MOLECULAR AND CELLULAR BIOLOGY, Oct. 1991, p. 5044–5052
0270-7306/91/105044-09$02.00/0

Vol. 11, No. 10

C/EBP-Like Proteins Binding to the Functional Box-α and Box-β of the Second Enhancer of Hepatitis B Virus

CHIOU-HWA YUH AND LING-PAI TING*

Graduate Institute of Microbiology and Immunology, National Yang-Ming Medical College, Shih-Pai, Taipei 11221, Taiwan, Republic of China

Received 11 April 1991/Accepted 11 July 1991

The second enhancer (enhancer II) of hepatitis B virus is functionally liver specific. Located within an open reading frame of the virus and immediately upstream of the initiation sites of viral major transcripts, enhancer II furnishes a unique model for use in investigating the structure and function of an enhancer. In this study, two functional constituents, a 23-bp box-α and a 12-bp box-β, are identified as being both necessary and sufficient for enhancer II function. Examination of the box-α and box-β sequences reveals a weak homology to the extended consensus for a C/EBP binding site. Gel shift and footprinting analyses indicate that multiple proteins bind to these sequences and thus are candidate transcription factors that mediate the enhancer function. One heat-resistant protein, protein a, and one heat-sensitive protein, protein b, bind to box-α. Protein a, which binds to box-α in a way indistinguishable from that seen with a recombinant C/EBP, appears not to be identical to C/EBP in that the binding of protein a requires a minimal sequence larger than the canonical C/EBP sites. Two box-β-binding proteins, c and d, show greater affinity for the C/EBP consensus than for box-β. However, both proteins c and d are relatively heat sensitive and display a distinct sequence preference from the recombinant C/EBP protein. Since the function of enhancer II is strictly dependent on a bipartite architecture, this system provides a unique model for studies of how the interactions of its binding proteins lead to the enhancer function.

Control of gene transcription is in part regulated by the presence of *cis*-acting DNA elements that interact with specific nuclear transcription factors. Enhancers, which act in a position- and orientation-independent manner, are key regulatory elements in the transcriptional regulation of viral and cellular genes (18, 26). The 72-bp repeat of simian virus 40 (SV40) is the best-characterized enhancer. The SV40 72-bp enhancer, which is composed of multiple regulatory *cis*-acting elements via the interaction with ubiquitous and cell-type-specific transcriptional factors, orchestrates the expression of viral genes in many cells (13, 28–30). We are interested in the study of gene regulation in hepatitis B virus (HBV), in particular its enhancer elements. HBV infection is one of the leading causes of acute and chronic hepatitis. The latter has been associated with liver cirrhosis and hepatocellular carcinoma (4). A better understanding of the molecular biology of HBV will provide deeper insight into its life cycle and the myriad disease processes in which it is involved.

The 3.2-kb HBV genome is one of the smallest known for animal viruses. The genome organization is very compact, with four overlapping open reading frames coding for the surface, core, polymerase, and X proteins (6, 14). Their transcription is under the control of four promoters: two for surface, one for core and polymerase, and one for X. A liver-specific enhancer element has been located between the surface and X open reading frames, mapping from nucleotides (nt) 1074 to 1234 (8, 35).

A second, liver-specific enhancer (enhancer II) within the X open reading frame of HBV was recently identified (36, 38, 39). We have mapped this enhancer to nt 1636 to 1741. Upon further dissection, this enhancer was found to comprise two interacting elements, II-A (nt 1636 to 1690) and II-B (nt 1704 to 1741), whose cooperation is indispensable for enhancer function in vivo (39). In this study, fine mapping of the enhancer region was analyzed with serial deletion mutants. We have identified a 23-bp sequence (designated box-α) in the II-A element and a 12-bp sequence (designated box-β) in the II-B element, the interaction of which is both necessary and sufficient for enhancer II activity. Furthermore, proteins that display binding activities to box-α and box-β are found to be present in nuclear extracts of the differentiated human hepatoma cell line HepG2 by DNase I footprinting and gel shift analyses. It is likely that some of these binding proteins are C/EBP-like proteins, although they are not identical to C/EBP.

