringer Mannheim. Human multiple tissue Northern blots were obtained from CLONTECH.

**Cloning of Human CTGF, *nov* Oncogene, and *cyr61* cDNAs.** To characterize the family of low-affinity IGFBPs, partial cDNAs corresponding to all three members of this family were generated by using reverse transcription-PCR (RT-PCR) as described previously (18). In brief, degenerate primers were designed to conserved sequences in this family: 5' primer GGNAT(A/C/T)(T/A)(C/G)NACN(A/C)GNGTNACNAA(C/T)GA(C/T)AA, corresponding to amino acids GIS-TRVTNDN, and 3' primer GGNGT(A/G)CA(G/A)CAN(C/T/G)NCC(G/A)TC, corresponding to amino acids DGRCTP (5' amino acids 226–236 and 3' amino acids 298–308 of GenBank sequence M92934). Five micrograms of total RNA isolated from the Hs578T human breast carcinoma cell line was used in RT-PCR as described previously (18). Fragments corresponding to all three cDNAs were identified



**FIG. 2.** (A) WLBs of various amounts of rhCTGF (0–500 pmol) and 3 pmol of IGFBP-3<sup>E. coli</sup> with <sup>125</sup>I-IGFs after denaturing nonreducing SDS/PAGE. (B) WLBs of various amounts of rhCTGF, IGFBP-7 (100 and 300 pmol), and 3 pmol of IGFBP-3<sup>E. coli</sup> with <sup>125</sup>I-IGFs after nonreducing PAGE.

by using gene-specific primers and by sequencing. Partial cDNAs were used to screen a cDNA library prepared from Hs578T cell line mRNA in the  $\lambda$ ZAP Express vector (Stratagene), and full-length cDNAs were isolated as described previously (19). The sequences of full-length CTGF and *nov* oncogene cDNAs were identical to the published sequences (10, 11), and the human *cyr61* sequence (S.R.N., unpublished data) shares 90% identity to mouse *cyr61* (12).

**Expression of rhCTGF Protein.** To facilitate purification of recombinant CTGF, a FLAG epitope sequence (DYKD-DDDK) was added at the COOH terminus by use of PCR. Primers 5'-GTCAGGCCTTGCGAAGCTGAC (845–868) and 5'-CGTGGTACCTTACTTGTTCATCGTCGTCCTTGTAGTCTGCCATGTCTCCGTACAT (1160–1179), which was designed with the FLAG sequence followed by a stop codon and a restriction site for *Kpn*I, were used. The resulting PCR product was digested with unique restriction enzymes *Stu*I and *Kpn*I and ligated into full-length cDNA digested with the same restriction enzymes, to replace the COOH terminus. After sequencing, the FLAG-tagged CTGF cDNA was subcloned into the baculovirus expression vector pFASTBAC1 (Life Technologies). The CTGF-pFASTBAC1 construct was transfected into Sf9 cells and positive viral recombinants were isolated by using the vendor's protocols. Western immunoblotting was performed with the FLAG sequence-specific anti-FLAG M2 antibody (Eastman Kodak) to confirm the expression of rhCTGF protein.

**Protein Purification.** Large-scale protein purification was performed by infecting 10<sup>8</sup> HI-5 insect cells at a multiplicity of infection of 3 at 27°C for 2 days. The media from the infected cells were collected and centrifuged. The supernatant was bound to an anti-FLAG M2 antibody affinity column at 4°C for 2 hr. The column was washed three times with 5 ml of HBS (20 mM HEPES, pH 7.8/150 mM NaCl), and the protein was eluted with four 1-ml washes with HBS containing 0.5 mg/ml FLAG peptide. The purified protein was subjected to sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS/PAGE) in a 12% polyacrylamide gel and stained with Coomassie blue or transferred to nitrocellulose membrane for immunodetection.

