A Novel Human Zinc Finger Protein That Interacts with the Core Promoter Element of a TATA Box-less Gene*

(Received for publication, September 17, 1996, and in revised form, January 6, 1997)

Nicolás P. Koritschoner‡, José L. Bocco, Graciela M. Panzetta-Dutari, Catherine I. Dumur§, Alfredo Flury, and Luis C. Patrício¶

From the Departamento de Bioquímica Clínica, Facultad de Ciencias Químicas, Universidad Nacional de Córdoba, Córdoba, Argentina

We describe a novel human cDNA isolated by target site screening of a placental expression library, using as a probe, an essential element of a TATA box-less promoter corresponding to a pregnancy-specific glycoprotein gene. The cDNA encoded a predicted protein of 290 amino acids, designated core promoter-binding protein (CPBP), which has three zinc fingers (type Cys2-His2) at the end of its C-terminal domain, a serine/threonine-rich central region and an acidic domain lying within the N-terminal region. Additional sequence analysis and data base searches revealed that only the zinc finger domains are conserved (60–80% identity) in other transcription factors. In cotransfection assays, CPBP increased the transcription from a minimal promoter containing its natural DNA-binding site. Moreover, a chimeric protein between CPBP and Gal4 DNA binding domain also increased the activity of an heterologous reporter gene containing Gal4 DNA binding sites. The tissue distribution analysis of CPBP mRNA revealed that it is differentially expressed with an apparent enrichment in placental cells. The DNA binding and transcriptional activity of CPBP, in conjunction with its expression pattern, strongly suggests that this protein may participate in the regulation and/or maintenance of the basal expression of PSG and possibly other TATA box-less genes.

The molecular mechanisms involved in the transcription of eukaryotic genes are controlled by the ordered interplay of DNA-protein and protein-protein contacts. The factors responsible for basal RNA-polymerase II transcription reaction are the core promoter elements and the general transcription factors (1, 2). In addition, the regulation of transcription rates results from the combined action of activating and repressing proteins that are brought to promoters, enhancers, and silencers through their interactions with specific sequences and sub...

* This work was supported by grants from the Consejo Nacional de Investigaciones Científicas y Tecnológicas of Argentina, the Consejo de Investigaciones Científicas y Tecnológicas de la Provincia de Córdoba, and the Secretaría de Ciencia y Técnica of the Universidad Nacional de Córdoba (SECyT). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ Recipient of a Consejo Nacional de Investigaciones Científicas y Técnicas fellowship. Present address: Max-Planck-Center for Molecular Medicine, Robert Rössle Str. 10, 13125 Berlin, Germany.

§ Recipient of a SECyT fellowship.

To whom correspondence should be addressed: Dept. de Bioquímica Clínica, Facultad de Ciencias Químicas, Universidad Nacional de Córdoba, C.C. 61-A.P. 4 (5000), Córdoba, Argentina. Tel.: 54-51-334164; Fax: 54-51-334174; E-mail: lapatrito@fqc.uncor.edu.

This paper is available online at http://www-jbc.stanford.edu/jbc/

1 The abbreviations used are: bp, base pair(s); kb, kilobase pair(s); IPTG, isopropyl-1-thio-β-D-galactopyranoside; PSG, pregnancy-specific glycoprotein; CPE, core promoter element; CPBP, core promoter-binding protein; PAGE, polyacrylamide gel electrophoresis; EMSA, electrophoretic mobility shift assay; PCR, polymerase chain reaction; CAT, chloramphenicol acetyltransferase; EKLF, erythroid Kruppel-like factor; BTEB, basic transcription element-binding protein.
(CPE). This cDNA encodes a novel protein that has some attractive features that may be used to predict its function as a transcription factor acting through its natural promoter context in a tissue- and gene-specific fashion.

EXPERIMENTAL PROCEDURES

Oligonucleotide Probes and Competitors—The sequences of the oligonucleotides used are as follows.

\[
\text{CPE} \quad 5'\text{AATTCTGCCCCCCTAGAAGGCATGTG}GTT\text{G}3' \\
3'\text{GACTGGTGGGTTACGATCGACTTCCAGC}5'
\]

SEQUENCE 1

\[
\text{CPEmut} \quad 5'\text{AATTCTGCCCCC}G\text{atatagGCGGCT}G\text{GAAGTGC}3' \\
3'\text{GACTGGTGGGTTACGATCGACTTCCAGC}5'
\]

SEQUENCE 2

Competition experiments were performed using different double-stranded oligonucleotides with the following sequences: NS1,5'-AAAGGAGCCATATAAAGGAAAT-3' and its complement NS2,5'-AAGATGATGTGAAGCTGCCTAC-3' and its complement NS3,5'-AATTCCCCCCTAGACATTACGCC-3', with the complementary strand 5'-GATCCGGTGACGTCACGGGGG-3'.