MATERIALS AND METHODS

Plasmid constructions. The plasmid with a 3′ deletion of the enhancer II region was generated by BAL31 exonuclease digestion and confirmed by double-strand DNA sequencing. Fragments with serial 3′ deletions were inserted downstream of the chloramphenicol acetyltransferase (CAT) gene of plasmid pA3SVpCAT (39) to generate a series of plasmids named pA3SVpCATENdl. To construct plasmids containing serial 3′ deletions in II-A but intact II-B elements, the fragment containing the II-B sequence (nt 1704 to 1741) was added downstream of the enhancer II sequence in the pA3SVpCATENdl series of plasmids, which have lost II-B in its entirety, to make a series of plasmids designated pA3SVpCATAdlB. Plasmids containing the intact box-α (nt 1636 to 1668) but the sequential deletions from the 3′ end of II-B were constructed by replacement of the II-A sequence (nt 1636 to 1703) with the box-α sequence (from nt 1645 to 1669). These plasmids are designated pA3SVpCATABdl.

To create mutations in the box-α sequence, oligonucleotides of the following sequences (coding strand on top, with underlined nucleotides indicating nt 1645 and 1669, respectively, of HBV, and with the mutated nucleotides in lower-

* Corresponding author.

case letters) were synthesized:

Wild type: 5′-ATCGATCAAGGTCTTACATAAGAGGACTCTT-3′
3′-TAGTTCCAGAATGTATTCTCCTGAGAA-5′
AB: 5′-ATCGATCAAGGTCTTACATccGctGACTCTT-3′
3′-TAGTTCCAGAATGTAggCgaCTGAGAA-5′
CD: 5′-ATCGATCAAGGgCggACATAAGAGGACTCTT-3′
3′-TAGTTCCcGccTGTATTCTCCTGAGAA-5′
EF: 5′-ATCGATCActGTCTTACATAAGAGGACTCTT-3′
3′-TAGTgaCAGAATGTATTCTCCTGAGAA-5′

Double strands were annealed by incubation in a boiling water bath for 2 min and then at 65°C for 10 min, 37°C for 10 min, and room temperature for 10 min. The annealed double-stranded oligonucleotides were inserted between the *Bam*HI and blunt-ended *Pst*I sites in the upstream position of the II-B sequence of plasmid pA3SVpCAT/II-B.

Cell lines, transfections, and CAT assays. The human hepatoma cell line HepG2, which maintains the characteristics of differentiated liver cells, was cultured as previously described (1, 8). Cells were transfected with plasmids containing the CAT gene by the calcium phosphate precipitation method. CAT assays were performed according to Gorman et al. (15), with a slight modification as previously described (8). The protein amount of all samples in each experiment varied by no more than 5%. The CAT activity of each sample was normalized to the CAT activity of pSV_2CAT, which was taken as 100%. If CAT activity was high, the cell lysate was serially diluted prior to performance of the CAT assay.

Preparation of nuclear extracts and heparin-agarose fractionation of extracts. Nuclear extracts from the differentiated human hepatoma cell line HepG2 were prepared as previously described (9). Extracts were stored in small aliquots at −70°C after being quickly frozen under liquid nitrogen. The nuclear extract was fractionated at 4°C essentially as described by Lichtsteiner et al. (24). Briefly, 25 mg of crude nuclear extract in 10 ml of nuclear dialysis buffer (NDB; 20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES]-KOH [pH 7.9], + 0.1 M KCl, 0.2 mM EDTA, 0.5 mM dithiothreitol, 20% glycerol, 0.5 mM phenylmethylsulfonyl fluoride) was loaded onto a 20-ml heparin-agarose column (heparin Ultrogel-A4R; 4 to 6 mg of heparin per ml; IBF, Garenne, France). The proteins bound to heparin were collected by step elution using NDB containing 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, or 1.0 M NaCl. The fractions obtained were precipitated with 50% $(NH_4)SO_4$, resuspended in NDB, and then dialyzed against two changes of 500 ml of NDB, each for 3 h. They were aliquoted, quickly frozen under liquid nitrogen, and kept frozen at −70°C. For experiments in which nuclear extracts were subjected to a thermolability test, the 0.4 and 0.3 M NaCl fractions were heated (50°C for 5 min, 70°C for 5 min, or 90°C for 15 min) and then centrifuged at 12,000 rpm and 4°C for 10 min. The supernatants were collected for the DNase I footprinting and gel shift analyses.