Eluted fractions from an anti-FLAG M2 antibody affinity column were pooled and quantitated by using a densitometer

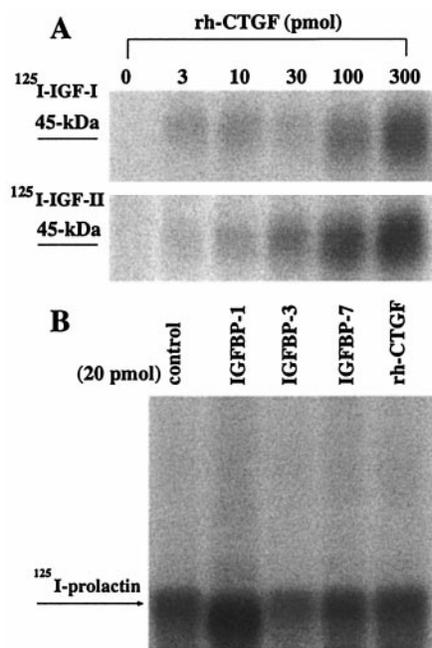


FIG. 3. Affinity cross-linking of rhCTGF. (A) Autoradiogram of  $^{125}\text{I}$ -IGF-I (Upper) or  $^{125}\text{I}$ -IGF-II (Lower) cross-linked to rhCTGF. An approximately 45-kDa band can be seen that is compatible with the 38-kDa rhCTGF bound to 7-kDa  $^{125}\text{I}$ -IGF-I or  $^{125}\text{I}$ -IGF-II. (B) Autoradiogram of  $^{125}\text{I}$ -prolactin cross-linked to 20 pmol of rh-IGFBP-1 (lane 2), rh-IGFBP-3 (lane 3), rh-IGFBP-7 (lane 4) or rhCTGF (lane 5). No binding of prolactin to any of the IGFBPs was observed. The band seen at 33 kDa represents unbound  $^{125}\text{I}$ -prolactin.

and comparing with known amounts of IGFBP-3<sup>E. coli</sup> after silver staining.

**Glycosylation Studies.** Proteins were deglycosylated with Endo F (9). As a positive control, acid-chromatographed normal human serum IGFBP fractions were prepared (20). Five hundred nanograms of FLAG-tagged-rhCTGF and 2  $\mu\text{l}$  of acid-chromatographed normal human serum IGFBP fraction were treated with 100–600 milliunits of Endo F as described previously (9). Subsequent Western immunoblots or ligand blots were prepared as described below.

**Western Ligand Blots (WLBs).** Unreduced samples of FLAG-tagged rhCTGF and IGFBP-3<sup>E. coli</sup> were subjected to SDS/PAGE in a 12% gel and electroblotted onto nitrocellulose membranes. The membranes were incubated overnight with  $1.0 \times 10^6$  cpm of  $^{125}\text{I}$ -labeled IGF-I and IGF-II after blocking with normal saline containing 1% BSA at room temperature for 2 hr. The filters were then washed with normal saline containing 0.1% Tween 20 two times and normal saline three times for 15 min each time, dried, and exposed to film. Nondenaturing WLBs were prepared by PAGE without SDS, in stacking (pH 7.4) or resolving (pH 8.8) gels (21).

**Affinity Cross-Linking.** FLAG-tagged rhCTGF and IGFBP-3<sup>E. coli</sup> were incubated overnight at 4°C with  $^{125}\text{I}$ -labeled IGF-I or IGF-II or prolactin (100,000 cpm) in the presence or absence of unlabeled peptides at the concentrations indicated in the text and figures. After cross-linking with disuccinimidyl suberate (0.6 mM), samples were subjected to SDS/PAGE and autoradiography (17). Bands were quantified by densitometry using the area under the curve, as calculated by an LKB densitometer.

**Northern Blot Analysis.** Blots of 2  $\mu\text{g}$  of poly(A)<sup>+</sup> RNA from normal human tissues, which had been subjected to electrophoresis in a formaldehyde/1.5% agarose gel before transfer to nylon membranes, were purchased from CLONTECH.  $^{32}\text{P}$ -labeled antisense complementary RNA probes for CTGF, transcribed from the plasmid constructs, were used.

