Nucleic Acids Research, 1994, Vol. 22, No. 22 4733–4741

Alternative promoter usage and splicing options result in the differential expression of mRNAs encoding four isoforms of chicken VBP, a member of the PAR subfamily of bZIP transcription factors

John B.E.Burch* and Dorene L.Davis
Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA

Received June 27, 1994; Revised and Accepted September 23, 1994 EMBL accession nos U09221–U09223 (incl.)

ABSTRACT

We previously isolated a set of overlapping cDNA clones that encoded a unique open reading frame for the chicken VBP transcription factor. We now report the isolation of a cDNA clone that encodes a complete open reading frame for a VBP isoform that differs from the previously reported sequence at both the amino-terminal and carboxyl-terminal ends. An analysis of the VBP gene revealed that the two different amino-terminal sequences map to alternative first exons and that the two different carboxyl-terminal sequences reflect an optional splicing event which can occur only on transcripts that are polyadenylated at the more distal of two polyadenylation sites. An RT–PCR analysis further revealed that a total of four VBP isoforms are encoded by the combinatorial use of these two splicing options. The mRNAs for these four isoforms are differentially expressed in different tissues and cell types. We provide evidence that one function of the amino-terminal domains is to impose cell type specificity on a core transactivation domain that is present in all four isoforms. Since it is known that VBP can heterodimerize with other members of the PAR subfamily of bZIP factors, our evidence for four VBP isoforms greatly expands the number of complexes that may be used to effect transcriptional regulation through PAR-factor binding sites.

INTRODUCTION

Transcriptional control regions are comprised of sets of transcription factor binding sites, with the information content of each promoter or enhancer being determined by both the qualitative nature of the individual binding sites and their spatial relationships to each other (for reviews see 1–4). The sequence of a given module is not necessarily sufficient, however, to dictate occupancy by a unique transcription factor. This is most evident in cases where several transcription factors belonging to a multigene family are co-expressed in the same cell type. The fact that many transcription factors bind to their target sequences as dimers, and are able to form heterodimeric complexes with related factors, further underscores the range of influences that must somehow be integrated at critical regions to ensure appropriate levels of target gene expression.

In the last several years it has become increasingly apparent that different transcription factor isoforms may be expressed from single copy genes (for a recent review see 5). In some cases this is achieved by the alternative usage of different translational initiation sites. Examples of this include the short and long forms of CREMt (6) and LAP/LIP (7). In other cases multiple isoforms are specified by alternative splicing pathways. Three antagonists (CREM$\alpha/\beta/\gamma$) and one agonist (CREMτ) of cAMP-responsive transcription are encoded by alternatively spliced CREM transcripts (8) and additional examples where alternative splicing pathways are used to generate functionally distinct isoforms of transcription factors are provided by the genes for FosB (9), the thyroid hormone receptor (10), CREB (11), and c-*myb* (12), to name only a few.

The chicken VBP transcription factor that is the subject of this report (13) belongs to the PAR subfamily of bZIP transcription factors (14). The members of this PAR subfamily have nearly identical basic regions and recognize similar, if not identical, binding sites. Although the leucine zipper regions for these factors are similar in only four of the seven helical spokes (13), these factors are able to bind as heterodimers (13–15). A proline and acid-rich (hence PAR) region located immediately upstream of the core basic region is also conserved between members of this subfamily (14), and evidence has been presented that these PAR domains contribute to the binding properties of these factors (14). In addition to VBP (13) and its presumptive rat homolog, TEF (14), the PAR subfamily also includes rat DBP (16) and human HLF (15,17). The rat DBP gene is remarkable in that it is transcribed only during a brief period each day in response to circadian cues (18). Superimposed upon this circadian control is a post-transcriptional control(s) which results in DBP being

*To whom correspondence should be addressed

a liver-enriched factor (16). It has been inferred that HLF is also liver-enriched, based on the finding that HLF mRNA is largely restricted to this tissue (15,17). Although TEF mRNA is initially restricted to thyrotroph cells in the developing pituitary (14), TEF and VBP mRNAs are expressed in many, if not all, somatic tissues in adult animals (13,14). These studies suggested that heterodimeric PAR bZIP complexes might be restricted to the liver. However, we now provide evidence that VBP is not a unique factor but rather is a collection of four distinct isoforms. Whereas VBP mRNAs are synthesized in all tissues examined, we show that tissues differ with respect to the particular VBP isoforms that are expressed.

MATERIALS AND METHODS

Isolation and characterization of a novel VBP cDNA clone

A chicken embryo cDNA library was screened (18) with a random prime labeled VBP cDNA probe (13) that was prepared using Multiprime reagents (Amersham). The cDNA inserts from the positive phage clones were isolated and cloned into pGEM-3 (Promega) using standard protocols (18). A preliminary sequence analysis using supercoiled plasmid templates (19) and Sequenase reagents (US Biochemicals) revealed that one insert was different from the VBP cDNA clones we had previously analyzed (13). The entire sequence of this cDNA insert was determined from two sets of nested *Bal*I deletion constructs.

Isolation and characterization of VBP genomic clones

Two chicken genomic libraries (kindly provided by Kathleen Conklin and Sin Jen Kung) were screened with random prime labeled VBP cDNA probes using standard protocols (18). Restriction maps for the two positive phage clones were determined by an indirect end labeling assay of partially digested phage DNA (18). Overlapping restriction fragments were cloned into pGEM-3Z (Promega) and the exonic regions were sequenced using a set of primers directed against various portions of the VBP cDNA sequences. The two alternative first exons (see text) were sequenced on both strands by making *Bal*31 deletion constructs from both ends of each of two fragments that span these respective regions.

RNase mapping and primer extension

Poly(A^+) RNAs were isolated from the liver and oviduct of an egg-laying hen using the Fast Track mRNA isolation kit (Invitrogen). An aliquot of each (5 μg) was added to an aliquot of yeast tRNA (45 μg) and precipitated with 2.5 M ammonium acetate (pH 4.5) and two volumes of absolute ethanol for 30 min at −70°C. As controls, aliquots of yeast tRNA were also precipitated without tissue RNA. The RNA pellets were dissolved in 30 ml of hybridization buffer [80% formamide, 40 mM PIPES (pH 6.4), 400 mM NaCl, 1 mM EDTA (pH 8.0)] containing 10^6 c.p.m. of an antisense riboprobe spanning the alpha first exon (for the liver RNA sample) or an antisense riboprobe spanning the beta first exon (for the oviduct RNA sample). The samples were heated for 5 min at 85°C and then for 16 h at either 60°C (for the alpha probe) or 45°C (for the beta probe).

