Nucleic Acids Research, 1993, Vol. 21, No. 7 ***1527–1532***

cDNA cloning and transcriptional properties of a novel GC box-binding protein, BTEB2

Kazuhiro Sogawa, Hiroaki Imataka[1], Yuichi Yamasaki, Hiroshi Kusume, Hisaku Abe and Yoshiaki Fujii-Kuriyama

Department of Chemistry, Faculty of Science, Tohoku University, Sendai 980 and [1]Institute for Virus Research, Kyoto University, Kyoto 606, Japan

Received February 5, 1993; Revised and Accepted March 8, 1993 DDBJ accession no. D14520

ABSTRACT

We have cloned a cDNA for a novel GC box-binding protein designated BTEB2 from a human placenta cDNA library using rat BTEB cDNA (Imataka et al. (1992). EMBO J. 11,3663–3671. as a hybridization probe. BTEB2 consists of 219 amino acids and contains three contiguous zinc finger motifs at its C-terminus. The zinc finger domains showed 59% and 64% sequence similarity to those of Sp1 and BTEB, respectively. Adjacent to the N-terminal of the zinc finger motifs, a short sequence rich in basic amino acids is conserved between BTEB2 and Sp1. Furthermore, This basic sequence concurs with the N-terminal half of the consensus sequence for basic domains of the proteins containing both helix-loop-helix and leucine zipper motifs. The other region of BTEB2 is notably rich in proline, serine, threonine, and alanine residues. BTEB2 expressed in *Escherichia coli* showed DNA-binding activity whose specificity was closely similar to that of Sp1. Cotransfection experiments using Hepa-1 cells (a mouse hepatoma cell line) with a BTEB2 expression plasmid and GC box-containing reporter plasmids revealed that BTEB2 apparently activated the expression of the CAT activity. Moreover, when BTEB2 was fused to GAL4 DNA-binding domain, the chimeric protein could enhance the transcription through promoters containing GAL4-binding sites. Analysis of the BTEB2 mRNA by RNA blot analysis demonstrated that the mRNA was expressed specifically in testis and placenta with different sizes, 20S and 28S, respectively, among various organs examined.

INTRODUCTION

The regulation of eukaryotic transcription is governed by the combined action of various sequence-specific DNA-binding factors. Extensive studies on the DNA-binding properties of these factors revealed that most of target sequences are recognized by multiple DNA-binding proteins acting positively or negatively on transcription. Presence of a set of factors which interact with common binding sites seemed favorable for the fine control of transcription of eukaryotic genes through a limited number of cis-acting elements located in their promoter region.

GC box sequence is one of the most widely distributed promoter elements in cellular and viral genes (1, 2). Tjian and his colleagues found Sp1 as the GC box-binding factor, and established its role on transcription by elucidating its molecular properties (3, 4). Sp1 contains three contiguous zinc finger motifs and glutamine-rich domains as the DNA-binding and transcriptional activation domains, respectively. Although Sp1 has long been thought to be a unique GC box-binding protein, we have recently cloned a cDNA for a GC box-binding protein designated BTEB (BTE-binding protein) (5). BTEB has three-time repeated zinc finger motifs like Sp1. Investigation of transcriptional activity of BTEB indicated that BTEB stimulates promoters with repeated GC box sequences, while promoters with a single GC box sequence are not activated by BTEB (5). Finding of BTEB encouraged us to explore further the presence of GC-box binding factors other than Sp1 and BTEB.

Here we report the isolation of human cDNA encoding a novel GC box binding protein designated BTEB2. BTEB2 consists of 219 amino acids and contains three contiguous zinc finger motifs at its C-terminus. Transient expression of BTEB2 in eukaryotic cells activates the expression of the CAT (chloramphenicol acetyltransferase) activity from a reporter gene consisting of a GC-box promoter and CAT structural gene. Transcriptional activation by BTEB2 was further confirmed by a fusion protein of BTEB2 and GAL4 DNA binding segment which enhances the transcription in a GAL4 binding site-dependent manner.

MATERIALS AND METHODS

Cloning of cDNAs

cDNA clones for BTEB2 were isolated from a human placenta cDNA library in λgt11 by the method of Benton and Davis (6) using BTEB cDNA, λBP26 (5), as a probe. Plaque hybridization was carried out with the nick-translated cDNA (specific activity, 8×10^7cpm/μg) as the probe at 65°C in 20mM Tris buffer, pH7.5, containing 1M NaCl, 10mM EDTA, 0.1% sodium N-lauroylsarcosinate, 0.2% polyvinylpyrrolidone, 0.2% Ficoll, 0.2% bovine serum albumin, and 100μg/ml salmon sperm DNA.

The filters were washed twice at 65°C for 30min in 2×SSC (1×SSC, 0.15M NaCl and 0.015M sodium citrate) and 0.1% sodium N-lauroylsarcosinate. Positive clones were plaque-purified four times and inserted DNAs were subcloned into pUC18 plasmid vector at the EcoRI site.