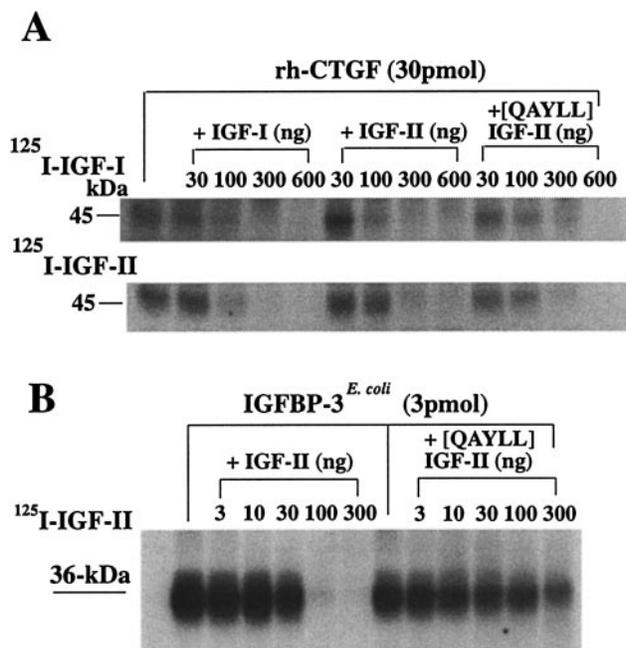


FIG. 4. Competitive affinity cross-linking of rhCTGF or IGFBP-3<sup>E. coli</sup>. (A) Autoradiogram of  $^{125}\text{I}$ -IGF-I (Upper) or  $^{125}\text{I}$ -IGF-II (Lower) cross-linked to rhCTGF. Radiolabeled ligands ( $1.0 \times 10^5$  cpm) were incubated with 30 pmol of rhCTGF alone or in the presence of unlabeled IGF-I, IGF-II or [QAYLL]IGF-II, which has 1/100 the affinity for IGFBPs 1–6. (B) Autoradiogram of  $^{125}\text{I}$ -IGF-II cross-linked to 3 pmol of IGFBP-3<sup>E. coli</sup> in the absence or presence of different amounts of unlabeled IGF-II or [QAYLL]IGF-II.

Blots were hybridized and washed at high stringency as described previously (18).

## RESULTS

**Construction and Expression of FLAG Epitope-Tagged rhCTGF.** The immunoblot of the fractions collected during purification of FLAG-tagged rhCTGF by reducing SDS/12% PAGE (Fig. 1A) identifies an M2 antibody-specific protein of 38 kDa, which is compatible with the predicted size of CTGF as reported previously (10). Application of various amounts of the purified rhCTGF protein on a reducing SDS/12% polyacrylamide gel and subsequent silver staining (Fig. 1B) showed that the protein has approximately 99% purity and a molecular mass of approximately 36–38 kDa.

Because the predicted amino acid sequence analysis revealed that CTGF contains two potential N-glycosylation sites, located at amino acids 28 and 225 (*Asn*-Cys-Ser, *Asn*-Ala-Ser, respectively), we treated rhCTGF with various concentrations

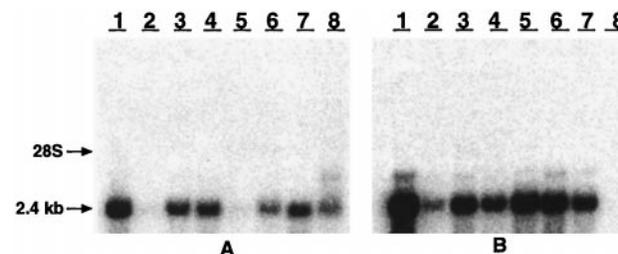


FIG. 5. Northern blot analysis of CTGF in normal human tissues. Human multiple tissue blots were probed with CTGF complementary RNA and exposed to Biomax film for 2 hr at  $-80^\circ\text{C}$ . (A) Lanes 1–8 are as follows: 1, heart; 2, brain; 3, placenta; 4, lung; 5, liver; 6, skeletal muscle; 7, kidney; and 8, pancreas. (B) Lanes 1–8: 1, spleen; 2, thymus; 3, prostate; 4, testis; 5, ovary; 6, small intestine; 7, colon; and 8, peripheral blood leukocyte.