DNase I footprinting and gel shift analyses. The end-labeled probes were generated via fill-in with [α-^{32}P]dATP (3,000 Ci/mmol; Amersham, Buckinghamshire, England) at either terminus of a *Bam*HI-*Hin*dIII fragment containing the 340-bp HBV sequence from nt 1402 to 1741 and the 198-bp SV40 early promoter. Plasmids carrying this insert were cut with one enzyme, end labeled, cut with the other enzyme, and then subjected to agarose gel electrophoresis. The labeled fragments were eluted from the agarose gel by a Spin-X filtering unit (Costar, Cambridge, Mass.). The DNase I footprinting assay was performed as previously described (9). Briefly, the crude or fractionated nuclear extracts were incubated with labeled DNA fragment for 30 min at 4°C. After addition of $MgCl_2$ and $CaCl_2$, the proper amount of DNase I was added to digest for 90 s at room temperature. Reactions were stopped by deproteination. The extracted DNA was analyzed on 4% polyacrylamide—8 M urea gels.

Annealed double-stranded oligomers (200 ng) corresponding to the box-α sequence (underlined nucleotides indicate nt 1645 and 1669, respectively, of the HBV sequence),

5′-GATCCATCGATCAAGGTCTTACATAAGAGGACTCTT-3′
3′-GTAGCTAGTTCCAGAATGTATTCTCCTGAGAA-5′

or the box-β sequence (underlined nucleotides indicate nt 1705 and 1721, respectively, of HBV),

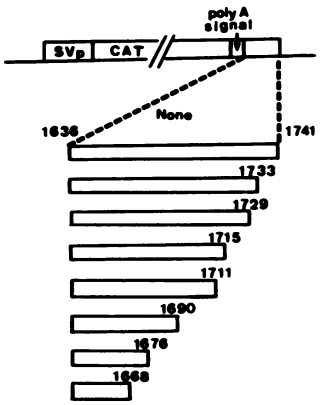
5′-GCTACTTCAAAGACTGTGGTACCCATG-3′
3′-ACGTCGATGAAGTTTCTGACACCATGG-5′

were end labeled with 80 U of polynucleotide kinase (Toyobo, Osaka, Japan) and 80 μCi of [γ-^{32}P]ATP (5,000 Ci/mmol; Amersham) and separated from the free [γ-^{32}P]ATP by an Ultrafree-C3 filter unit (UFC3 LGC NB; Millipore, Bedford, Mass.).

Proteins (10 to 20 μg) were incubated in a 20-μl reaction mixture containing 1 μg of poly(dI-dC) · poly(dI-dC) (Pharmacia, Inc.), 17 mM HEPES (pH 7.9), 12.5% glycerol, 0.42 mM EDTA, 0.3 mM dithiothreitol, 10 mM KCl, 6.25 mM $MgCl_2$, and 6×10^4 cpm of labeled DNA. Reaction mixtures were incubated at 30°C for 30 min, and protein-DNA complexes were resolved on 4% polyacrylamide (acrylamide/bisacrylamide weight ratio, 30:1) made in 1× TBE (90 mM Tris-borate [pH 8.0], 2 mM EDTA). Electrophoresis was performed at 150 V for 2.5 h at 4°C. Gels were dried and autoradiographed. For competition experiments, the 5-, 25-, 125-, and 625-fold molar excesses of unlabeled double-stranded oligonucleotides were preincubated with nuclear extracts on ice for 5 min before addition of probe. The double-stranded HNF-1 oligomer (5′-CAAACTGTCAAAT ATTAACTAAAGGGAG-3′ and 3′-GTTTGACAGTTTATA ATTGATTTCCCTC-5′), previously described (9), was used as the nonspecific competitor. The double-stranded C/EBP consensus sequence oligomer (5′-GATCGAACATATTGCG CAATACATTTCCCAAGT-3′ and 3′-CTTGTATAACGCG TTATGTAAAGGGTTCA-5′) was kindly provided by S. C. Lee (National Taiwan University, Taipei). Recombinant rat C/EBP was kindly provided by Steven L. McKnight (Carnegie Institution of Washington, Baltimore, Md.).

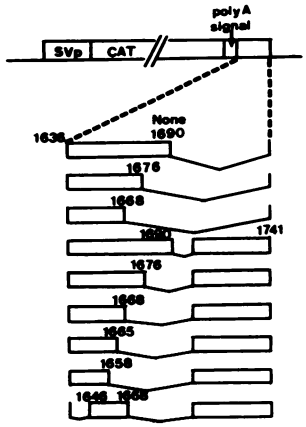
RESULTS

Identification of sequences required for the activating function of enhancer II. HBV enhancer II has been previously localized to nt 1636 to 1741. This 106-bp sequence, when placed downstream of a CAT reporter gene, strongly stimulates expression of the CAT gene driven by a heterologous SV40 early promoter (39). To delineate the minimal sequence element required for this activation, a series of plasmids that contained sequential deletions from the 3′ end of this enhancer element was derived. As shown in Fig. 1, removal of the sequence from nt 1741 to 1716 did not cause any significant change in enhancer activity (from 62-fold to 57-fold). However, a deletion of four more base pairs from nt 1715 to 1712 resulted in an almost complete loss of enhancer