Following hybridization, each RNA sample was ethanol precipitated in the presence of 2.5 M ammonium acetate (pH 4.5) and then resuspended in 350 μl of digestion buffer [300 mM NaCl, 10 mM Tris (pH 8.0), 5 mM EDTA (pH 7.5)]. The samples were incubated for 60 min at 37°C after adding either 6 μg of ribonuclease T1 (for the alpha probe) or 2 mg of ribonuclease T1 and 40 mg of ribonuclease A (for the beta probe). After adding 20 μl of 10% SDS and 5 μl of a stock solution (10 mg/ml) of proteinase K, each sample was extracted with 400 μl of phenol:chloroform (1:1) and the aqueous phase was transferred to a fresh Eppendorf tube that contained 10 μg of yeast tRNA. The RNAs were precipitated as described above, dried under vacuum and then each pellet was dissolved in 3 μl of RNA loading buffer [80% formamide, 1 mM EDTA (pH 7.5), 0.1% bromophenol blue, 0.1% xylene cyanol]. The samples were heated for 3 min at 85°C and then resolved on a 5% denaturing polyacrylamide cast in TBE (18). The dried gel was exposed to X-OMAT film between two intensifying screens (Dupont Lightening Plus) for 16 h at −70°C. The sizes of the protected fragments were determined by reference to a sequence ladder that was also included on the gel.

The primer extension assay for the beta first exon was carried out using a synthetic oligo (5′ TGCCAAATAACTCCGGCTG-AGAAAACCCGAGGAGG 3′). A 100 ng aliquot of this oligo was 5′ end labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP, purified on a 15% acrylamide denaturing gel (18) and resuspended in 0.3 M sodium acetate to a concentration of 2×10^5 c.p.m./ml. A 5 μg aliquot of poly(A^+) RNA, prepared using a Fast Track kit (Invitrogen), was mixed with 45 mg of yeast tRNA and 2×10^5 c.p.m. of the beta primer in 0.3 M sodium acetate. A control sample which lacked poly(A^+) RNA was also analyzed as a negative control. Following ethanol precipitation, the pellets were resuspended in 30 ml of hybridization buffer [80% formamide, 40 mM PIPES (pH 6.4), 400 mM NaCl, 1 mM EDTA (pH 8.0)], heated to 85°C for 5 min and then incubated for 16 h at 34°C. The samples were again ethanol precipitated and then resuspended in 25 μl of reverse transcriptase buffer [50 mM Tris (pH 8.2), 5 mM $MgCl_2$, 5 mM dithiothriotol, 50 mM KCl, 50 μg/ml bovine serum albumin, 5.6 mM of each dNTP]. Following the addition of 1 μl of RNasin (Promega) and 4 μl (40 units) of AMV reverse transcriptase (Life Sciences), the samples were incubated for 90 min at 42°C. The reactions were stopped by the addition of 1 μl of 0.5 M EDTA (pH 8.0) and the samples were then heated for 30 min at 37°C in the presence of 1 μg of ribonuclease A. The samples were diluted with 100 μl of 2.5 M ammonium acetate, extracted with an equal volume of phenol:chloroform (1:1), and the aqueous phases were ethanol precipitated. Pellets were washed with 70% ethanol, dried under vacuum, resuspended in 3 μl of formamide loading buffer (Maniatis) and the DNAs were resolved on a 5% denaturing polyacrylamide gel (Maniatis). A DNA sequencing ladder was included on the gel to provide a size reference.

Construction of expression and reporter vectors

Sense oligos for VBP(α/α) and VBP(β/β) were designed with *Nco*I sites engineered over the respective initiation codons (CCATGG). *Nco*I sites were also engineered into antisense oligos that flank the VBP(α/α) and VBP(β/β) stop codons. These sense and antisense primers (50 pmol of each) for VBP(α/α) or VBP(β/β) were used to PCR amplify the respective ORFs from 1 ng aliquots of the appropriate plasmid. Standard PCR conditions were used (1 min at 95°C, 1 min at 50°C, 1.5 min at 72°C, 30 cycles, followed by 4 min at 72°C). After digestion with *Nco*I, these two PCR products were ligated into the *Nco*I site of an *in vitro* expression vector [pT7βglobin (20); kindly provided by Tom Kadesch]. Subclones bearing the correct orientations of the

VBP(α/α) and VBP(β/β) ORFs were identified by restriction digests, and the absence of inadvertent mutations was verified by completely sequencing each insert using synthetic oligos as described above. In order to construct *in vitro* expression vectors for VBP(α/β) and VBP(β/α), we took advantage of a *Hin*dIII site that is common to both VBP(α/α) and VBP(β/β). The 5′ *Nco*I−*Hin*dIII fragment from VBP(α/β) and the 3′ *Hin*dIII−*Nco*I fragment from VBP(β/β) were ligated at the *Hin*dIII site and cloned into the *Nco*I site of pT7βglobin. A similar strategy was used to engineer the VBP(β/α) ORF into pT7βglobin.

A segment of the VBP(α/α) ORF (aa 79−313; see Figure 1), which we have denoted VBP(t/α) (see text), was also PCR cloned into the pT7βglobin expression vector using a sense oligo with an *Nco*I site engineered over the methionine codon (residue 79) and an antisense oligo that flanks the termination codon (see above). This insert was also completely sequenced to confirm that no mutations were inadvertently introduced.

The *Nco*I fragments from these pT7βglobin/VBP vectors were cloned in place of the C/EBP ORF in the R18 RSV-driven expression vector (21) for use in co-transfection assays. The reporter plasmid used in these studies contains eight copies of the VBP binding site from the chicken vitellogenin II promoter (13) cloned in a head-to-tail arrangement upstream of a minimal promoter containing a TATA box from the liver/bone/kidney alkaline phosphatase gene promoter [(20); kindly provided by Tom Kadesch]. To make this reporter plasmid, we first cloned a single VBP binding site with *Bam*HI and *Bgl*II ends into the *Bam*HI and *Bgl*II sites located upstream of the TATA box so as to recreate both sites and yield the plasmid p(VBP)$_1$/TATA/CAT. The oligos used (5′ GATCCTTATCAGTTTATGTAAACACTCCTA 3′ and 5′ GATCTAGGAGTGTTTACATAAACTGATAAG 3′) in this first step contain the −106 to −83 region from the chicken vitellogenin II promoter (13). Since the parental plasmid contains a unique *Bgl*I site that is located further upstream of the promoter, by cloning the *Bgl*I−*Bgl*II fragment from p(VBP)$_1$/TATA/CAT into the *Bgl*I−*Bam*HI fragment of p(VBP)$_1$/TATA/CAT, we were able to generate p(VBP)$_2$/TATA/CAT. Two subsequent rounds of this strategy yielded p(VBP)$_8$/TATA/CAT.

Segments of the VBP(α/α) ORF were cloned into the GAL4(1-147) vector (22) using a PCR protocol similar to the one described above for cloning the VBP(α/α) ORF into the pT7βglobin vector. In this case, however, a *Kpn*I site was engineered into the sense oligo and an *Xba*I site was engineered into the antisense oligo in order to be able to directionally clone the PCR products. The endpoints listed in Figure 8 for the various GAL4(1-147)/VBP constructs denote the first VBP codon following the *Kpn*I site. The sequence of each cloned VBP segment was determined to verify that no mutations had been introduced during the PCR amplification. The reporter plasmid used in these studies (kindly provided by Tom Kadesch) contains five GAL4 binding sites cloned upstream of a minimal TATA box promoter (22).