Screening of the cDNA Expression Library with a Specific DNA Target Site Probe—The human placental cDNA library constructed in the expression vector agt11 was kindly provided by Dr. J. L. Millian (14). The library contains approximately 106 independent clones (95% recombimants) with an average insert size of 1.3 kb and was amplified as described previously (15). Protein replica membranes were prepared according to the procedure described by Singh (13). We applied our previous data concerning binding assays of immobilized DNA-binding proteins to improve the conditions for a reliable and specific DNA-protein interaction. Briefly, the filters were incubated for at least 5 h at 4 °C in solution A (10 mM HEPES, pH 7.9, 50 mM KCl, 0.1 mM EDTA, 1 mM dithiothreitol, 10% (v/v) glycerol) supplemented with 5% nonfat dry milk. For screening, the treated filters were immediately immersed in separate recipient containing aliquots (15–20 ml) of solution A and exposed for autoradiography. The filters were washed twice (10 min each wash) with solution A and exposed for autoradiography. Least 4 °C with gentle agitation, the filters were washed twice (10 min each wash) with solution A and exposed for autoradiography. The filters were washed twice (10 min each wash) with solution A and exposed for autoradiography. Least 4 °C with gentle agitation, the filters were washed twice (10 min each wash) with solution A and exposed for autoradiography.

Confirmation of Protein Extracts Derived from Escherichia coli Lysogens—The generation of E. coli (Y1088) lysogens was achieved according to the procedure described in Glover (17), and those bacterial clones harboring agt11, ACPBP, and Anr phages were isolated and incubated to synthesize their respective β-galactosidase fusion proteins by the addition of IPTG (10 μM). Bacterial cells derived from induced and uninduced cultures (3 ml) were centrifuged, and the pellets were resuspended in 100-μl aliquots of solution B (20 mM HEPES, pH 7.9, 100 mM KCl, 0.2 mM EDTA, 20% (v/v) glycerol, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, and 0.5 mg/ml lysozyme). After 15 min, the concentration of NaCl was adjusted to 1 M, and the solutions were mixed by inverting the tubes every 3 min for a total period of 15 min. Cell lysates were centrifuged at 4 °C for 30 min at 30,000 × g, and the supernatants were dialyzed using Millipore filters (type VS, 0.0025 μm) against buffer B (without lysozyme) for 1 h at 4 °C. The dialyzed extracts were frozen in liquid nitrogen and stored at −70 °C. Protein concentrations were determined by the method of Bradford (18).

Western and Southwestern Blot Assays—Crude protein extracts derived from lysogens were separated by 7.5% SDS-PAGE and electrophoretically transferred to nitrocellulose membranes. For immunoblot analysis, the blotting buffer contained 25 mM Tris, 190 mM glycine, 20% methanol, while for the Southwestern assay the conditions were the same as those previously described (11). Blocking, washing, and incubation of the membranes with antibodies were carried out in Tris-buffered saline (20 mM Tris-HCl, pH 7.5, and 150 mM NaCl) containing 5% nonfat dry milk and 0.1% Tween 20. The monoclonal antibody against β-galactosidase (Sigma) was diluted 1:2500 and used as primary antibody. After incubation for 1 h at room temperature, the antibody was further incubated (1 h at room temperature) with the secondary antibody, horseradish peroxidase-linked goat anti-mouse IgG (dilution 1:5000). The immune complexes were visualized by the ECL chemiluminescent detection system (Amersham Corp.) according to the instructions of the manufacturer.

Electrophoretic Mobility Shift Assay (EMSA) —EMSA experiments were carried out essentially as described previously (11). Briefly, the 32P-labeled oligonucleotide probes were incubated with bacterial extracts in a total volume of 20 μl that contained 10 mM Hepes, pH 7.9, 50 mM KCl, 0.1 mM EDTA, 10% (v/v) glycerol, 0.5 mM dithiothreitol, 0.5 mM phenylmethylsulfonyl fluoride, 1 μg of poly(dI-C), and 1 μg of denatured salmon sperm DNA. For competition experiments, the protein fractions were preincubated (10 min on ice) with the appropriate unlabeled oligonucleotide before the addition of the labeled probe. After a 15-min incubation on ice, the binding reactions were analyzed by electrophoresis on a nondenaturing 5% polyacrylamide gel.

Plasmid Constructions—The cDNA inserted in the ACPBP cloned was amplified by PCR, using the forward and reverse primers of agt11 phage, and then digested with EcoRI (CPBP cDNA does not have internal EcoRI sites) and ligated into the Bluescript plasmid (Stratagene) previously digested with the same restriction enzyme. Other enzymes were used to map restriction sites in the CPBP cDNA to obtain different fragments, which were subsequently cloned into pUC18 vector. These subclones contained the following inserts with the respective sizes: 0.6-kb BamHI-PstI (200 bp and 500 bp), and BamHI-PstI (400 bp). The inserts corresponding to these subclones and that obtained in the Bluescript plasmid were sequenced in both directions using the forward and reverse primers for pUC18 or T3 and T7 primers for Bluescript. The sequencing reactions were performed by the chain termination method (19) using denatured double-stranded DNA templates.