DNA sequence analysis

The cDNA sequence was determined by the dideoxynucleotide method using M13mp18 and M13mp19 as cloning vectors (7). Some parts of the sequence were determined by the chemical method of Maxam and Gilbert (8).

Construction of plasmids

For the expression of BTEB2 in *E.coli*, 832bp BglII/BamHI fragment (The BamHI site is located at the polylinker site of the plasmid vector) was cloned into the BamHI site of pAR2106 (9) which has the T7 promoter to generate pT7BTEB2. For the expression of BTEB2 in cultured cells, 775bp Xba fragment (The downstream XbaI site is located at the polylinker site of the plasmid vector) was cloned into the Xba I site of pCMSV (10) which has the cytomegalovirus enhancer and promoter. An Xho fragment which covers the entire protein-coding sequence of Sp1 was isolated from pRSVSp1 (5) and cloned into the Xho I site of pCMSV. A GAL4-BTEB2 chimeric plasmid was constructed by inserting 462bp XbaI/DraI fragmant of BTEB2 at the BamHI site in pSG424(11) after filled in with *E.coli* DNA polymerase I, Klenow fragment.

Gel mobility shift assay

E.coli strain BL21 was transformed with pT7BTEB2, and the transformant was grown in L-broth. Induction with 0.4mM IPTG (isopropyl-β-D(−)-thiogalactopyranoside) was performed when the culture reached an A_{600} of 0.1. Three hours after induction, extracts were prepared by treatment with lysozyme (1mg/ml) in 20mM Tris buffer, pH7.5, containing 20mM NaCl at 4°C for 30 min followed by sonication. After centrifugation, the supernatant was used for gel mobility shift assay. Hepa-1 cells from a mouse hepatoma (12) were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and 0.35% glucose. Nuclear extracts from Hepa-1 cells were prepared by the method of Dignam et al. (13). Gel mobility shift assays were performed as described previously (14).

DNA transfection and CAT assay

DNA transfection was performed by the calcium phosphate-DNA coprecipitation method as described (15). CAT assays were carried out by the method of Gorman et al. (16).

RNA blot analysis

Total RNA was isolated using the guanidine thiocyanate-CsCl procedure (17). RNA (15μg) was electrophoresed on 0.8% agarose/formaldehyde gel and blotted onto nitrocellulose filters. The filters were hybridized at 42°C overnight with ^{32}P-labeled random-primed BTEB2 cDNA (896bp EcoRI/HpaI fragment, specific activity 1.0×10^8cpm/μg) as described (18), and then washed twice at 50°C for 30min in 2×SSC or 0.1×SSC.

RESULTS

Cloning and structure of BTEB2 cDNA

We screened approximately 5×10^5 plaques from the human placenta cDNA library using rat BTEB cDNA as a probe, and obtained three positive clones. Two were for human BTEB cDNA