Each of these expression and reporter supercoiled plasmids was isolated on a CsCl gradient (18) and resuspended to a final concentration of 0.5−1.0 mg/ml in TE [10 mM Tris, 1 mM EDTA (pH 7.5)] as assayed by UV absorbance.

Co-transfection assays

The methods used to culture LMH cells (23) and chicken embryo fibroblast (CEF) cells have been described in detail in a previous publication (24). The cells were cultured on 60 mm diameter dishes and co-transfected with 5 μg each of the relevant expression and reporter plasmids using the calcium phosphate method (25). After culturing for an additional 48 h, cells were harvested and total cell extracts were prepared according to a standard protocol (26). Protein concentrations were determined using the method of Bradford (27) and chloramphenicol acetyltransferase (CAT) enzyme assays (26) were performed using 10−50 μg of total cell protein. An Ambis radioanalytic imaging system was used to count directly the amounts of [^{14}C]chloramphenicol and acetylated products. The data presented represent the results obtained from two completely independent co-transfection assays.

In vitro transcription and translation

The pT7βglobin/VBP expression vectors (see above) were used to program *in vitro* transcription and translation reactions as previously described (13), except that *Sal*I was used to linearize the plasmid templates downstream of the VBP inserts.

Gel shift assays

The same pair of oligos that were used to introduce the VBP binding site into the p(VBP)$_1$/TATA/CAT construct (see above) were also used to make a probe for gel shift assays. A 100 ng aliquot of one oligo was labeled using [γ-^{32}P]ATP and polynucleotide kinase and then passed through a Sephadex G-50 column to remove the unincorporated [γ-^{32}P]ATP (18). This labeled oligo was annealed with 300 ng of the other (unlabeled) oligo and the resultant double-stranded probe was resolved on, and purified from, a 12% non-denaturing polyacrylamide gel (18). Probe aliquots (0.1 ng) were mixed with 1 μl aliquots of *in vitro* translated proteins in a final volume of 30 μl of incubation buffer [250 mM KCl, 12 mM Tris (pH 7.8), 1 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1 mM dithiothreitol, 5 mM NaCl, 0.1 mg/ml bovine serum albumin, 5% glycerol, 8 μg poly(dI−dC)]. After incubating on ice for 1 h, free DNA and protein−DNA complexes were resolved on a 6% non-denaturing polyacrylamide gel that had been prerun for 1 h at 4°C. Gels were run at a constant voltage of 100 V, and the electrophoresis buffer [0.25× TBE (18)] was circulated manually every 30 min. The dried gel was exposed to X-OMAT film for 6 h at −70°C between two (Dupont Lightening Plus) intensifying screens.

RT−PCR analysis

PolyA(+) RNAs were isolated from the indicated tissues and cell types (see Figure 8) using the Fast Track mRNA isolation kit (Invitrogen). An aliquot of each RNA (1 μg) was copied into cDNA using 500 ng of oligo(dT)$_{17}$ and 15 units of AMV reverse transcriptase (Life Sciences). The reactions were performed for 60 min at 42°C followed by 30 min at 52°C in the reverse transcriptase buffer described in a preceding section, diluted to a final volume of 1 ml by the addition of TE [10 mM Tris (pH 7.5), 1 mM EDTA], and stored at −20°C. A 10 μl aliquot of each cDNA sample was used for each RT−PCR analysis. The sense oligos used for this analysis were specific for either the alpha first exon (5′ CTGAAGAAGCTGATGGAGAAC 3′), the beta first exon (5′ TTCTCAGCCGGAGTTATTTGG 3′) or the common second exon (5′ AAACCCATGGGATCCATGGCAGTCTCAGCTTCCCTCAT 3′). The antisense oligos used for this analysis were specific for either the common third exon (5′ GCGTCTAGACTTTTGCTCATCTGGGACAAA 3′), a sequence flanking the VBP(α/α) termination codon (5′ CGCC-

CATGGAGATCTCTGAGAGTCTCTACAGAAGGC 3'), or a sequence spanning the exon−exon junction present in the VBP(β/β) cDNA clone (5' AGATAAGTCAAAGGGTCCGT-A 3'). Control experiments were used to optimize conditions for amplifying each of the products shown in Figure 5. This was done using 1 ng aliquots of the four VBP isoform control plasmids described in a preceding section. The specificities of these reactions were confirmed using individual plasmid templates and mixtures of all four templates. The conditions used to PCR amplify the VBP(α/α), VBP(α/β), and internal control (exon 2−exon 3) PCR products shown in Figure 5 were: 1 min at 95°C, 1 min at 55°C, 1.5 min at 72°C, 30 cycles, followed by 4 min at 72°C. The conditions used to PCR amplify the VBP(β/α) and VBP(β/β) PCR products shown in Figure 5 were: 1 min at 95°C, 1 min at 58°C, 1.5 min at 72°C, 30 cycles, followed by 4 min at 72°C. The second set of PCR reactions were done in the presence of 1 μl of a polymerase enhancer (Perfect Match; Strategene). Following amplification, 20 μl aliquots of each of the 100 μl PCR reactions were resolved on 1.5% TAE agarose gels (18) and photographed using Polaroid 57 film.

RESULTS

The VBP gene has a complex structural organization

In a previous study we screened a chicken liver cDNA expression library with a binding site probe from the chicken vitellogenin II promoter and isolated a cDNA clone for a transcription factor that we denoted VBP, for vitellogenin gene binding protein (13). A unique open reading frame (ORF) for VBP was deduced by isolating and sequencing a set of overlapping cDNA clones. These cDNA clones could be grouped into two classes according to the use of alternative polyadenylation sites spaced 2.3 kb apart. Consistent with this analysis of cDNA clones, Northern blots of poly(A$^+$) RNA revealed two bands that differed in size by 2.3 kb in all somatic tissues and cell types examined (13).

We now report the isolation of a cDNA clone that specifies a VBP ORF with amino-terminal and carboxyl-terminal ends distinct from those described previously (Figure 1). In order to examine the basis for this heterogeneity, genomic clones spanning the VBP gene were isolated and all of the coding regions were mapped and sequenced. The previously described VBP ORF [which we will refer to henceforth as VBP(α/α)] was found to be encoded by four exons, and we infer that polyadenylation occurs either 0.1 kb (panel A of Figure 2) or 2.4 kb (panel B of Figure 2) downstream of the ORF (see above). The newly described VBP ORF [which we will refer to as VBP(β/β)] was found to be encoded by five exons (panel C of Figure 2). The first exon for VBP(β/β) is located ~4 kb upstream of the first exon for VBP(α/α) whereas the fourth and fifth exons are delineated by a splicing event that can occur only on transcripts that use the distal polyadenylation site. It is interesting to note that the two splice sites that deviate significantly from consensus sequences (sites 8 and 9; see panel D of Figure 2) correspond to sites where splicing must be regulated. The splice between sites 8 and 9 can occur (see panel C of Figure 2), but does not necessarily occur (see panel B of Figure 2) on transcripts that use the downstream polyadenylation site. The carboxyl-terminal extension for the VBP(β/β) isoform results from a splicing event that deletes the termination codon (TAA; see site 8 in panel D of Figure 2) that is otherwise used to terminate translation of the VBP(α/α) isoform. The locations of the exon−exon junctions within the VBP(α/α) and VBP(β/β) ORFs are indicated in Figure 1 and are in perfect agreement with the sequence analysis of the genomic exon junctions (data not shown).