The reporter plasmid pSG5-CAT was constructed by inserting the minimal promoter region of pSG5 (positions −254/−43 with respect to the translation start site) upstream from the chloramphenicol acetyltransferase (CAT) gene in the promoterless pBlCAT3 vector as described (20). The (17mer/2-TK-CAT reporter plasmid, which contains two Gal4 DNA binding sites in the thymidine kinase promoter, has been previously described (21). For expression of CPBP in cultured cells, the complete cDNA was subcloned into the EcoRI site of pgy plasmid downstream from the cytomegalovirus promoter and enhancer. A Gal4-CPBP chimeric plasmid was constructed by cloning the complete CPBP cDNA into the pSG424 plasmid (22).

Cell Culture, Transfections, and Cat Assays—COS-7 cells were grown in Dulbecco’s modified Eagle’s medium, supplemented with 5% fetal bovine serum, streptomycin (0.1 mg/ml), and penicillin (100 units/ml). The cells were transfected either using the lipofectamine method following the manufacturer’s recommendations (Life Technologies, Inc.) or by calcium phosphate coprecipitation (23) with the amounts of recombinant DNA indicated in the figure legends, adjusted to a total DNA amount of 16 μg with Bluescript DNA. Cells were harvested after 48 h, and protein extracts were prepared and assayed for CAT activity as described previously (24). Translation of the plasmid-derived cat gene was controlled by β-galactosidase staining of COS-7 cells transfected with the pCH110 expression plasmid (Pharmacia Biotech Inc.) as described (25). To compare the CAT activities, the amounts of cell extracts were normalized to equivalent protein concentrations. Each transfection experiment was repeated at least three times with different effector plasmid concentrations. Percentage of acetylation of chloramphenicol was determined by thin layer chromatography followed by scintillation counting.

Multiple Tissue Northern Blot—The Northern transfer of RNAs (poly(A)+) derived from several human tissues was purchased from Clontech and used according to the instructions of the manufacturer. The DNA template employed in the preparation of CPBP radio-labeled probe was obtained by PCR amplification of a 500-bp fragment from CPBP cDNA. The PCR reaction was performed with an internal primer specific for CPBP cDNA (5'-CGCGATGTCCCTAGAAGGGTTG-CCTGTCCT-3') and the Bluescript reverse primer, using as a template a recombinant Bluescript plasmid carrying the complete CPBP cDNA. As an internal control, the blot was probed with the GAPDH and β-actin cDNAs. The generation of radiolabeled fragments was performed as described previously (26). To quantify the expression of CPBP mRNA in the different detected tissues, the Northern blot was scanned with a Shimadzu dual wavelength chromatoscanner. The obtained CPBP area values were normalized to the cognate GAPDH area values.

Cell-free Transcription and Translation of CPBP cDNA—The Bluescript plasmid carrying the CPBP cDNA was digested with suitable restriction enzymes to obtain the sense and antisense orientation with...
under such conditions, allowing us to isolate two and the free probe (lanes 3) are indicated by arrows (Lane 1, without proteins). Respect to the T7 and T3 phage promoters, respectively. The linearized plasmids were subsequently transcribed in vitro by T7 or T3 RNA polymerases. After transcription, the mRNAs (1–2.5 μg) were used separately for translation reactions in a rabbit reticulocyte lysate, according to the specifications recommended by the manufacturer (Promega). The reactions (25 μl) were performed in the presence of 1.5 μl of [35S]methionine (353 mCi mmol−1). The radilabeled polypeptides were resolved by SDS-PAGE and visualized by fluorography.

RESULTS

Isolation of CPBP cDNA—Since we have previously succeeded in detecting placental DNA-binding proteins able to interact with a core promoter element (11) in solid supports (i.e. nitrocellulose membranes), an expression library containing human plasmid cDNAs was screened according to well-established conditions. Specifically, 5 × 106 clones were processed under such conditions, allowing us to isolate two Δgt11 clones containing inserts of 1.35 kb encoding core promoter element-specific CPBP-encoded protein. The DNA binding specificity of the ΔCPBP-encoded protein was determined in EMSA only with extracts derived from IPTG-induced ΔCPBP lysogen, indicating that it was a product of the lacZ fusion gene (Fig. 1B). The incubation of the induced ΔCPBP extracts with a monoclonal antibody against β-galactosidase, before or after the addition of the CPE probe, produced a supershift of the DNA-protein complex (Fig. 2A, lanes 2 and 3, respectively). This experiment clearly indicates the identity of the protein complex as a β-galactosidase chimeric protein. Even more informative was the competitive EMSA analysis performed with extracts derived from ΔCPBP lysogenic bacteria and the wild type or a mutant version of the core promoter element contained in CPE and CPEmut oligonucleotides, respectively (Fig. 2B). Formation of the DNA-protein complex was completely abolished by the addition of a 100-fold molar excess of unlabeled CPE oligonucleotide (Fig. 2B, lane 3). To test the specificity of the binding, a 100-fold molar excess of the CPEmut and three other different nonrelated oligonucleotides, named NS1, NS2, and NS3, were incubated with the induced bacterial extracts prior to the addition of the labeled CPE probe. The results of the competition experiments are shown in lanes 4–7 (Fig. 2B). In all of the binding reactions with nonrelated competitor oligonucleotides, the formation of the CPE-protein complex remained unaffected, indicating that the chimeric protein possesses sequence specificity. Additionally, when the CPEmut DNA was used as probe (Fig. 2B, lanes 8 and 9) the presence of the DNA-protein complex was no longer detected. Taken together, these results indicate that the interaction of the fusion protein with the CPE motif is highly specific and that an essential sequence within this element, which was substituted in CPEmut (i.e. 5′-ACC CAT-3′ → 5′-GATATC-3′), is crucial for DNA recognition.