A

```
                                                                GGGCACGCGCACCACCGC   18
CCGCAGCGCAGCCCGGCGCCGCGCAGGCCCCGCAGCCGGCCCAGCCCGCCGCCACCGGCCGCGGCTGCCTCCAGAGGACCTGGTCCAGAC  108
AAGATGTGAAATGGAGAAGTATCTGACACCTCAGCTTCCTCCAGTTCCTATAATTCCAGAGCATAAAAAGTATAGACGAGACAGTGCCTC  198
AGTCGTAGACCAGTTCTTCACTGACACTGAAGGGTTACCTTACAGTATCAACATGAACGTCTTCCTCCCTGACATCACTCACCTGAGAAC  288
TGGCCTCTACAAATCCCAGAGACCGTGCGTAACACACATCAAGACAGAACCTGTTGCCATTTTCAGCCACCAGAGTGAAACGACTGCCCC  378
TCTCCGGCCCCGACCCAGGCCCTCCCTGAGTTCACCAGTATATTCAGCTCACACCAGACCGCAGCTCCAGAGGTGAACAATATTTTCATC  468
AAACAAGAACTTCCTACACCAGATCTTCATCTTTCTGTCCCTACCCAGCAGGGCCACCTGTACCAGCTACTGAATACACCGGATCTAGAT  558
ATGCCCAGTTCTACAAATCAGACAGCAGCAATGGACACTCTTAATGTTTCTATGTCAGCTGCCATGGCAGGCCTTAACACACACACCTCT  648
MetProSerSerThrAsnGlnThrAlaAlaMetAspThrLeuAsnValSerMetSerAlaAlaMetAlaGlyLeuAsnThrHisThrSer   30
GCTGTTCCGCAGACTGCAGTGAAACAATTCCAGGGCATGCCCCCTTGCACATACACAATGCCAAGTCAGTTTCTTCCACAACAGGCCACT  738
AlaValProGlnThrAlaValLysGlnPheGlnGlyMetProProCysHisTyrThrMetProSerGlnPheLeuProGlnGlnAlaThr   60
TACTTTCCCCCGTCACCACCAAGCTCAGAGCCTGGAAGTCCAGATAGACAAGCAGAGATGCTCCAGAATTTAACCCCACCTCCATCCTAT  828
TyrPheProProSerProProSerSerGluProGlySerProAspArgGlnAlaGluMetLeuGlnAsnLeuThrProProProSerTyr   90
GCTGCTACAATTGCTTCTAAACTGGCAATTCACAATCCAAATTTACCCACCACCCTGCCAGTTAACTCACAAAACATCCAACCTGTCAGA  918
AlaAlaThrIleAlaSerLysLeuAlaIleHisAsnProAsnLeuProThrThrLeuProValAsnSerGlnAsnIleGlnProValArg  120
TACAATAGAAGGAGTAACCCCGATTTGGAGAAACGACGCATCCACTACTGCGATTACCCTGGTTGCACAAAAGTTTATACCAAGTCTTCT 1008
TyrAsnArgArgSerAsnProAspLeuGluLysArgArgIleHisTyrCysAspTyrProGlyCysThrLysValTyrThrLysSerSer  150
CATTTAAAAGCTCACCTGAGGACTCACACTGGTGAAAAGCCATACAAGTGTACCTGGGAAGGCTGCGACTGGAGGTTCGCGCGATCGGAT 1098
HisLeuLysAlaHisLeuArgThrHisThrGlyGluLysProTyrLysCysThrTrpGluGlyCysAspTrpArgPheAlaArgSerAsp  180
GAGCTGACCCGCCACTACCGGAAGCACACAGGCGCCAAGCCCTTCCAGTGCGGGGGTGTGCAACCGCAGCTTCTCGCGCTCTGACCACCTG 1188
GluLeuThrArgHisTyrArgLysHisThrGlyAlaLysProPheGlnCysGlyValCysAsnArgSerPheSerArgSerAspHisLeu  210
GCCCTGCATATGAAGAGGCACCAGAACTGAGCACTGCCCGTGTGACCCGTTCCAGGTCCCCTGGGCTCCCTCAAATGACAGACCTAACTA 1278
AlaLeuHisMetLysArgHisGlnAsnStop                                                              219
TTCCTGTGTAAAAACAACAACCC                                                                     1301
```

B

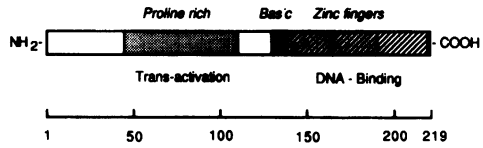

Figure 1. (A) Nucleotide and deduced amino acid sequences of human BTEB2 cDNA. The amino acid sequence is numbered from the presumed initiator methionine. The zinc finger motifs and a short basic sequence were underlined. (B) Schematic representation of structure of human BTEB2, The proline-rich region and the basic region are indicated by shaded and solid boxes, respectively. Three zinc finger motifs are shown by striped boxes.

A

```
BTEB2  CDVPGCDKVYLKSSHLKAHLRTHTGEKPYK
BTEB   CPVSGCGKVYGKSSHLKAHVRVHTGERPFP
Sp1    CHIQGCGKVYGKTSHLRAHLRWHTGERPFM

       CTWEGCDWRFARSDELTRHYRKHTGHKPFQ
       CTWPDCLRKFSRSDELTRHYRTHTGEKQFR
       CNWSYCGKRFTRSDELDRHFRTHTGEKKFA

       CGV--CNRSFSRSDHLALHMKRH
       CPL--CEKRFMRSDHLTKHARRH
       CPE--CPKRFMRSDHLSKHIKTH
```

B

```
BTEB2        SNPDLEKRRIHYC---Zinc Finger
Sp1          GSGDPGKKKQHIC---Zinc Finger
c-Myc        NTEENVKRRTHNVLERQRRNE
TFEB         LAKERQKKDNHNLIERRRRFN

                E  KK  H        R
consensus   ***-**—**-N**ER*R-**
                D  RR  N        N
```

Figure 2. (A) Comparison of amino acid sequences of zinc finger motifs among BTEB2, BTEB, and Sp1. Asterisks indicate cysteine and histidine residues that possibly coordinate zinc. Dots indicate amino acids which are considered to be important for base recognition. Gaps in the third finger motif are introduced to align the conserved cysteine and histidine residues among the three zinc finger motifs. The amino acids conserved among BTEB2, BTEB (5), and Sp1 (4) are enclosed in boxes. (B) Alignments of basic sequences for BTEB2, Sp1, and two HLH-LZ proteins. Amino acids of basic sequences in BTEB2, Sp1, TFEB (29), and c-Myc (34) are aligned, and the consensus sequence for basic domain of HLH-LZ proteins is shown (29).