It is noteworthy that the PAR and bZIP domains of VBP, which were previously delineated simply on the basis of sequence comparisons between members of the PAR subfamily of bZIP factors (14), are revealed by this analysis to be encoded by different exons.

The VBP gene has two alternative first exons

We previously inferred that the 1179 bp VBP(α/α) cDNA sequence [exclusive of poly(A)] corresponds to a nearly full-length sequence because it is essentially the same size as the

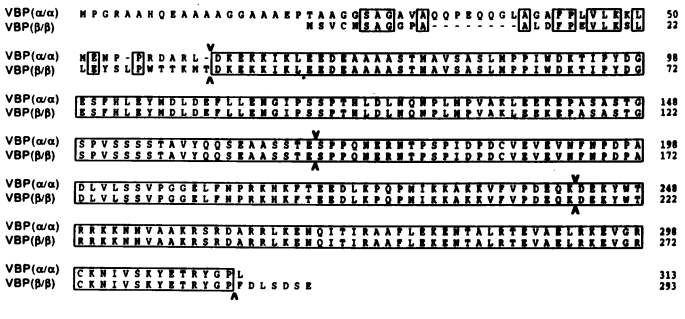

Figure 1. Evidence for a novel VBP isoform. A cDNA clone analyzed in the present study encodes a 293 aa VBP isoform [denoted VBP(β/β)] which is shown to differ from the 313 aa VBP isoform [denoted VBP(α/α)] that we previously reported (13) at both the amino-terminal and carboxyl-terminal ends. The positions of the exon−exon junctions for VBP(α/α) and VBP(β/β) are indicated by the arrowheads that are marked above and below the respective open reading frames.

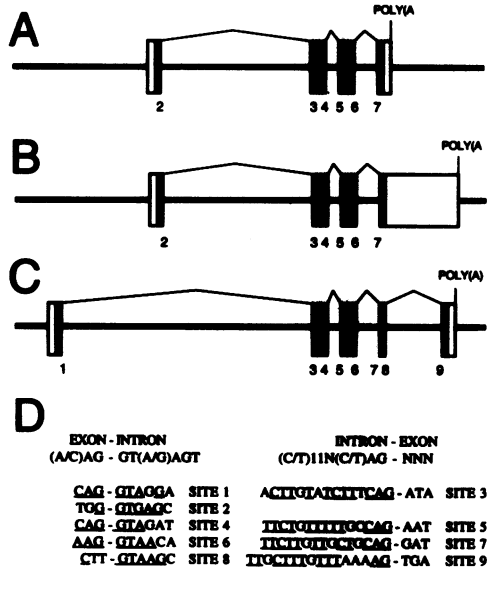

Figure 2. Alternative splicing pathways are used to generate mRNAs that encode VBP(α/α) and VBP(β/β). (A, B) The splicing and polyadenylation sites used to generate 1.2 kb and 3.5 kb mRNAs that encode VBP(α/α). (C) The splicing and polyadenylation sites used to generate a 1.2 kb mRNA that encodes VBP(β/β). The various splice sites are numbered in panels A−C, and the sequences of these sites are shown in (D) below the respective splice donor and acceptor consensus sequences.

smaller (1.2 kb) mRNA species that would result from the use of the proximal polyadenylation site (13). In order to confirm this inference, a genomic antisense probe spanning the alpha first exon was used in an RNase protection assay. As shown in panel A of Figure 3, a cluster of bands was generated in the presence of liver poly(A^+) RNA, which suggests that transcriptional initiation occurs at multiple sites. As anticipated, the 5′ ends of the two largest cDNA clones for the VBP(α/α) isoform (13) map within the region of multiple transcriptional start sites (panel A of Figure 4). This region is not preceded by a consensus TATA box, which is consistent with the lack of a unique transcriptional start site. As shown in panel A of Figure 4, this presumptive promoter contains numerous motifs that are potential binding sites for the ubiquitous SP1 transcription factor, which is common for promoters lacking TATA boxes (28). Although we have been unable to find conditions that would allow us to use a primer extension assay for this exon, presumably due to the fact that this region is (80%) GC-rich, the lack of a consensus splice acceptor site in this region gives us confidence that this is indeed a first exon. It is important to note that this caveat in no way undermines the conclusions that follow.

A similar analysis was carried out to delineate the beta first exon, except that poly(A^+) RNA was assayed from oviduct rather than liver since we have found (see below) that the liver is essentially devoid of transcripts for VBP(β/β). The results of the RNase protection assay for the beta first exon are shown in panel B of Figure 3. Importantly, the cluster of protected bands seen in panel B for the beta first exon probe are of comparable size to the cluster of bands seen in panel A for the alpha first exon probe, indicating that the two first exons are of similar length. These two samples were resolved on the same sequencing gel and are properly aligned in this figure. The primer extension analysis presented in panel C of Figure 3 confirms that this exon

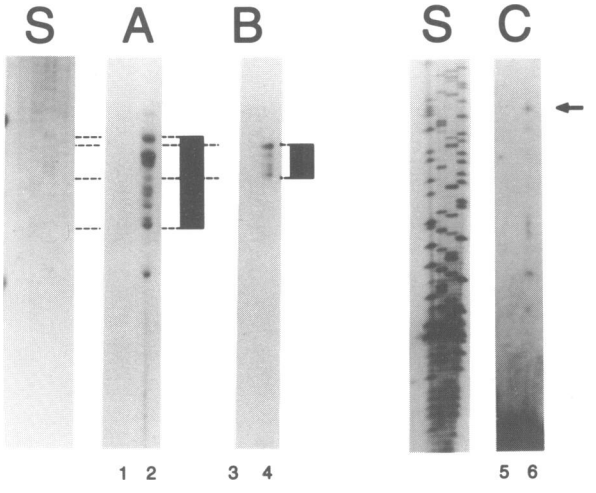

Figure 3. Mapping the transcription start sites for the alpha and beta first exons. (**A**) The products of an RNase protection assay obtained using an antisense riboprobe spanning the alpha first exon and either tRNA (lane 1) or liver RNA (lane 2). (**B**) The products of an RNase protection assay obtained using an antisense riboprobe spanning the beta first exon and either tRNA (lane 1) or oviduct RNA (lane 2). The transcriptional start sites that are inferred from this analysis cluster within regions (indicated by solid bars), the sequences of which are shown below in Figure 4. (**C**) The primer extension products obtained using a primer specific to the beta first exon and either tRNA (lane 5) or oviduct RNA (lane 6). The transcriptional start site inferred from the size of the largest product (indicated by an arrow) is also indicated by an arrow below in Figure 4. Note that one sequencing ladder (left panel S) was used for panels A and B, whereas a second sequencing ladder (right panel S) was used for panel C.