As mentioned before, we established that ΔCPBP1 and ΔCPBP2 contained identical inserts; thus, we continued our experiments with only one cDNA (CPBP1). This clone was PCR-amplified and subcloned in the Bluescript plasmid to perform further manipulations. Several subclones were obtained, and the complete sequence (1350 bp) was accomplished in both strands of each insert. The nucleotide sequences of CPBP cDNA and its deduced protein are depicted in Fig. 3. On the basis of sequence analyses and data base searches, we con-
cluded that the CPBP cDNA and its conceptual polypeptide have not been reported previously; however, it has only partial homologies with some transcription factors (see below).

Structural Features of CPBP—The CPBP cDNA has an open reading frame of 290 amino acids, which in turn constitutes a polypeptide with a calculated molecular mass of 33 kDa and an isoelectric point of 9.404. In contrast, the presence of the AUG codon in the farthest 5' position matches almost completely (8 of 9 nucleotides) with the consensus sequence (cc(g/a)ccAUGg) proposed by Kozak (27, 28) to be an essential element in the scanning model of translation. This initiation codon may, therefore, allow for the synthesis of a protein with a primary structure of 283 amino acids. The molecular mass of the fusion β-galactosidase-CPBP protein was determined by SDS-PAGE to be approximately 150 kDa. This chimeric protein was identified with two different methodologies, taking advantage of the β-galactosidase portion and the DNA binding activity of CPBP (Fig. 4, A and B, respectively). In Western blots, the induced fusion protein was revealed with a monoclonal antibody against β-galactosidase and migrated with a relative mass of 150 kDa (Fig. 4 A). Accordingly, the Southwestern blotting analysis indicated that the same band accounts for the DNA-binding activity when probed with the CPE oligonucleotide (Fig. 4 B). These results confirmed the identity of the fusion protein and also allowed us to deduce the molecular mass of the CPBP fraction by subtracting the β-galactosidase portion (116 kDa) from the 150 kDa corresponding to the chimeric protein; thus, the CPBP fraction is responsible for approximately 34 kDa of the total molecular mass. To confirm that the predicted CPBP protein can be synthesized in a eukaryotic system, we performed in vitro translation of the CPBP mRNA (1.35 kb). To this end, the mRNA encoding the CPBP protein or the antisense was transcribed in vitro, and the cognate protein was translated in a rabbit reticulocyte lysate system. The synthesized polypeptide had an apparent molecular mass of 32 kDa in SDS-PAGE (Fig. 4 C), which is in good agreement with the calculated mass deduced from its primary structure and from that determined by Western and Southwestern blot assays. The shorter polypeptides detected in the presence of the sense mRNA may be due to incomplete synthesis of the CPBP protein (Fig. 4 C). In conclusion, the CPBP mRNA has the potential to be translated into a polypeptide of 32 kDa in a eukaryotic system.

As a first step toward the identification of characteristic domains for DNA-binding proteins, several sequence alignments were performed with other known transcription factors. These analyses allowed us to determine that CPBP has three contiguous zinc fingers (type Cys2-His2) at the end of its C-terminal portion (Fig. 5, A and B). The cysteine and histidine residues as well as other conserved amino acids are present in the zinc finger structures of CPBP (Fig. 5 A). Additionally, the variations of key residues in different arrays of zinc fingers determine different affinities and specificities that are displayed by such DNA binding domains (29). The target sequences recognized by several related zinc-finger proteins (e.g., EKLF and BTEB) share strong similarity with each other and have been proposed to be a guanine-rich binding site (30, 31,
lysates from wild type agt11 (wt) and ACPBP lysogens, as indicated, were analyzed by Western blot (A) and Southwestern blot (B). The bacterial extracts were resolved by SDS-PAGE and electrotransferred to nitrocellulose. The immobilized proteins were incubated with a monoclonal antibody against β-galactosidase, and the bands were visualized by incubation of the nitrocellulose membrane with the CPE-labeled probe and autoradiography. The arrows indicate the migration and the estimated molecular mass markers in kDa.