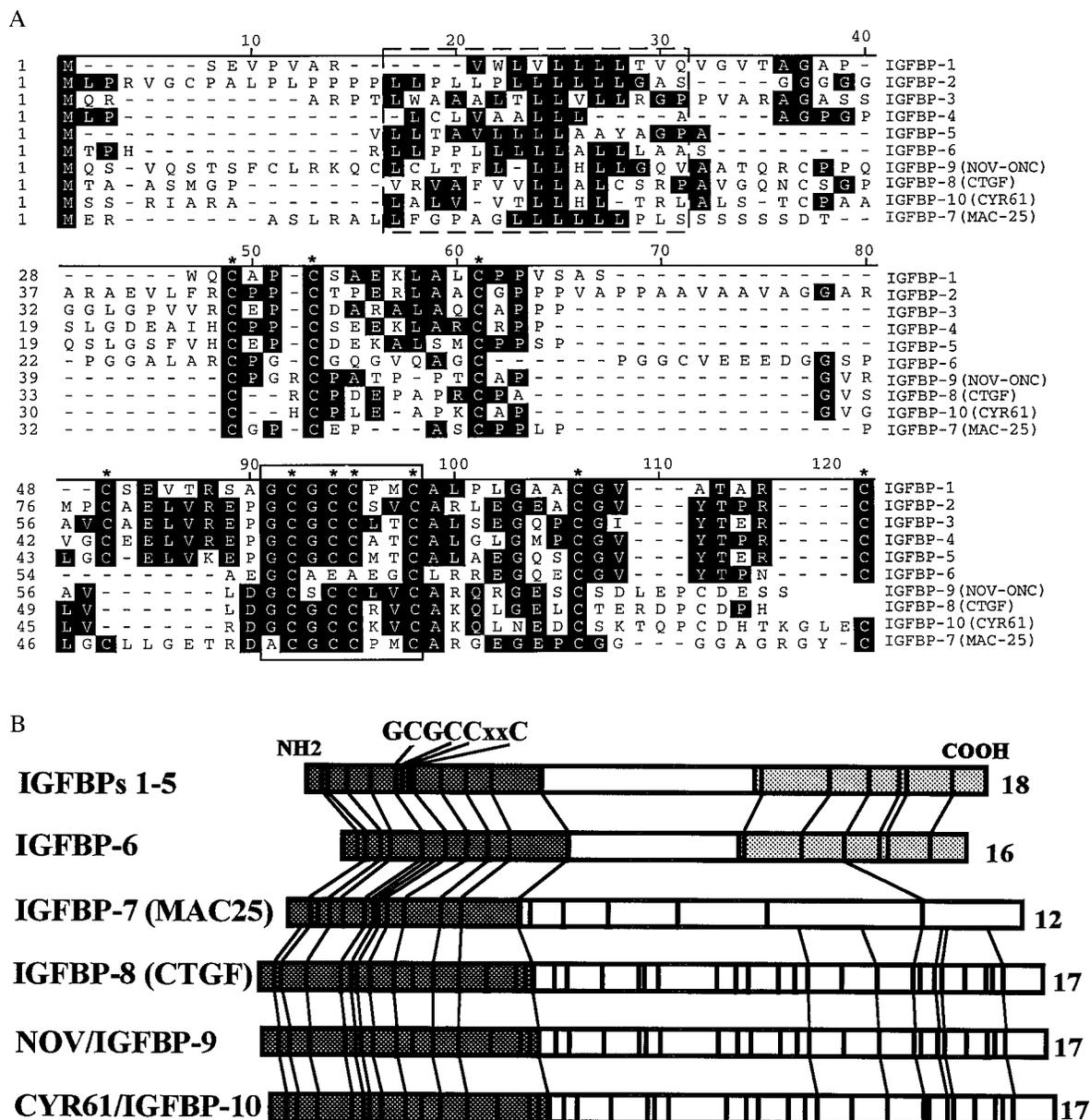


FIG. 6. Comparison of NH<sub>2</sub>-terminal amino acid sequences (A) and total cysteine residues (B) of IGFBP superfamily. (A) Sequences of IGFBPs 1-7 (1-7), CTGF (10), *nov* oncogene (22), and *cyr61* (S.R.N., unpublished data) were aligned using the DNASTAR MEGALIGN program (Madison, WI). Consensus signal sequences are boxed by a broken line. The characteristic IGFBP motif (GCGCCXXC) is boxed by a solid line. Conserved cysteine residues in the family are marked with asterisks. We have named mac-25 IGFBP-7 and have named CTGF IGFBP-8. We have tentatively named the protein products of the *nov* and *cyr61* genes IGFBP-9 and IGFBP-10, respectively. (B) The conservation of the cysteine residues is shown by the vertical lines. The numbers on the right represent the total conserved cysteines among IGFBPs. The shaded regions represent the conserved domain of IGFBPs in their NH<sub>2</sub> or COOH terminus.

of Endo F (100-600 milliunits) to cleave the N-glycosylated carbohydrates. The size of rhCTGF (36 kDa under nonreducing conditions) was decreased to approximately 32 kDa and 30 kDa by treatment with 100-600 milliunits of Endo F, indicating that the secreted rhCTGF is a glycosylated protein with 2 kDa and 4 kDa of N-linked sugars and a 30-kDa core (Fig. 1C).