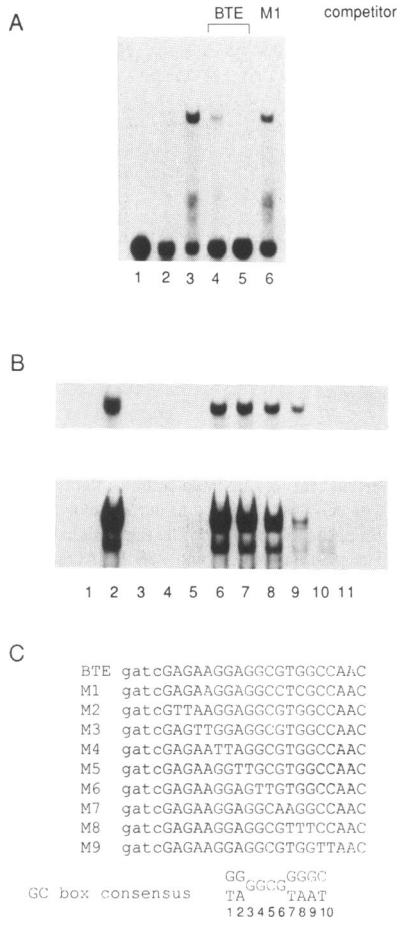

Figure 3. Binding specificity of BTEB2 expressed in *E.coli*. **(A)** Extracts were prepared from *E.coli* treated with or without IPTG and used for gel mobility shift assay. ^{32}P-labeled BTE was used as a labeled probe. Lane 1, no extracts: lane 2, uninduced extracts; lanes 3−6, induced extracts with no competitor (lane 3), 0.1μg (lane 4), 0.3μg (lane 5) of unlabeled BTE, and 0.3μg unlabeled M1 (lane 6) were used for gel mobility shift assay. **(B)** Effect of a series of mutations in the BTE sequence on the binding activity of BTEB2 to BTE. ^{32}P-labeled BTE was used as a labeled probe. IPTG-induced *E.coli* extracts (upper panel) or Hepa-1 nuclear extracts (lower panel) were used for the binding assay. Unlabeled BTE and its mutants were used as competitors (0.5μg for *E.coli* extracts or 0.04μg for Hepa-1 nuclear extracts). lane 1, no extracts; lane 2, no competitors; lane 3, BTE; lane 4−11, M2−M9. **(C)** Nucleotide sequence of BTE and its 9 mutants, M1−M9. Four nucleotides denoted with lowercase letters were introduced for ^{32}P-end-labeling (35). The consensus sequence of GC boxes is shown below.