In this regard, the sequence recognized by CPBP fulfills these criteria and can also be included in this subset of zinc finger proteins.

These findings can be rationalized in the light of other reports. For instance, Klevitt (34) proposed that certain basic residues in the finger are responsible for the contacts with guanine nucleotides present within triplets, which in turn constitute the binding site. These amino acids are termed X, Y, and Z and are depicted in Fig. 5C. Therefore, assuming that Klevitt’s predictions apply for CPBP contacts with DNA, the target site can be deduced to be 3′-GGN GNG GGN-5′. In fact, this sequence is in perfect correspondence with the natural context detected by previous in vitro approaches (11) (Fig. 5C). These data, along with the results accomplished by competitive EMSA using a mutant CPE, are compelling evidence that the zinc fingers of CPBP are responsible for the specific DNA binding with the guanine-rich CPE.

Other interesting features can also be distinguished from the amino acid composition analysis of CPBP. For instance, the central region of the CPBP polypeptide, located between positions 114 and 208, is serine/threonine-rich and could be a potential target for some post-translational modifications like phosphorylation (Fig. 5B). In addition, the high content of acidic residues found from amino acid 19 to 112 conferred a predominant negative interphase (net charge = −11) to the CPBP zinc finger domain shown at the bottom (N indicates any of the four deoxynucleotides).

Fig. 4. The CPBP cDNA is translated into a polypeptide of 32 kDa. Noninduced (lanes 1 and 3) or IPTG-induced (lanes 2 and 4) crude lysates from wild type agt11 (wt) and ACPBP lysogens, as indicated, were analyzed by Western blot (A) and Southwestern blot (B). A, the bacterial extracts were resolved by SDS-PAGE and electrotransferred to nitrocellulose. The immobilized proteins were incubated with a monoclonal antibody against β-galactosidase, and the bands were visualized by incubation of the nitrocellulose membrane with the CPE-labeled probe and autoradiography. The arrows indicate the migration and the estimated molecular mass markers in kDa.

Fig. 5. Structural analysis of the CPBP polypeptide. A, CPBP contains a zinc finger domain of the Cys2-His2 type. Alignment of the putative zinc finger sequences from CPBP and BTEB or EKLF proteins. x, xenopus; h, human; r, rat; m, mouse. Identical regions (one-letter code) are represented by white letters boxed in black. The conserved cysteine and histidine residues are shown in black and gray boxes, respectively. Letters outside boxes correspond to the nonconsensus regions. Hyphens correspond to gaps introduced to optimize the alignment. B, the polypeptide corresponding to the CPBP open reading frame is schematically represented. The gray boxes depict the amino acids involved in the Cys2-His2-type zinc finger structure. The black box and the white box containing minus symbols or S (serine) and T (threonine) represent the predicted acidic and serine/threonine-rich domains, respectively. The bar with numbers indicating amino acids positions for each domain is shown at the top of the scheme. C, sequence comparison between the natural and predicted DNA-binding site for CPBP zinc finger domain. The peptide sequence (one-letter code) of the three putative CPBP zinc finger domains numbered from 1 to 3 are shown in white letters boxed in black. The X, Y, and Z represent the key coordinates used for the prediction of DNA-binding sites of known zinc finger domains according to Klevitt (34) (note above the amino acid sequence). The sequence of the deduced and natural DNA-binding sites for the CPBP zinc finger domain are shown at the bottom (N indicates any of the four deoxynucleotides).
A Novel Core Promoter-binding Protein, CPBP

**A**

**Effectors**

|          | CMV+E | CPBP-DNA |
|----------|-------|----------|
| Reporters| CPE   | CAT      |
|          | PSG5prom | (17mer)TK-CAT |

**B**

**Fig. 6. The CPBP protein stimulates transcription.** A, the structure of the plasmids expressing the CPBP open reading frame alone or fused to the region encoding the Gal4 DNA binding domain (Effectors) and the structure of the reporter plasmids. PSG5-CAT and (17mer/2-TK-CAT, expressing the chloramphenicol acetyltransferase gene driven by the PSG5 and thymidine kinase promoters are schematically represented. CMV+ E and SV40 refer to the cytomegalovirus promoter and enhancer and the SV40 promoter present in the pβ and pβG424 expression plasmids, respectively. The Gal4 DNA binding domain (Gal4DBD) is shown as a black box. The CPE box and the two Gal4 binding sites, (17mer/2) present in the PSG5 and thymidine kinase promoters respectively, are shown as shaded boxes. B, using calcium phosphate coprecipitation (lanes 1–4) or lipofectamine (lanes 5–8) methods, COS-7 cells were cotransfected with the PSG5-CAT (5 μg, lanes 1–4) or (17mer/2-TK-CAT (0.5 μg, lanes 5–8) reporter plasmids and different amounts of the effector plasmids, as indicated. In lanes 1–4, the total DNA amount of the effector plasmid was supplemented up to 5 μg with the empty pCMV+E plasmid. After 48 h, the cells were harvested, and CAT activities were determined. For each set of transfections, the relative CAT activity values were calculated as the percentage of chloramphenicol acetylation of each assay divided by the percentage of acetylation of the cognate control assay (lanes 1 and 5). The chromatograms of representative CAT assays for each reporter/effector transfection experiments are shown at the bottom (lanes 1–4 and lanes 5–8). The CAT activity represents the average of values from three independent experiments and is shown as indicated.