**Characterization of rhCTGF as IGFBP-8.** We prepared WLBs and performed IGF affinity cross-linking to characterize the affinity of rhCTGF for IGF peptides. In WLBs with <sup>125</sup>I-labeled IGF-I and IGF-II under denaturing conditions, a 36-kDa band was detected under nonreducing conditions at amounts of rhCTGF ranging from 10 to 500 pmol, although with considerably lower binding ability than observed with IGFBP-3<sup>E. coli</sup> (Fig. 2A). A minor band at 26 kDa presumably represents a CTGF degradation product. In WLB with <sup>125</sup>I-IGF-I and <sup>125</sup>I-IGF-II after nondenaturing polyacrylamide gel

electrophoresis, bands were also detected at CTGF amounts ranging from 30 to 300 pmol (Fig. 2B).

When we performed affinity cross-linking with <sup>125</sup>I-IGF-I or -IGF-II and concentrations of rhCTGF similar to those used in WLB, a 45-kDa band was detected on a reducing SDS/12% polyacrylamide gel with concentrations of rhCTGF as low as 3 pmol (Fig. 3A). This 45-kDa band is consistent with the size of 38-kDa rhCTGF bound to 7-kDa <sup>125</sup>I-IGF-I or <sup>125</sup>I-IGF-II. Binding of <sup>125</sup>I-IGF-I and <sup>125</sup>I-IGF-II was linearly dependent on CTGF amount ranging from 3 to 300 pmol. To exclude the possibility that rhCTGF nonspecifically binds IGFs, we performed affinity cross-linking with <sup>125</sup>I-prolactin and rhCTGF or other IGFBPs (IGFBPs-1, -3, -7). As shown in Fig. 3B, no band was detected for prolactin binding, indicating that the binding of IGF by CTGF is specific. Specificity of binding was further confirmed by competitive affinity binding assays using

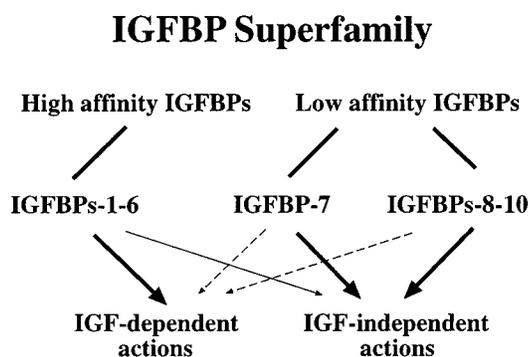


FIG. 7. The IGFBP superfamily composed of high-affinity IGFBPs and low-affinity IGFBPs. The thick and thin arrows indicate proposed primary and secondary biological actions, respectively, and the dashed arrows represent potential actions of IGFBPs that have not been verified experimentally.

$^{125}\text{I}$ -IGFs and unlabeled IGF-I and IGF-II (Fig. 4A). Displacement of  $^{125}\text{I}$ -IGF-I or -II from CTGF was observed at IGF-I and IGF-II amounts of 100–600 ng per lane. These data clearly demonstrate that the affinity of rhCTGF for IGF-I and IGF-II is significantly lower than that of rh-IGFBP-3<sup>E. coli</sup>. Interestingly, [QAYLL]IGF-II, which binds to IGFBPs 1–6 with 1/100 the affinity of IGF-II, appears to have an affinity for CTGF approximately equal to that of IGF-II (Fig. 4B).

**Expression of CTGF mRNA in Normal Human Tissues.** We performed Northern blot analysis to assess the expression of the CTGF gene in normal human tissues. The 2.4-kb CTGF mRNA was detected in a broad spectrum of normal human tissues (Fig. 5). In particular, CTGF mRNA was expressed at high levels in spleen, ovary, gastrointestinal tract, prostate, heart, and testis.