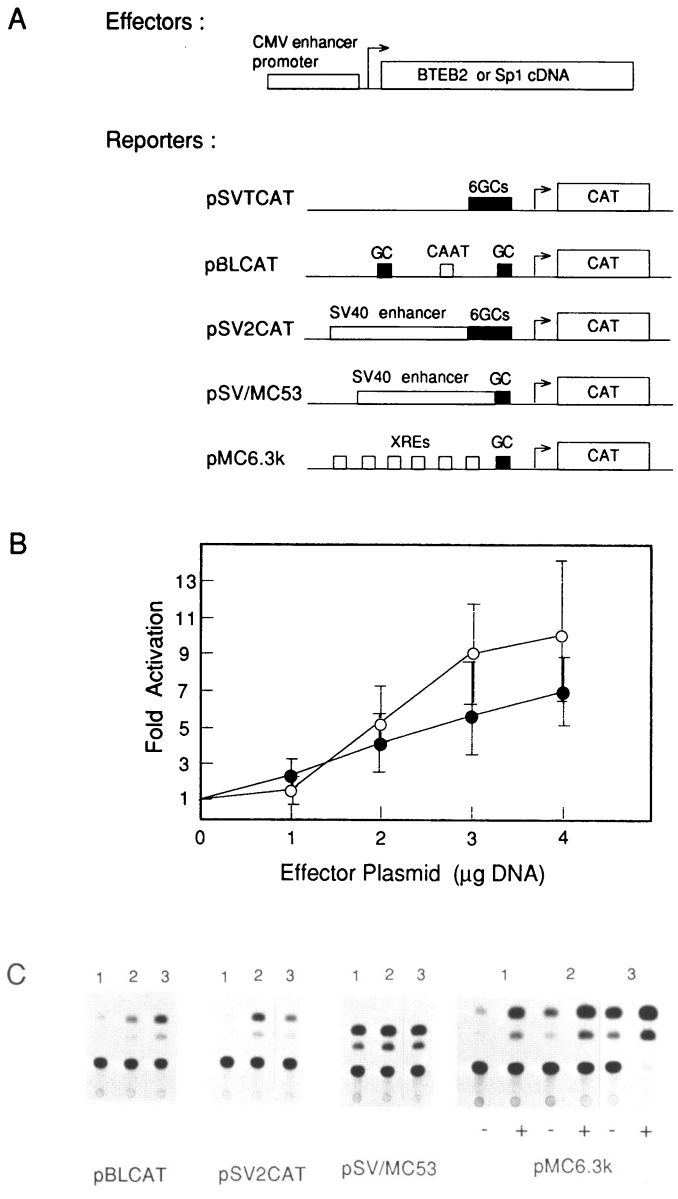

Figure 4. Effects of BTEB2 and Sp1 on the expression of CAT activity directed by promoters with GC box sequences. **(A)** Schematic representations of effector plasmids, pCMBTEB2 and pCMSp1, and of reporter plasmids, pSVTCAT (SV40 promoter) (36), pBLCAT (TK promoter) (37), pSV2CAT (SV40 promoter and enhancer) (16), pSV/MC53 (P-450c promoter and SV40 enhancer) (35), and pMC6,3k(P-450c promoter and enhancer) (15). Scales for genes are arbitrary. (B, C) CAT assay. A reporter plasmid, pSVTCAT (2μg), was cotransfected with increasing amounts of the effector plasmids, pCMBTEB2 (○) and pCMSp1 (●), into Hepa-1 cells and relative CAT activity was quantified **(B)**. pCMSV was used to adjust the total amounts of transfected plasmids to a constant level. Results are average of three independent experiments. Several GC box-containing reporter plasmids (2μg) were cotransfected with 4μg of effector plasmids **(C)**, 1. pCMSV. 2. pCMBTEB2. 3. pCMSp1. Plus and minus indicate the presence and absence of the inducer 3-methylcholanthrene(1μM), respectively, in the culture medium. Representative autoradiograms were shown obtained from three independent experiments.

and the other encoded a novel zinc finger protein different from BTEB. Fig.1 shows the cDNA and its deduced amino acid sequences. Since the nucleotide sequence around the first methionine poorly matched with the Kozak's rule (19) on the translation initiation, we performed *in vitro* translation experiments using reticulocyte lysate and BTEB2 mRNA synthesized *in vitro*. The results suggested that the first two methionine residues (Met 1 and Met 11) were relatively evenly used as the initiator methionine although the nucleotides around the second one fitted better in the Kozak's consensus sequence than the first (data not shown). The cDNA codes for a protein consistng of 219 amino acids with three Cys-Cys/His-His zinc finger motifs near the C-terminus of the protein. The zinc finger sequence showed 64% and 59% similarity to that of BTEB and Sp1, respectively, as shown in Fig. 2A. Furthermore, a short basic region followed by zinc finger motifs was also conserved between the deduced protein and Sp1. This basic sequence satisfied the N-terminal half of the consensus sequence of the basic domain of a group of DNA-binding proteins containing helix-loop-helix and leucine zipper (HLH-LZ) motifs as shown in Fig. 2B. This sequence conservativeness was not marked in

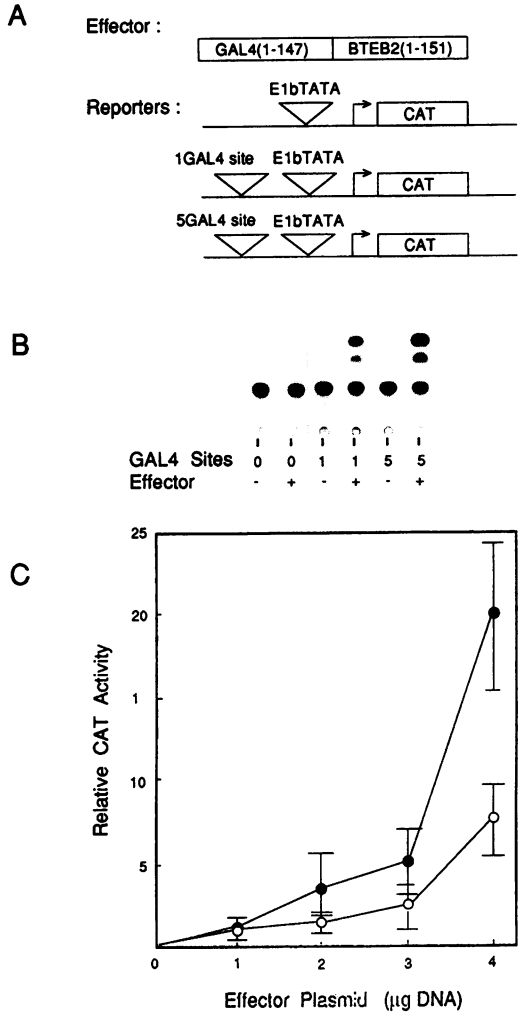

Figure 5. **(A)** Schematic representations of the GAL4-BTEB2 chimeric plasmid and reporter plasmids. pGAL4-BTEB2 contains the N-terminal half (1−151) of the coding region of BTEB2 cDNA fused to the C-terminus of GAL4 (1−147). Reporter plasmids contain E1b TATA box with (one or five) or without GAL4 binding sites (25). (B, C) CAT assay. Reporter plasmids (2μg), containing one (○) or five (●) GAL4 binding sites, were cotransfected with increasing amounts of the effector plasmids into Hepa-1 cells and relative CAT activity was quantified. pSG424 without BTEB2 cDNA sequence was used to adjust the total amounts of transfected plasmids to a constant level. Results are average of three independent experiments. A representative autoradiogram of CAT expression (4μg of pGAL4-BTEB2 was used) was also shown in **(B)**.