**DISCUSSION**

We have previously reported the characterization of an essential element of the PSG5 gene promoter that contributes to its basal activity in different cell types. Furthermore, we have demonstrated that this element is recognized by distinct proteins (11). In an attempt to identify these proteins, we have performed a target site screening of a placental expression library, and a cDNA encoding a polypeptide with specific CPE binding activity, designated CPBP, was isolated. DNA sequencing and subsequent data base searches indicated that the CPBP cDNA encodes a previously undiscovered protein that bears a three-zinc finger (Cys2-His2) motif at its C-terminal domain.

We have also determined that a fusion protein constituted by N-terminal domain (Fig. 5B). The possible role of such structures will be interpreted under “Discussion.” Finally, the presence of potential phosphorylation sites, notably in serine and/or threonine residues, which in turn constitute the 22.6% of the CPBP polypeptide, most probably indicates a post-translational regulation of CPBP protein.

Functional Analysis of CPBP—To determine the impact of CPBP on transcription, its cDNA was cotransfected in COS-7 cells with a reporter plasmid carrying the CAT gene driven by the PSG5 promoter sequence (positions −254/−43) (20). This reporter vector contains the CPE sequence located between positions −150 and −124, overlapping one cluster (−130/−134) of the two transcription start sites described for the PSG5 gene. Subsequently, CAT activity was measured, and the results obtained are shown in Fig. 6A. Increasing amounts of effector plasmid in the presence of adjusted quantities of reporter DNA resulted in a clear and dose-dependent stimulation of CAT activity (Fig. 6B, lanes 1–4). To further extend our results, we constructed a chimeric version of CPBP fused with the DNA binding domain of Gal4 and tested the activity of this chimera on a promoter containing two Gal4 binding sites (Fig. 6B, lanes 5–8). Taken together, these approaches enable us to unequivocally indicate that CPBP is capable of activating transcription approximately 4-fold either on homologous or heterologous promoters.

Expression of CPBP Transcript in Human Tissues—Hybridization of a radiolabeled 5′ fragment (outside the zinc fingers) of the CPBP cDNA to a Northern transfer of poly(A) RNAs from different human tissues allowed us to determine that the CPBP transcript is differentially expressed as unique species of 4.5 kb (Fig. 7). Considering the normalized mRNA levels, the CPBP transcript appears to be enriched in placental cells (Fig. 7, lane 3). In contrast, the transcript level in other tissues (i.e. pancreas, lung, liver, heart, and skeletal muscle) was present in decreasing amounts or was undetectable (kidney and brain). In sum, the in vivo mRNA levels observed for CPBP in different human organs indicate that transcriptional regulation plays a crucial role in CPBP expression. It is important to remark that essentially the same results were obtained when the values for CPBP mRNA expression were normalized using the β-actin mRNA instead of GAPDH (not shown).
CPBP and β-galactosidase exerted its DNA binding activity by recognizing a particular GC-rich sequence of the CPE element. It is worth mentioning that although the competitive EMSAs were performed with a fusion protein, essentially the same results were achieved either with a bacterially synthesized CPBP or a CPBP polypeptide translated in a reticulocyte lysate.3

So far, we have sequenced 1.35 kb of the CPBP cDNA, and when it was used as a probe in a Northern blot analysis a single 4.5-kb band was identified (Fig. 7). At present, we are using distinct strategies to isolate the full-length cDNA to decipher the complete primary structure of CPBP. Nevertheless, a consensus sequence for translational initiation encompasses the first AUG codon at the 5′ end of CPBP cDNA, although the reading frame remains open to its upstream side (Fig. 3). It seems likely that this AUG codon serves as the regular translation start site, since it is recognized at high efficiency in a reticulocyte lysate to produce a polypeptide that migrates in SDS-PAGE as a band of 32 kDa. These data are in good agreement with the calculated molecular mass of the CPBP protein as deduced from its primary structure. However, the coding potential for CPBP mRNA is not limited to the 32-kDa protein as mentioned above. In a previous report, we determined that at least two placental proteins with molecular masses of 78 and 53 kDa were involved in the formation of specific DNA-protein complexes (11). In this work, we found that CPBP has the same sequence requirements for DNA binding as the complexes detected with protein extracts from culture cells and has the ability to stimulate the activity of both homologous and heterologous promoters about 4-fold. Considering these data, it would be essential to identify the in vivo synthesized CPBP to correlate its molecular mass to those previously described (Ref. 11 and this work). To this end, we are currently working to produce antibodies against specific CPBP peptides.