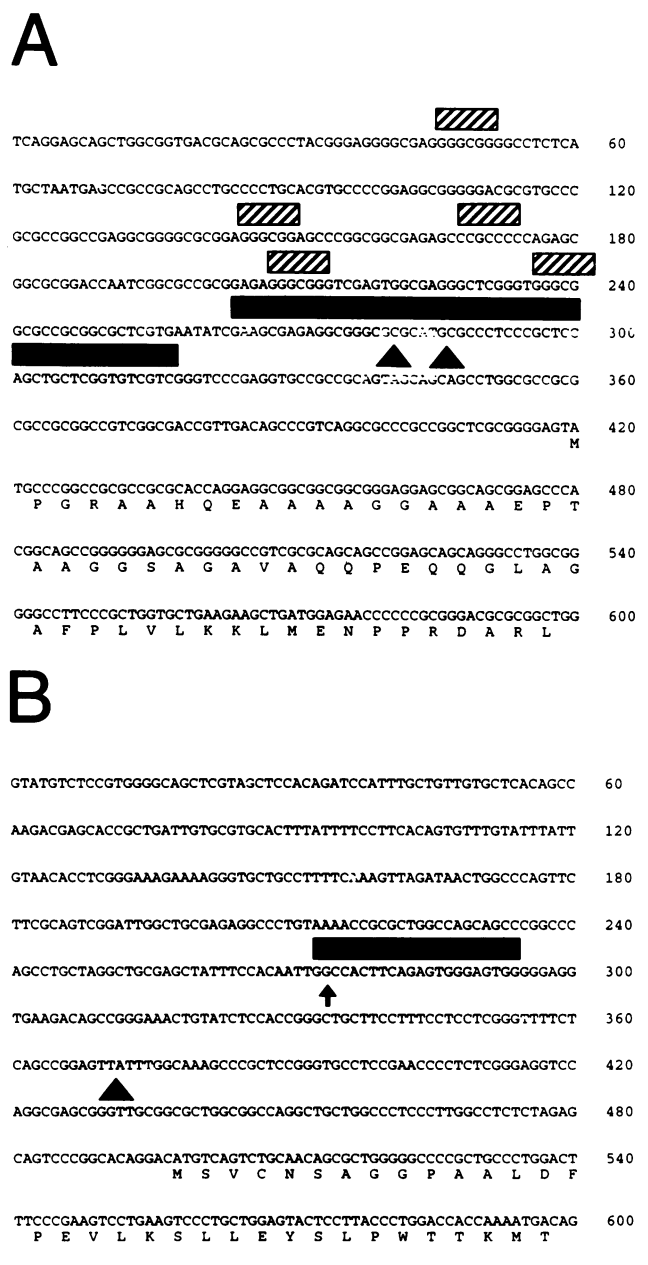

Figure 4. Sequences of the VBP alternative first exons and presumptive promoter regions. (**A, B**) The 600 bp regions immediately upstream of the introns that flank the alpha and beta first exons, respectively. The black bars shown above the sequences delimit the regions where multiple transcriptional starts sites are inferred to occur, as deduced from the RNase protection assays shown in panels A and B of Figure 3. The limit of primer extension products for mRNAs bearing the beta first exon is marked with an arrow and corresponds to the result obtained in panel C of Figure 3. The positions of potential SP1 binding sites are also indicated (by hatched boxes) in the region flanking the alpha first exon in panel A. The triangles shown below the sequences in panels A and B denote the 5′ endpoints of cDNA clones for VBP(α/α) (13) and VBP(β/β), respectively. Note that the presumptive initiation codons for the alpha and beta first exons are each preceded in the −3 position by a purine, consistent with the Kozak rules for translational initiation (35).

is indeed a first exon. The sequence of the beta first exon and the immediate upstream flanking region is shown in panel B of Figure 4. As in the case of the presumptive alpha promoter, the beta promoter lacks an obvious TATA box. However, in keeping with the more restricted expression from this promoter (see below), consensus binding sites for SP1 are not evident in the beta promoter region.

Messages for four VBP isoforms are differentially expressed

The modular structure of the VBP gene shown in Figure 2 led us to question whether additional isoforms might be encoded by other splice combinations. For example, in the event that the VBP(α/α) transcript shown in panel B of Figure 2 were to undergo the carboxyl-terminal splicing event shown in panel C of Figure 2, the resultant message would be predicted to encode a VBP(α/β) isoform; namely, an isoform with an amino-teminal end comparable to VBP(α/α) and a carboxyl-terminal end comparable to VBP(β/β). Conversely, the failure to undergo this splice in the VBP(β/β) transcript shown in panel C of Figure 2 would result in a message encoding a VBP(β/α) isoform. To examine whether mRNAs are expressed that encode these two other VBP isoforms, and to address whether any VBP mRNAs are differentially expressed, oligos were designed that enabled us to assay for each of the four possible isoforms in parallel RT–PCR reactions (see Materials and Methods). Figure 5 shows the results obtained using embryonic tissues and cultured cells (panel A) and adult tissues (panel B). The bottom strip in each panel is a control for the quality of the cDNAs used in this analysis and represents the PCR product obtained using primers that bind to exons 2 and 3, which are present in all four isoforms (see Figure 2). This positive control corroborates our previous conclusion that the VBP gene is ubiquitously expressed (13). However, the other strips in panels A and B reveal differences in expression among the mRNAs that encode different VBP isoforms. For example, as shown in panel A, chicken hepatoma (LMH) cells and chick embryo fibroblast (CEF) cells both contain detectable amounts of mRNAs for VBP(α/β) but CEF cells lack detectable amounts of mRNA for VBP(α/α), and neither cell type has detectable amounts of mRNAs for either VBP(β/α) or VBP(β/β). The latter two mRNAs are, however, evident in certain tissues, most notably in the adult heart and intestine (panel B). The restricted expression of VBP mRNA species containing the beta first exon sequence in comparison to the alpha first exon sequence suggests that the beta promoter may function in only a subset of cell types.