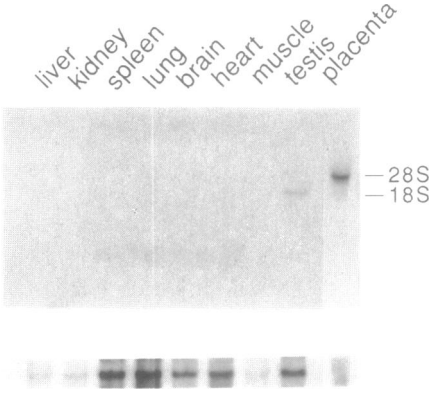

Figure 6. BTEB2 mRNA expression in various organs. Total RNA (15μg) was prepared from various rat organs and rabbit placenta, and separated on agarose/formaldehyde gel, and transfered to nitrocellulose filters. The filters were hybridized with a BTEB2 specific probe. The same filters were rehybridized with a β-actin probe. The position of 18S and 28S RNAs are indicated.

BTEB, although the corresponding region in BTEB was rich in basic amino acids (5). Outside of the two domains described above, the protein (residues 1−127) was characteristically rich in proline (20/127), serine (14/127), threonine (13/127), and alanine (13/127), which are known to constitute activation domains of a number of transcription factors. Comparison of the predicted sequence with those of the available protein sequence databases revealed no protein with significant similarity, apart from regions of homology discussed above. We, therefore, designated this protein BTEB2.

DNA-binding activity of BTEB2

In order to examine DNA-binding activity of BTEB2, we expressed it as a fusion protein in *E.coli* After induction by IPTG, extracts were prepared and used for binding experiments. As shown in Fig.3A, induced extracts showed a strong binding activity to the synthetic oligonucleotide of BTE (a GC box sequence in the promoter of the CYP1A1(P450c) gene(5)). This binding was competed out with unlabeled BTE but not with BTE carrying double point mutations, suggesting that the DNA-binding was sequence-specific. Some other GC boxes such as those of SV40 early promoter (20), HIV LTR (21), and adenovirus E1B promoter (22) showed a binding affinity to the *E.coli*-expressed BTEB2 (data not shown). Furthermore, we compared the binding specificity of BTEB2 with that of cellular Sp1 using eight synthetic mutant BTEs as competitors in gel mobility shift assays. As shown in Fig 3B, typical shifted bands of Sp1-GC box complexes (23) and again one band of a BTEB2-GC box complex were observed. Although these bands were competed out with unlabeled BTE, mutations present in the middle of the consensus GC box sequence abolished the affinity toward Sp1 and BTEB2. Competition patterns of BTEB2- or Sp1-BTE complexes with the mutant oligonucleotides were very similar to each other. In similar experiments, *E.coli*-expressed Sp1 as well as BTEB gave essentially the same results (data not shown). This observation indicated that BTEB2 and Sp1 show very similar DNA-binding specificity.

BTEB2 activates transcription *in vivo*

Fig.4A shows structures of expression plasmids of the zinc finger proteins and several reporter plasmids containing GC boxes and the CAT structural gene. As shown in Fig. 4B, BTEB2 as well as Sp1 activated the CAT expression from a reporter gene carrying SV40 early promoter in a dose-dependent manner. The transactivation by BTEB2 was somewhat stronger than that by Sp1. We further tested transactivation by BTEB2 using some other reporter plasmids containing GC box sequences in their promoter region. As shown in Fig. 4C, transcription of all the reporter genes was activated, although activation was weak with promoters containing a single GC box. Interestingly, pRSVCAT having no GC boxes in its promoter was also weakly activated by BTEB2 (data not shown). Similar observations were reported when Sp1 and BTEB were used as effectors (5, 24). To confirm the transcription-enhancing activity of BTEB2, a plasmid was constructed to express a fusion protein that contains the GAL4 DNA-binding domain (residues 1−147) at its N-terminus (Fig

5A) and N-terminal 151 amino acids of BTEB2. Cotransfection of the GAL4-BTEB2 plasmid with the target plasmid containing one GAL4-binding site, (Fig 5A, (25)) markedly enhanced the CAT activity (Fig. 5B). Additional GAL4 sites in tandem arrangement markedly amplified the expression presumably by increasing the number of bound GAL4-BTEB2 molecules. The expressed CAT activity was progressively increased as the amount of the transfected effector plasmids was increased (Fig. 5C). When reporter plasmids lacking GAL4 binding sites were used, the activation of the CAT expression was very weak, indicating that the activation was dependent on the presence of the GAL4-binding sites. Taken together, these results clearly demonstrate that BTEB2 is a transcriptional activator.