Regarding the in vivo expression patterns of CPBP mRNA, we found that its level of transcription varies in different human tissues, as determined by Northern blot analysis. The highest mRNA level was detected in placenta, which is the organ where PSG genes are predominantly expressed. Thus, one might speculate that the CPBP factor plays a pivotal role in the regulation of PSG expression in placental cells. In contrast, the CPBP mRNA level in other organs was moderate or undetectable. These observations are consistent with the idea that CPBP expression is differentially regulated at the transcriptional level and that CPBP actions may lead to a tissue-specific regulation of its target gene(s).

Having identified a novel DNA-binding protein that interacts with fundamental promoter elements, we aimed at searching for the potential transcriptional activity of CPBP. In cotransfection experiments, the functional properties of CPBP were determined, indicating that this novel protein increases transcription levels. We also asked whether the activating functions of CPBP are specific for the PSG5 promoter or can be elicited in different promoter contexts. As a first step in this direction, a chimera consisting of CPBP and the DNA binding domain of Gal4 was constructed and cotransfected with a reporter containing Gal4 binding sites. On the basis of the data shown here, CPBP is able to stimulate transcription either in its natural context or in a heterologous promoter context.

The PSG5 promoter contains neither TATA box nor typical pyrimidine-rich initiator motifs; however, one core promoter element (CPE box) overlaps the farthest transcription start site and is essential for DNA-protein interactions and promoter activity in different cell types (11). Interestingly, the CPE box spans a region positioned between nucleotides −24 and −50 bp upstream from the major transcription start site (+1), thus resembling the TATA box location (−30/−35 bp) found in most eukaryotic genes. This observation implies that CPBP most probably acts as a tethering factor that serves to recruit and/or maintain the transcription machinery within the boundaries of the promoter and allows for the accurate transcriptional initiation from this particular TATA-less promoter. Moreover, the CPE box is crucial to sustain roughly similar levels of PSG5 transcriptional activity in different cell types, independently of whether the cells do or do not express PSG genes, suggesting that this sequence and its cognate CPBP are mainly involved in basal promoter activity. These observations are not in disagreement with the CPBP dependent transcription enhancement observed, since CPBP may behave as a rate-limiting factor for basal activity, whereas overexpression of this factor may overcome the rate-limiting step, and consequently an activated transcription level is sustained.

A comparison at the amino acid level of the CPBP sequence with its most closely related transcription factors (i.e. EKLF and BTEB) revealed a high degree of conservation of the zinc finger regions (Fig. 5A), but no other significant similarities have been found in the rest of the molecule. The regulatory proteins that share homology between the zinc finger domains, most specifically in the residues contacting the DNA (positions X, Y, and Z; Fig. 5A), may act as potential regulators of transcription by interacting with their cognate sequence as well as with other similar motifs. For instance, transcription factors such as EKLF and BTEB, whose transcriptional activities and DNA-binding specificities have been well established (29–31),2 may contact the CPE motif present in the PSG5 promoter via their zinc fingers. Alternatively, other promoters that are targets for the aforementioned regulatory proteins may also be regulated by CPBP, which in turn contributes to a higher diversity in promoter recognition.

Interestingly, the central and N-terminal portions of the CPBP protein are well distinguished between each other and bear some attractive features. In this regard the central region

---

3 N. P. Koritschoner and L. C. Patrito, unpublished results.
A Novel Core Promoter-binding Protein, CPBP

of CPBP (amino acids 114–208) is a serine/threonine-rich domain (28.7%) that might be involved in activation or post-translational regulatory pathways. In contrast, the N-terminal domain (residues 19–112) contains acidic amino acids conforming to a predominant negative interface with a net charge of −11. These acidic residues are interspersed among numerous hydrophobic amino acids (40% of the domain) of which approximately 50% correspond to leucine and isoleucine. Such a type of structure has been proposed to play an important role in the process of transcriptional activation, where the interactions of the activation domain with its target protein might be mediated by hydrophobic forces (3). Additionally, the basic domains present in some general transcription factors (e.g. TFIIB, TATA box-binding protein) have been postulated to be the targets of acidic activators (32, 35). In conjunction, all of these features delineate a modular structure for CPBP, which might function as a tethering factor that binds TATA box-less promoters or an activator itself by mediating interactions, e.g. via its acidic domain, with one (or several) general transcription factors. Further investigations of the structure-function relationship of CPBP as well as full elucidation of its role in transcription will shed more light on the molecular mechanisms that are regulated by this novel protein.