## DISCUSSION

CTGF is a cysteine-rich mitogen secreted by human umbilical vein endothelial (HUVE) cells. It was initially purified from conditioned media of HUVE cells subjected to PDGF-IgG affinity chromatography, but it was shown to not be composed of PDGF A or B chain peptides (10). Subsequent expression screening of an HUVE cell cDNA library with the anti-PDGF antibody led to the cloning and sequencing of a cDNA with an open reading frame encoding a 38-kDa protein (10). This cysteine-rich protein was shown to be the major PDGF-related mitogen and chemotactic factor secreted by HUVE cells and to compete with PDGF for binding to the PDGF cell-surface receptor on fibroblasts (10, 22). CTGF has, in fact, little peptide sequence homology with PDGF; it is believed that CTGF and PDGF monomers must share ternary structure, resulting in both common antigenic epitopes and competition for receptor binding. CTGF contains 39 cysteine residues, suggesting the presence of multiple intramolecular disulfide bonds and a complex ternary structure (14). The hydrophobic NH<sub>2</sub>-terminal sequence is consistent with its potential role as a signal peptide directing the secretion of processed CTGF.

Bork (14) has noted that CTGF is one of six different proteins, varying between 348 and 379 amino acids (including the signal peptides), that are the products of a group of “immediate-early genes” expressed after induction by growth factors or certain oncogenes. These genes are (i) human CTGF (10) and mouse CTGF (also known as *fisp-12* or  $\beta\text{IG-M2}$ ) (23), (ii) the chicken (24) and human (11) *nov* oncogenes, and (iii) the chicken gene *cef10* (25) and the related mouse gene *cyr61* (also known as  $\beta\text{IG-M1}$ ) (12). Thirty-eight cysteines are conserved among the six proteins. It was further noted that these proteins are characterized by a modular architecture, with domains homologous to (i) IGFBPs, (ii) the von Willebrand factor type C repeat, (iii) thrombospondin type I repeat, and (iv) a COOH-terminal module. In addition to conferring the potential for binding IGF peptides, these modules may be involved in oligomerization, cell attachment through binding motifs for sulfated glycoconjugates, and dimerization (14).

A number of characteristics of this family of proteins are of potential relevance to the IGFBPs: (i) as has been suggested for some of the IGFBPs, these proteins may be capable of IGF-independent regulation of cell growth (16, 17, 26); (ii) several of the proteins have been shown to interact with both cell surfaces and extracellular matrix and to be capable of binding to heparin, properties also shared by some members of the IGFBP family (27); and (iii) several genes from this family are induced by transforming growth factor  $\beta$ , as is the case for some IGFBPs (15, 28). Most importantly, these proteins share significant sequence homology with the IGFBPs, including preservation of as many as 17 of the 18 cysteines conserved in IGFBPs 1–6, and conservation of the “GCGCCXXC” motif found in the NH<sub>2</sub> terminus of the IGFBPs (Fig. 6).

CTGF shares an overall 28–38% amino acid identity to IGFBPs 1–6. Both WLB, after either denaturing or nondenaturing PAGE, and affinity cross-linking demonstrate that CTGF specifically binds IGF-I and -II, although with relatively low affinity as compared with IGFBPs 1–6. Nevertheless, the specificity of IGF binding, coupled with the structural similarity to IGFBPs 1–7, indicates that CTGF meets the criteria necessary to define it as a member of the IGFBP family, IGFBP-8. Specificity of binding was demonstrated by the failure of CTGF (like IGFBPs 1, 3, and 7) to bind prolactin, as well as by the dose-dependent inhibition of binding by unlabeled IGF-I or -II. Interestingly, although the affinity of IGFBP-8 for IGFs is significantly reduced, the affinity of IGFBP-8 for [QAYLL]IGF-II appears to be approximately equivalent to its affinity for IGF-II, in contrast to the situation for IGFBPs 1–6, where [QAYLL]IGF-II binds with 1/100 the affinity. These results suggest that the IGF binding site in IGFBP-8 may differ somewhat from the binding site in IGFBPs 1–6. Whether this reflects the larger, more complex structure of IGFBP-8, or results from the less rigorous conservation of structures in the NH<sub>2</sub> terminus or lack of conserved sequences in the COOH terminus is unclear. It is of note that the IGFBP-3 fragments derived from proteolysis appear to exhibit significantly reduced affinity, similar to that of IGFBP-8 (29). Furthermore, a baculovirus-expressed IGFBP-3 fragment corresponding to the NH<sub>2</sub>-terminal portion of IGFBP-3 (amino acids 1–97) was able to bind IGF-I and