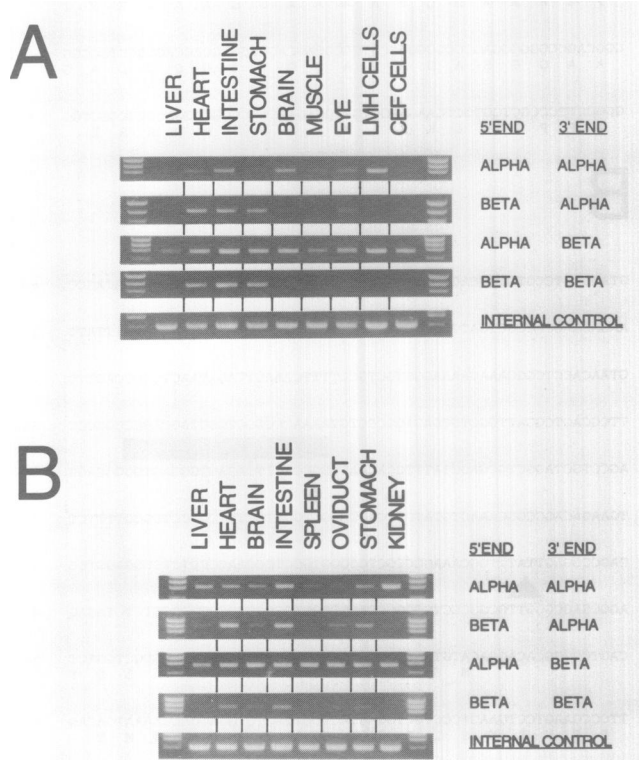

Figure 5. Evidence that mRNAs encoding four distinct VBP isoforms are differentially expressed. PCR protocols were devised that allowed us to detect cDNAs for all four VBP isoforms, as well as each unique VBP isoform (see Materials and Methods). These protocols were used to examine total cDNA samples prepared from various embryonic (**A**) or adult (**B**) tissues, as well as CEF and LMH cells grown in culture (A). When RT–PCR assays were carried out using a pair of primers that map to exons two and three, which are common to all four VBP isoforms (see Figure 2), a product of the expected size was obtained from each cDNA sample, as shown by the strips labeled 'internal control'. In contrast, mRNAs for the four VBP isoforms [VBP(α/α), VBP(β/α), VBP(α/β), and VBP(β/β)] are shown to be differentially expressed in these tissues and cell types. The virtual absence of particular PCR products in certain cDNA samples is noteworthy since the PCR parameters that were used (see Materials and Methods) would tend to overemphasize minor mRNA species.

Delineation of a core transactivation domain and cell type-specific modulating domains for VBP isoforms

In an attempt to gain insights regarding possible functional differences between the predicted VBP isoforms, we first tested whether all four isoforms are able to bind to a known VBP target sequence from the chicken vitellogenin II promoter (13). To this

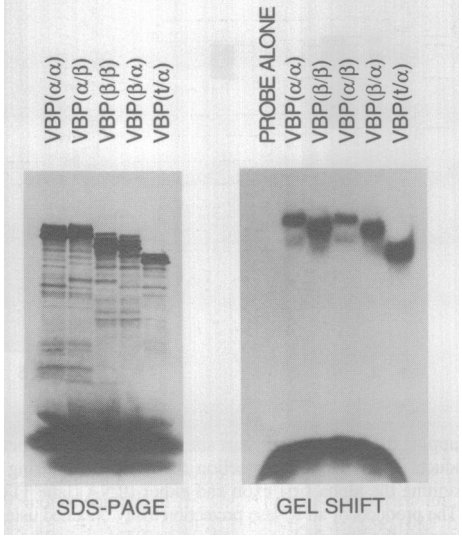

Figure 6. The four VBP isoforms can each bind to a known target site. The four VBP isoforms, as well as a truncated form of VBP(α/α) which spans aa 79–313 (see Figure 1) and is denoted VBP(t/α), were translated *in vitro* as labeled or unlabeled proteins and then analyzed on a denaturing protein gel (left panel) or in a gel shift assay (right panel), respectively. The probe used for the gel shift assay corresponds to the VBP binding site from the chicken vitellogenin II promoter (13).

end, the ORFs for each of the four VBP isoforms were cloned into an *in vitro* expression vector (20). As shown in panel A of Figure 6, the four isoforms, as well as a truncated form of VBP(α/α) that spans aa 79–313 (see Figure 1), were each translated efficiently. This autoradiogram also suggests that these VBP isoforms may be subject to post-translational modifications, since a collection of bands is evident above each of the major translation products. Parallel *in vitro* translation reactions were carried out using non-labeled amino acids for use in the gel shift assay shown in panel B of Figure 6. It is clear that each isoform is able to bind to this probe, although the amino-terminal alpha sequence may decrease binding to a modest degree.

A co-transfection assay that was carried out before we cloned the cDNA for VBP(β/β) revealed that an amino-terminal truncated (aa 79–313) form of VBP(α/α) was able to transactivate a reporter plasmid that contained eight VBP binding sites cloned upstream of a minimal TATA box promoter. An analogous (see Figure 1) amino-terminal truncated (aa 53–293) form of VBP(β/β) was cloned into the same expression vector and the two truncated forms, which we denote VBP(t/α) and VBP(t/β), were directly compared in co-transfection assays using both CEF cells and LMH cells. As shown in Figure 7, VBP(t/α) and VBP(t/β) displayed similar transactivation properties. These two carboxyl-terminal tails thus appear not to affect the transactivation properties of VBP, at least in these two cell types.

We also cloned cDNA fragments for the four VBP isoform ORFs into the same expression vector in order to examine whether the isoform-specific amino-terminal domains might modulate the transactivation properties of VBP. As shown in Figure 7, the four full-length VBP isoforms transactivated the target promoter less efficiently than VBP(t/α) and VBP(t/β) in CEF cells, whereas in LMH cells they functioned as well or better than VBP(t/α) and VBP(t/β). This result suggests that the activity of the VBP core transactivation domain, which is present in VBP(t/α) and VBP(t/β) (see below), may be modulated in a cell type-specific manner by the alpha and beta amino-terminal domains. However, it is also possible that these results may be due to differences in the steady-state levels of the truncated versus full-length VBP isoforms in these two cell types. We are currently

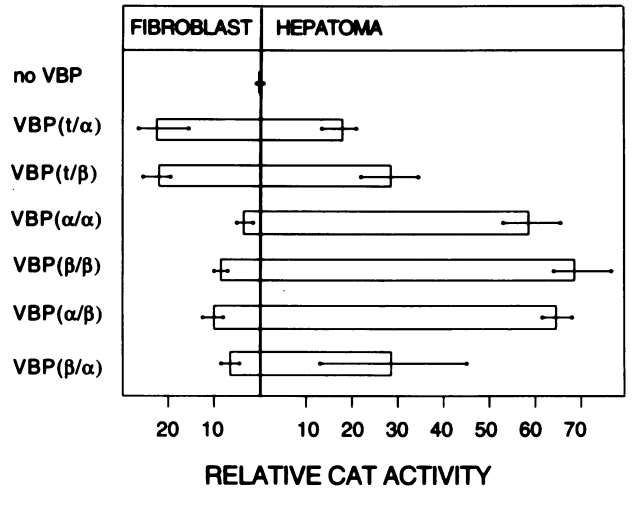

Figure 7. Differential transactivation by full-length VBP isoforms and truncated VBP proteins in CEF cells and LMH cells. Co-transfection assays were performed in CEF and LMH cells using RSV-driven VBP expression vectors and a reporter CAT plasmid bearing a bank of eight VBP binding sites located immediately upstream of a minimal (TATA only) promoter. In addition to four VBP isoforms, two truncated VBP proteins, denoted VBP(t/α) and VBP(t/β), were expressed using an internal methionine codon located in exon 2: VBP(t/α) spans aa 79–313 of the VBP(α/α) sequence and VBP(t/β) spans aa 53–293 of the VBP(β/β) sequence (see Figure 1). The range of values obtained in two independent sets of co-transfection assays are indicated in arbitrary CAT units. Note that, whereas the two truncated VBP proteins transactivate equivalently in either CEF cells (left portion of figure) or LMH cells (right portion of figure), the full-length VBP isoforms are relatively less efficient in transactivating the reporter plasmid in CEF cells but are relatively more efficient in transactivating the reporter plasmid in LMH cells.