Expression of the BTEB2 mRNA in various organs

In order to investigate tissue specificity of expression of BTEB2, we prepared total RNA from various rat organs and used for RNA blot analysis, because partial sequence analysis of rat BTEB2 genes revealed more than 90% sequence similarity between human and rat BTEB2 genes. Fig.6 shows that BTEB2 mRNA was detected only in testis and placenta with different sizes, 20S and 28S, respectively. Both mRNA species remained detectable even when hybridized filters were washed under high stringent conditions. This tissue-specific expression pattern was in contrast with ubiquitous expression of the Sp1 and BTEB mRNA (5). Since we isolated three different genomic clones with the human BTEB2 cDNA used as a hybridization probe from a rat genomic library, the two mRNA species may come from different BTEB2 or BTEB2-related genes, although we cannot exclude the possibility that different processes of mRNA maturation may be involved (eg. alternative splicing and/or different usage of poly(A) signals).

DISCUSSION

BTEB2 contains three characteristic domains for transcriptional regulators. One is zinc finger motifs which have been established as a DNA binding domain (26, 27), and another is a short basic sequence adjacent to N-terminal of the zinc finger sequences. Among zinc finger sequences of BTEB2, BTEB, and Sp1, the amino acids which are considered to be important for recognition of the DNA sequence by X-ray crystallography (28) are invariable (Fig. 2A). Together with the results obtained from gel shift competition assay as shown in Fig. 3B, this fact strongly suggests the specificity of DNA-binding was closely similar among BTEB2, BTEB and Sp1. Furthermore, we could not detect any difference in the specificity of the binding sequences among the three proteins, as tested by hydroxy radical footprinting and methylation interference footprinting (data to be published). The basic sequence concurs with the N-terminal half of the consensus sequence for the basic domain of proteins containing HLH-LZ structure (29). The basic domain of HLH-LZ is responsible for the DNA binding and recognizes the consensus DNA sequence, CACGTG, after forming dimers or tetramers (29). By analogy, it seems possible that the basic sequence of BTEB2 and Sp1 may play an auxiliary role for DNA binding. Since the first zinc finger motif would interact with 3'-downstream triplet bases (no 7–9, see Fig. 3) of the consensus GC box sequence (28, 30), the basic sequence might interact with bases located further 3'-downstream of the consensus sequence. According to this hypothesis, the most downstream base C or T in the consensus, $^{G}/_{T}{}^{G}/_{A}$GGCG$^{G}/_{T}{}^{G}/_{A}{}^{G}/_{A}{}^{C}/_{T}$ (31), and a few bases further downstream may correspond to the region recognized by the basic sequence of BTEB2 and Sp1. It is also possible that this basic sequence of BTEB2 fulfils other functions such as a signal for nuclear translocation postulated for the basic sequence of Sp1 (4).

Thirdly, proline residue is the most abundant amino acid in BTEB2, especially in a region of residues 44–110 (16 out of 67). A proline-rich domain is known to be one of the motifs which are directly responsible for transcription activation as clarified by the studies on CTF/NF-I (32). Although the amino acid sequence of the proline-rich region of BTEB2 was totally different from that of CTF/NF-I, relative contents of proline residues and sizes of the domain of the two regulatory proteins seem to be comparable to each other. When the polypeptide of residues 37–111 of BTEB2 was expressed as a chimeric protein with GAL4 DNA binding domain, the expressed protein could enhance transcription of reporter genes containing GAL4 DNA binding sites in their promoter region (K. Sogawa et al. unpublished observation). This result strongly suggests that the proline-rich sequence of BTEB2 constitutes an independent activation domain. Further molecular dissection and mutational analysis which define the amino acid residues responsible for the transactivation are necessary and studies along this line are now in progress.

The expression of the BTEB2 mRNA is limited to testis and placenta as shown in Fig. 6. Physiological significance for this tissue specificity is unknown. Recently, several data reported that intracellular levels of Sp1 vary from tissue to tissue (33) and we found that the Sp1 expression is relatively low in testis and brain (5). It may be considered that BTEB2 replaces functionally Sp1 in these tissues or that some testis or placenta-specific genes specifically require BTEB2 for their expression for reasons to be clarified. Recently, it has been revealed that many of the DNA regulatory elements are recognized by multiple trans-acting regulatory factors cooperatively, sometime in dimer formation or competitively to exert their transcriptional effects in subtle and complex manners. The finding of novel GC-box binding factors such as BTEB2 and BTEB in addition to Sp1 strongly suggests that complex regulatory mechanisms work also on the expression of genes through GC-box DNA elements.