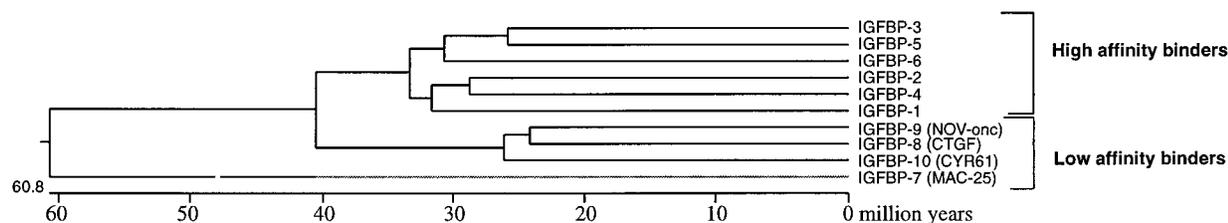


FIG. 8. Phylogenetic tree of the IGFBP superfamily, derived from multiple comparisons of full-length sequences of known IGFBPs 1–7, CTGF, *nov* oncogene, and *cyr61*, using the method of Hein (30) as implemented by the DNASTAR MEGALIGN program.

IGF-II with low affinity, which is comparable to those of low-affinity IGF binders (IGFBPs 7 and 8) (unpublished results). Nevertheless, it appears that the IGFBP superfamily, which now includes 10 potential members, can be divided into at least two subgroups: high-affinity IGF binders (IGFBPs 1–6) and low-affinity IGF binders (IGFBPs 7 and 8 and, potentially, the protein products of the *nov* and *cyr61* genes) (Fig. 7). Although the latter two proteins have yet to be evaluated for IGF affinity, given their structural similarity with CTGF/IGFBP-8, it seems reasonable to consider them as potential members of the IGFBP superfamily, pending further studies.

The dendrogram depicted in Fig. 8 indicates that, on the basis of structural similarities, all 10 members of the superfamily can be traced back to an ancestral gene approximately 60 million years ago (30). IGFBP-7 is the most divergent, followed by the CTGF (IGFBP-8), *nov* oncogene (provisionally IGFBP-9) and *cyr61* (provisionally IGFBP-10). It is of note that IGFBPs or IGF-like activity have been identified in fish, amphibians, reptiles, and birds, indicating that significant conservation exists among vertebrates (31). The detection of an IGFBP in lamprey serum suggests that IGFBPs have been present since early vertebrate evolution (32). On the other hand, the fact that the genes for IGFBPs 1–8 are located on six different chromosomes demonstrates that some degree of divergence has also occurred.

The existence of an IGFBP superfamily has important implications for our understanding of the biological roles of these proteins, including the conventional IGFBPs. While the lower affinity for IGFs exhibited by IGFBP-7 and -8 does not exclude a role for them as IGF transporters or modulators of IGF action, it is more likely that these proteins regulate cell proliferation and chemotaxis in an IGF-independent manner. It is noteworthy that a number of recent studies have indicated that several of the conventional IGFBPs, most notably IGFBP-3, are also capable of IGF-independent effects, and that these actions may be of importance in modulation of normal and malignant cell growth (16, 17, 26).

We speculate that the IGFBP superfamily is derived from an ancestral gene/protein that was critically involved in the regulation of cell growth and was capable of binding IGF peptides. Over the course of evolution, some members evolved into high-affinity IGF binders and others into low-affinity IGF binders, thereby conferring on the IGFBP superfamily the ability to influence cell growth by both IGF-dependent and IGF-independent mechanisms.

This work was supported in part by National Institutes of Health Grant CA58110 (to R.G.R.), National Institutes of Health Grant DK50810 (to C.T.R.), and U.S. Army Grant DAMD17-96-1-6204 (to Y.O.).

1. Lee, Y.-L., Hintz, R. L., James, P. M., Lee, P. D. K., Shively, J. E. & Powell, D. R. (1988) *Mol. Endocrinol.* **2**, 404–411.
2. Binkert, C., Landwehr, J., Mary, J.-L., Schwander, J. & Heinrich, G. (1989) *EMBO J.* **8**, 2497–2502.