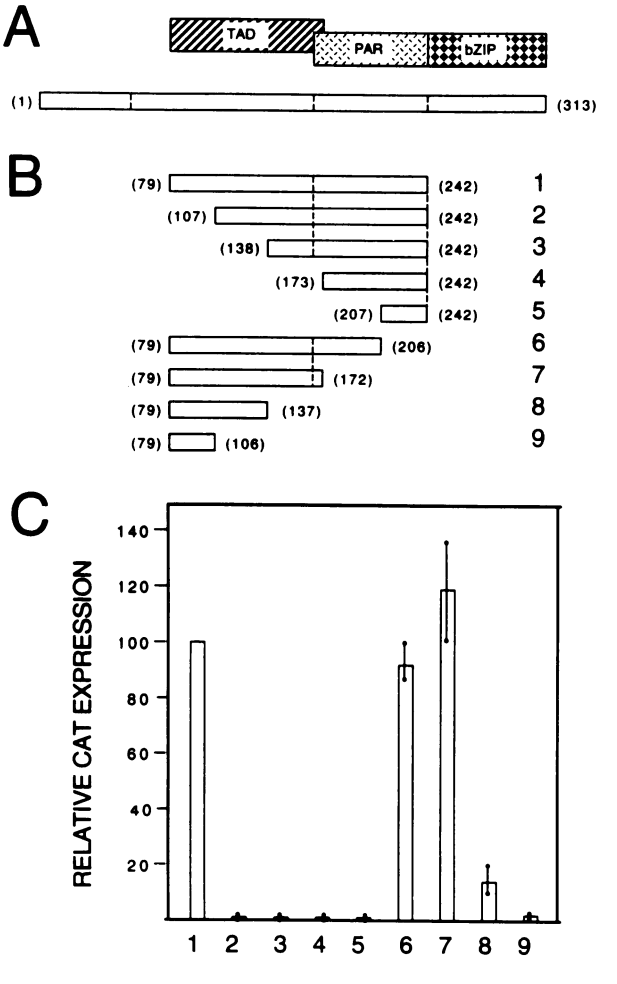

Figure 8. Delineation of a core transactivation domain for VBP. The positions of the exon–exon junctions for the 313 aa VBP(α/α) isoform are indicated by dotted lines in (**A**). Note that the PAR and bZIP domains map to exons 3 and 4, respectively. Based on the results shown in (**C**) (see below), it is clear that a core transactivation domain (TAD) maps primarily, if not exclusively, to exon two. The segments of VBP that were expressed as GAL4(1-147)/VBP fusion proteins (22) are shown in (**B**). The endpoints shown for each segment correspond to the numbering scheme presented in Figure 1 for VBP(α/α). Each of these GAL4(1-147)/VBP fusion proteins was tested for its ability to transactivate a minimal promoter bearing five GAL4 binding sites (22). The range of values obtained in two independent co-transfection assays are indicated for each GAL4(1-147)/VBP fusion protein. Note that these results are expressed in arbitrary CAT units, with 100 being defined as the CAT activity obtained using the parental GAL4-VBP(79-242) construct shown in lane 1. Less than one unit of CAT activity was obtained in co-transfection assays using the empty GAL(4-147) vector (data not shown).

attempting to raise antibodies agent VBP in order to resolve this issue.

In order to map the VBP core transactivation domain, we analyzed GAL4(1-147)/VBP fusion proteins for their abilities to transactivate a promoter containing five GAL4 binding sites cloned upstream of a TATA box (22). The parental GAL4(1-147)/VBP fusion protein used in these studies contains the amino acid 79−242 region of VBP(α/α) (see Figure 1) fused to the DNA binding domain of GAL4 (22). The choice of this particular amino-terminal endpoint was based on the observation that VBP(t/α) functioned in the transactivation assays shown in Figure 7. We chose not to include the bZIP region of VBP so as to avoid complications that might result from having this DNA binding domain acting in competition with the GAL4 DNA binding domain. The results obtained using this, and the other GAL4(1-147)/VBP fusion proteins indicated in panel B of Figure 8, are shown in panel C of Figure 8. It is clear from this analysis that the core VBP transactivation domain (denoted TAD in the summary shown in panel A) is encoded predominantly, if not exclusively, in exon two. This information is thus present in all four VBP isoforms.

DISCUSSION

Three members of the PAR subfamily of bZIP transcription factors (rat DBP, human HLF, and rat TEF) were previously shown to be differentially expressed. Rat DBP (16) and human HLF (15,17) are liver-enriched factors, with the former being expressed in response to circadian cues (29). Although rat TEF is ubiquitously expressed in adults, this factor appears to be selectively expressed in a subset of pituitary cells during embryogenesis (14). In apparent contrast to the tissue-specific expression of these factors, chicken VBP was previously thought to be ubiquitously expressed in both embryonic and adult tissues (13). While correct in its outlines, this picture of VBP must now be revised in light of the finding that four distinct VBP isoforms are encoded by differentially processed transcripts. Notwithstanding other complications that likely arise from dimeric complexes between chicken VBP and other PAR bZIP factors such as chicken HLF (manuscript in preparation), we infer from the RT−PCR analysis presented in Figure 5 that tissues differ with respect to their repertoires of homodimeric and heterodimeric VBP isoform complexes. Additional complexities are indicated by the fact that many target sites which are recognized by PAR bZIP factors are also recognized by members of the C/EBP multigene family of bZIP factors (16,30). PAR sites are thus potential targets for a diverse range of transcriptional regulatory influences.

Given that most transcription factors have modular structures (31,32), it is not surprising that alternative splicing pathways have been exploited as a means of generating functionally distinct transcription factor isoforms. Simple mutations would presumably suffice to generate regulatable splice sites, since modest deviations from consensus splice sites are sufficient to allow alternative splicing to occur (33). In contrast, many genetic changes are required to create differentially expressed first exons. Presumably for this reason, there have been only a few reports of transcription factor isoforms being generated by alternative first exon usage. Since the coding segments of the two first exons of VBP are related (see Figures 1 and 4), we presume that these alternative first exons are attributable to an exon duplication event followed by divergent evolution. The fact that the promoter regions flanking these two first exons appear unrelated (see Figure 4) further suggests that this presumptive exon duplication was an ancient event. It is interesting to contemplate whether the presumptive rat homolog of VBP (TEF) might also be expressed from alternative first exons, with one of the first exon promoters being specific for thyrotrope cells of the developing pituitary and the other being ubiquitously expressed in adult tissues.