Acknowledgments—We are grateful to Dr. Susana Genti-Raimondi and Dr. Daniel Perez for many helpful discussions.

REFERENCES

1. Buratowsky, S. (1994) Cell 77, 1–3
2. Goodrich, J. A., Cutler, G., and Tjian, R. (1996) Cell 84, 825–830
3. Tjian, R., and Maniatis, T. (1994) Cell 77, 5–8
4. Hernandez, N. (1993) Genes & Dev. 7, 1291–1308
5. Roy, A. L., Meisterernst, M., Pognonec, P., and Roeder R. G. (1991) Nature 354, 245–248
6. Shi, Y., Seto, E., Chang, L.-S., and Shenk, T. (1991) Cell 67, 377–388
7. Roeder, R. G. (1991) Trends Biochem. Sci. 16, 402–408
8. Hauck, W., and Stanners, C. P. (1995) J. Biol. Chem. 270, 3602–3610
9. Hauck, W., Nedellec, P., Turcotte, C., Stanners, C. P., Barnett, T. R., and Beauchemin, N. (1994) Eur. J. Biochem. 223, 529–541
10. Keck, U., Nedellec, P., Beauchemin, N., Thompson, J., and Zimmermann, W. (1995) Eur. J. Biochem. 229, 455–464
11. Kiritschonner, N. P., Ponzetta-Dutari, G. M., Bocca, J. L., Dumur, C. I., Flury, A., and Patrito, L. C. (1996) Eur. J. Biochem. 236, 365–372
12. Chamberlin, M. E., Lei, K.-J., and Chou, J. Y. (1994) J. Biol. Chem. 269, 17152–17159
13. Singh, H., LeBowitz, J. H., Baldwin, A. S., and Sharp, P. A. (1988) Cell 52, 415–423
14. Millian, J. L. (1986) J. Biol. Chem. 261, 3112–3115
15. Young, R. A., and Davis, R. W. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 1194–1198
16. Hoeffer, J. P., Meyer, T. E., Yun, Y., Jameson, J. L., and Habenerm, J. F. (1988) Science 242, 1430–1433
17. Glover, D. M. (1985) DNA Cloning: A Practical Approach, Vol. I, pp. 76–77, IRL Press Ltd., Oxford
18. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
19. Sanger, F., Nicklen, N., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
20. Ponzetta-Dutari, G. M., Bocca, J. L., Reimund, B., Flury, A., and Patrito L. C. (1992) Mol. Biol. Rep. 16, 255–262
21. Webster, N., Jin, J. R., Green, S., Hollis, M., and Chambon P. (1988) Cell 52, 169–178
22. Sadowski, I., and Pruschnke, M. (1989) Nucleic Acids Res. 17, 17539–17539
23. Bamberji, J. S., Rusconi, S., and Schaffner, W. (1981) Cell 27, 299–308
24. Gorman, C. M., Moffat, L. F., and Howard, B. H. (1982) Mol. Cell. Biol. 2, 1034–1051
25. Hall, C. V., Jacob, P. E., Ringold, G. M., and Lee, P. (1983) J. Mol. Appl. Genet. 2, 101–109
26. Feinberg, A. P., and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13
27. Kazaz, M. (1987) J. Mol. Biol. 196, 947–950
28. Kazaz, M. (1989) J. Cell Biol. 108, 229–241
29. Berg, J. M., and Shi, Y. (1996) Science 271, 1081–1085
30. Miller, I. J., and Bieler, J. V. (1993) Mol. Cell. Biol. 13, 2776–2786
31. Imataka, H., Sogawa, K., Yasumoto, K., Kikuchi, Y., Sasano, K., Kobayashi, A., Hayami, M., and Fujii, K. Y. (1992) EMBO J. 11, 3663–3671
32. Hisatake, K., Roeder, R. G. and Horikoshi, M. (1993) Nature 363, 744–747
33. Sogawa, K., Imataka, H., Yamasaki, Y., Kusume, H., Abe, H., and Fujii, K. Y. (1990) Nucleic Acids Res. 21, 1527–1532
34. Klevit, R. E. (1991) Science 253, 1367
35. Malik, S., Hisatake, K., Sumimoto, H., Horikoshi, M., and Roeder, R. G. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 9553–9557
A Novel Human Zinc Finger Protein That Interacts with the Core Promoter Element of a TATA Box-less Gene
Nicolás P. Koritschoner, José L. Bocco, Graciela M. Panzetta-Dutari, Catherine I. Dumur, Alfredo Flury and Luis C. Patrito

*J. Biol. Chem.* 1997, 272:9573-9580. doi: 10.1074/jbc.272.14.9573

Access the most updated version of this article at [http://www.jbc.org/content/272/14/9573](http://www.jbc.org/content/272/14/9573)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 35 references, 13 of which can be accessed free at [http://www.jbc.org/content/272/14/9573.full.html#ref-list-1](http://www.jbc.org/content/272/14/9573.full.html#ref-list-1)