3. Wood, W. I., Cachianes, G., Henzel, W. J., Winslow, G. A., Spencer, S. A., Hellmiss, R., Martin, J. L. & Baxter, R. C. (1989) *Mol. Endocrinol.* **2**, 1176–1185.
4. LaTour, D., Mohan, S., Linkhart, T. A., Baylink, D. J. & Strong, D. D. (1990) *Mol. Endocrinol.* **4**, 1806–1814.
5. Kiefer, M. C., Ioh, R. S., Bauer, D. M. & Zapf, J. (1991) *Biochem. Biophys. Res. Commun.* **176**, 219–225.
6. Kiefer, M. C., Masiarz, F. R., Bauer, D. M. & Zapf, J. (1991) *J. Biol. Chem.* **266**, 9043–9049.
7. Shimasaki, S., Shimonaka, M., Zhang, H.-P. & Ling, N. (1991) *J. Biol. Chem.* **266**, 10646–10653.
8. Shimasaki, S. & Ling, N. (1992) *Prog. Growth Factor Res.* **3**, 243–266.
9. Oh, Y., Nagalla, S. R., Yamanaka, Y., Kim, H. S., Wilson, E. & Rosenfeld, R. G. (1996) *J. Biol. Chem.* **271**, 30322–30325.
10. Bradham, D. M., Igarashi, A., Potter, R. L. & Grotendorst, G. R. (1991) *J. Cell Biol.* **114**, 1285–1294.
11. Martinierie, C., Viegas-Pequignot, E., Guenard, I., Dutrillaux, B., Nguyen, V. C., Bernheim, A. & Perbal, B. (1992) *Oncogene* **7**, 2529–2534.
12. O'Brien, T. P., Yang, G. P., Sanders, L. & Lau, L. F. (1990) *Mol. Cell. Biol.* **10**, 3569–3577.
13. Kireeva, M. L., Mo, F.-E., Yang, G. P. & Lau, L. F. (1996) *Mol. Cell. Biol.* **16**, 1326–1334.
14. Bork, P. (1993) *FEBS Lett.* **327**, 125–130.
15. Igarashi, A., Okochi, H., Bradham, D. M. & Grotendorst, G. R. (1993) *Mol. Biol. Cell* **4**, 637–645.
16. Oh, Y., Muller, H. L., Zhang, H., Ling, N. & Rosenfeld, R. G. (1994) *Adv. Exp. Med. Biol.* **343**, 41–54.
17. Oh, Y., Muller, H. L., Pham, H. M. & Rosenfeld, R. G. (1993) *J. Biol. Chem.* **268**, 26045–26048.
18. Nagalla, S. R., Barry, B. J., Falick, A. M., Gibson, B. W., Taylor, J. E., Dong, J. Z. & Spindel, E. R. (1996) *J. Biol. Chem.* **271**, 7731–7737.
19. Nagalla, S. R., Barry, B. J. & Spindel, E. R. (1994) *Mol. Endocrinol.* **8**, 943–951.
20. Gargosky, S. E., Pham, H. P., Wilson, K. F., Liu, F., Giudice, L. C. & Rosenfeld, R. G. (1992) *Endocrinology* **131**, 3051–3060.
21. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
22. DiCorleto, P. E. (1984) *Exp. Cell Res.* **153**, 167–172.
23. Ryseck, R.-P., Macdonald-Bravo, H., Mattei, M.-G. & Bravo, R. (1991) *Cell Growth Differ.* **2**, 225–233.
24. Joliot, V., Martinierie, C., Dambrine, G., Plassiat, G., Brisac, M., Crochet, J. & Perbal, B. (1992) *Mol. Cell. Biol.* **12**, 10–21.
25. Simmons, D. L., Levy, D. B., Yannoni, Y. & Erikson, R. L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1178–1182.
26. Oh, Y., Muller, H. L., Lamson, G. & Rosenfeld, R. G. (1993) *J. Biol. Chem.* **268**, 14964–14971.
27. Yang, G. P. & Lau, L. F. (1991) *Cell Growth Differ.* **2**, 351–357.
28. Oh, Y., Muller, H. L., Ng, L. & Rosenfeld, R. G. (1995) *J. Biol. Chem.* **270**, 13589–13592.
29. Lalou, C., Lassarre, C. & Binoux, M. (1996) *Endocrinology* **137**, 3206–3212.
30. Hein, J. (1990) *Methods Enzymol.* **183**, 626–645.
31. Kelley, K. M., Oh, Y., Gargosky, S. E., Gucev, Z., Matsumoto, T., Hwa, V., Ng, L., Simpson, D. M. & Rosenfeld, R. G. (1996) *Intl. J. Biochem. Cell Biol.* **28**, 619–637.
32. Upton, A., Chan, S. J., Steiner, D. F., Wallace, J. C. & Ballard, F. J. (1992) *Growth Regul.* **3**, 29–32.