The other splicing option available to VBP transcripts is also not trivial, since the use of the more distal of two possible polyadenylation sites is a necessary, but not sufficient, condition that must be satisfied in order for this splice to occur (see panels C and B of Figure 2). In some respects this regulation resembles the switch from the membrane form to the secreted form of the immunoglogulin μ heavy chain (33). We infer that, just as in the case of the regulated splicing of the μ heavy chain transcript, the suboptimal nature of the two relevant VBP splice sequences are critical in allowing this event to be regulated. However, in contrast to the situation with the immunoglogulin μ heavy chain transcript, this VBP splicing event does not appear to occur in strict competition with the usage of the proximal polyadenylation site. As a parenthetical note, the fact that this suboptimal splice donor is conserved between chicken VBP and rat TEF raises the possibility that rat TEF transcripts may be alternatively processed in similar fashion.

The finding that all four predicted isoforms contain a common core transactivation domain that functions efficiently when fused to the DNA binding domain of GAL4(1-147) indicates that each isoform has the potential to function as a positive-acting transcription factor. However, this potential appears to be either restricted or enhanced in a cell type-specific manner by the alternative amino-terminal domains that flank the core transactivation domain. Whereas the alpha and beta amino-terminal domains both inhibit the core transactivation domain in fibroblasts and both augment the activity of the core transactivation domain in hepatoma cells, the fact that the alpha domain has a unique alanine-rich region as well as a region that is similar to the beta domain suggests that these domains may function discordantly in other cell types. Future experiments will be required to address this issue and to determine whether the amino-terminal domains of VBP are targets for cell type-specific proteins that serve to modulate the activity of the flanking core transactivation domain akin to what has been shown for the δ and ϵ modulating domains that flank a core transactivation domain of c-Jun (34).

The functional significance of the alternative carboxyl-terminal domains of VBP is presently obscure. The fact that all four VBP isoforms are able to homodimerize and bind to a known VBP target sequence (Figure 6) is not surprising considering that the carboxyl-terminal differences map outside the basic and leucine zipper regions. We have also been unable to demonstrate any biases toward the formation of particular VBP isoform heterodimers (data not shown), although the possibility exists that particular VBP isoforms may preferentially form heterodimers with other members of the PAR subfamily of bZIP factors. Similarly, although these alternative carboxyl-terminal tails do not grossly affect the ability of VBP isoforms to transactivate a simple reporter construct, functional differences may become evident when more complex target promoters are analyzed. These issues will be the focus of future experiments.

ACKNOWLEDGEMENTS

We thank Subramanian Iyer for constructing some of the expression vectors used in this study; Leonard Cohen, Charles Emerson, and Robert Perry for critically reading the manuscript; Glen Miller (Oligonucleotide Synthesis Facility, Institute for Cancer Research) for oligonucleotides; and Michelle Slook, Mary Ann Newhouse, and Sarah Costello (Secretarial Services, Institute for Cancer Research) for help in preparing the manuscript. This work was supported by NIH grants 35535 and CA-06927, and by an appropriation by the Commonwealth of Pennsylvania.

REFERENCES

1. Johnson, P.F. and McKnight, S.L. (1989) *Ann. Rev. Biochem.*, **58**, 799–839.
2. Mitchell, P.J. and Tjian, R. (1989) *Science*, **245**, 371–291.
3. Pabo, C.O. and Sauer, R.T. (1992) *Ann. Rev. Biochem.*, **61**, 1053–1095.
4. Schaffner, W. (1989) *TIG*, **5**, 37–39.
5. Foulkes, N.S. and Sassone-Corsi, P. (1992) *Cell*, **68**, 411–414.
6. Delmas, V., Laoide, B.M., Masquilier, D., de Groot, R.P., Foulkes, N. and Sassone-Corsi, P. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 4226–4230.
7. Descombes, P. and Schibler, U. (1991) *Cell*, **67**, 569–579.
8. Foulkes, N.S., Mellstrom, B., Benusiglio, E. and Sassone-Corsi, P. (1992) *Nature*, **355**, 80–84.
9. Mumberg, D., Lucibello, F.C., Schuermann, M. and Muller, R. (1991) *Genes Dev.*, **5**, 1212–1223.
10. Koenig, R.J., Lazar, M.A., Hodin, R.A., Brent, G.A., Larsen, P.R., Chin, W.W. and Moore, D.D. (1989) *Nature*, **337**, 659–661.
11. Yamamoto, K.K., Gonzalez, G.A., Menzel, P., Rivier, J. and Montminy, M.R. (1990) *Cell*, **60**, 611–617.
12. Weber, B.L., Westin, E.H. and Clarke, M.F. (1990) *Science*, **249**, 1291–1293.
13. Iyer, S.V., Davis, D.L., Seal, S.N. and Burch, J.B.E. (1991) *Mol. Cell. Biol.*, **11**, 4863–4875.
14. Drolet, D.W., Scully, K.M., Simmons, D.M., Wegner, M., Chu, K., Swanson, L.W. and Rosenfeld, M.G. (1991) *Genes Dev.*, **5**, 1739–1753.
15. Hunger, S.P., Ohyashiki, K., Toyoma, K. and Cleary, M.L. (1992) *Genes Dev.*, **6**, 1608–1620.
16. Mueller, C.R., Maire, P. and Schibler, U. (1990) *Cell*, **61**, 279–291.
17. Inaba, T., Roberts, W.M., Shapiro, L.H., Jolly, K.W., Raimondi, S.C., Smith, S.D. and Look, A.T. (1992) *Science*, **257**, 531–534.
18. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor University Press, Cold Spring Harbor.
19. Chen, E.Y. and Seeburg, P.H. (1985) *DNA*, **4**,165–170.
20. Beckmann, H., Su, L.-K. and Kadesch, T. (1990) *Genes Dev.*, **4**, 167–179.
21. Pei, D. and Shih, C. (1990) *J. Virol.*, **64**, 1517–1522.
22. Lillie, J.W. and Green, M.R. (1989) *Nature*, **338**, 39–44.
23. Kawaguchi, T., Nomura, K., Hirayama, Y. and Kitagawa, T. (1987) *Cancer Res.*, **47**, 4460–4464.
24. Binder, R., MacDonald, C.C., Burch, J.B.E., Lazier, C.B. and Williams, D.L. (1990) *Mol. Endocrinol.*, **4**, 201–208.
25. Graham, F.L. and van der Eb, A.J. (1973) *Virology*, **52**, 456–467.
26. Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
27. Bradford, M. (1976) *Anal. Biochem.*, **72**, 248–254.
28. Kadanoga, J.T., Jones, K.A. and Tijian, R. (1986) *Trends Biochem. Sci.*, **11**, 20–23.
29. Wuarin, J. and Schibler, U. (1990) *Cell*, **63**, 1257–1266.
30. Landschulz, W.H., Johnson, P.F., Adashi, E.Y., Graves, B.J. and McKnight, S.L. (1988) *Genes Dev.*, **2**, 786–800.
31. Brent, R. and Ptashne, M. (1985) *Cell*, **43**, 729–736.
32. Frankel, A.D. and Kim, P.S. (1991) *Cell*, **65**, 717–719.
33. Peterson, M.L. and Perry, R.P. (1989) *Mol. Cell. Biol.*, **9**, 726–738.
34. Baichwal, V.R., Park, A. and Tijian, R. (1992) *Genes Dev.*, **6**, 1493–1502.
35. Kozak, M. (1984) *Nucleic Acids Res.*, **12**, 857–872.