ACKNOWLEDGEMENTS

We would like to thank Dr O.Gotoh for computer analyses. We also thank Drs A.Adachi, M.Ptashne and M.Obinata for providing pCMSV, pSG424 and pBLCAT, respectively and Dr M.Green for pG_5EC, pG_1EC and pE_1bCAT. K.S. thanks Mr Ema for assistance in sequence analysis. This work was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science and Culture of Japan, and by a grant from the Mitsubishi Foundation.

REFERENCES

1. Dynan, W. S., and R. Tjian. (1985). Nature **316**:774–778.
2. Mitchell, P. J., and R. Tjian. (1989). Science **245**:371–378.
3. Courey, A. J., and R. Tjian. (1988). Cell **55**:887–898.
4. Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian. (1987). Cell **51**:1079–1090.
5. Imataka, H., K. Sogawa, K. Yasumoto, Y. Kikuchi, K. Sasano, and Y. Fujii-Kuriyama. (1992). EMBO J. **11**:3663–3671.
6. Benton, W. D., and R. W. Davis. (1977). Science **196**:180–182.
7. Sanger, F., S. Nicklen, and A. R. Coulson. (1977). Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
8. Maxam, A. M., and W. Gilbert. (1977). Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.

9. Studier, F. W., and B. A. Moffatt. (1986). J. Mol. Biol. **189**:113–130.
10. Ema, M., K. Sogawa, N. Watanabe, Y. Chujoh, N. Matsushita, O. Gotoh, Y. Funae, and Y. Fujii-Kuriyama. (1992). Biochem. Biophys. Res. Commun. **184**:246–253.
11. Sadowski, I., and M. Ptashne. (1989). Nucleic Acids Res. **17**:7539.
12. Bernhard, H. P., G. J. Darlington, and F. H. Ruddle. (1973). Dev. Biol. **35**:83–96.
13. Dignum, J. D., R. M. Lebovitz, and R. G. Roeder. (1983). **11**:1475–1489.
14. Banerji, J., S. Rusconi, and W. Schaffner. (1981). Cell **27**:299–308.
15. Fujisawa-Sehara, A., K. Sogawa, C. Nishi, and Y. Fujii-Kuriyama. (1986). Nucleic Acids Res. **14**:1465–1477.
16. Gorman, C. M., L. F. Moffat, and B. H. Howard. (1982). Mol. Cell. Biol. **2**:1044–1051.
17. Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter. (1979). Biochemistry **18**:5294–5299.
18. Thomas, P. S. (1980). Proc. Natl. Acad. Sci. U.S.A. **77**:5201–5204.
19. Kozak, M. (1984). Nucleic Acids Res. **12**:857–872.
20. Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian. (1986). Science **234**:47–52.
21. Jones, K. A., J. T. Kadonaga, P. A. Luciw, and R. Tjian. (1986). Science **232**:755–759.
22. Schmidt, M. C., Q. Zhou, and A. J. Berk. (1989). Mol. Cell. Biol. **9**:3299–3307.
23. Schaufele, F., B. L. West, and T. L. Reudelhuber. (1990). J. Biol. Chem. **265**:17189–17196.
24. Swick, A. G., M. C. Blake, J. W. Kahn, and J. C. Azizkhan. (1989.) Nucleic Acids Res. **17**:9291–9304.
25. Lillie, J. W., and M. R. Green. (1989). Nature **338**:39–44.
26. Brown, R. S., C. Sander, and P. Argos. (1985). FEBS Lett. **186**:271–274.
27. Miller, S., A. D. Mclachlan, and A. Klug. (1985). EMBO J. **4**:1609–1614.
28. Pavletich, N. P., C. O. Pabo. (1991). Science **252**:809–817.
29. Fisher, D. E., C. S. Carr, L. A. Parent, and P. A. Sharp. (1991). Genes & Dev. **5**:2342–2352.
30. Nardelli, J., T. J. Gibson, C. Vesque, and P. Charnay. (1991). Nature **349**:175–178.
31. Kadonaga, J. T., K. A. Hones, and R. Tjian. (1986). Trends Biochem. Sci. **11**:20–23.
32. Mermod, N., E. A. Oneill, T. J. Kelly, and R. Tjian. (1989). Cell **58**:741–753.
33. Saffer, J. D., S. P. Jackson, and M. B. Annarella. (1991). Mol. Cell. Biol. **11**:2189–2199.
34. DePinho, R., K. S. Hatton, A. Tesfaye, G. D. Yancopoulos, and F. W. Alt. (1987). Genes & Dev. **1**:1311–1326.
35. Yanagida, A., K. Sogawa, K. Yasumoto, and Y. Fujii-Kuriyama. (1990) Mol Cell. Biol. **10**:1470–1475.
36. Inoue, H., N. Watanabe, Y. Higashi, and Y. Fujii-Kuriyama. (1991). Eur. J. Biochem. **195**:563–569.
37. Luckow, B., and G. Schütz. (1987). Nucleic Acids Res. **15**:5490