CAT activity (%)	Stimulating fold
1.5	1.0
93.2	62.1
92.5	61.6
70.7	47.0
85.6	57.0
4.6	3.0
3.7	2.4
2.4	1.6
3.9	2.5

FIG. 1. Determination of the 3′ boundary of enhancer II in differentiated human hepatoma HepG2 cells. A schematic diagram of different fragments containing sequential deletions from the 3′ end of enhancer II in the pA3SVpCATENdl series of plasmids is shown on the left; normalized CAT activities and fold stimulations are summarized on the right. The enhancer II sequences are represented by open boxes, with numerals indicating nucleotides in the HBV genome (the *Eco*RI site is nt 1). SVp, SV40 early promoter; poly A signal, SV40 polyadenylation signal. Plasmids were transfected into HepG2 cells, and a CAT assay was performed for 1 h at 37°C. Plasmid pSV_2CAT was used as a control (set at 100%) against which CAT activities of tested plasmids were normalized. The cell lysate of each sample was diluted in a series to quantitatively calculate CAT activities. The values are the means of four experiments, with a standard deviation of 5%.



CAT activity (%)	Stimulating fold
1.5	1.0
3.7	2.5
2.4	1.6
3.9	2.6
45.8	30.6
42.3	28.2
46.2	30.8
1.6	1.1
1.3	0.9
46.0	30.7

FIG. 2. Delineation of the 3′ and 5′ boundary of the enhancer II-A element in HepG2 cells. The enhancer II sequences with sequential deletions in two series of plasmids, pA3SVpCATENdl and pA3SVpCATAdlB, are shown on the left; normalized CAT activities and fold stimulations are summarized on the right. Open boxes represent the sequences contained; thin lines indicate the deleted sequences. Experimental details and symbols are as described in the legend to Fig. 1.

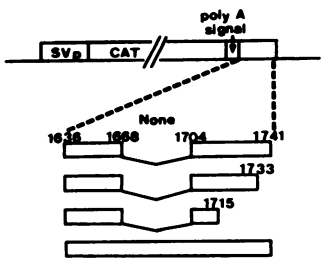
function (from 57-fold to 3-fold). No additional change was seen when more sequences were removed. This result unequivocally places the 3′ boundary of enhancer II between nt 1715 and 1711.

To further refine the minimal sequence requirement for enhancer activity in both elements II-A and II-B, similar strategies were applied. For II-A, plasmids containing serial deletions starting from the 3′ end of this 55-bp element (nt 1636 to 1690) were tested for the activation function. As shown in Fig. 2, neither the 55-bp element II-A nor any of its deletions had any significant enhancer activity by itself. In contrast, via cooperation with element II-B, the 55-bp element II-A activated transcription of the SV40 early promoter by 30-fold, which is consistent with previous results (39). Removal of the 22-bp sequence from nt 1690 to 1669 in element II-A did not alter the stimulating activity. This enhancer function was totally abolished, however, when three more nucleotides (to nt 1665) were left out. This identifies II-A as a positive regulatory element and marks its 3′ boundary at nt 1668 to 1665. In addition, a deletion of 10 more nucleotides from nt 1636 to 1645 did not result in any significant change, indicating that the 5′ boundary is not distal to nt 1646. Taken together, the data indicate that a 23-bp sequence (nt 1646 to 1668) is sufficient to account for the activation function of element II-A.

For II-B, a series of 3′-deletion mutants of the 38-bp II-B sequence (nt 1704 to 1741) in conjunction with the 33-bp (nt 1636 to 1668) sequence of II-A were placed downstream of the CAT gene and tested for transcriptional activation. As shown in Fig. 3, the 12-bp sequence from nt 1704 to 1715 in element II-B was able to cooperate with the 23-bp sequence of element II-A to achieve the expected level of transcription activation. Therefore, the coexistence of a 23-bp sequence (nt 1646 to 1668) of II-A, designated box-α, and a 12-bp sequence (nt 1704 to 1715) of II-B, designated box-β, is sufficient for the enhancer II activity.

Footprinting analysis of nuclear proteins binding to enhancer II. To search for a candidate transcription factor(s) that regulates the function of enhancer II, DNase I footprinting analysis with crude nuclear extracts from HepG2 cells was performed. A 538-bp DNA fragment containing the 340-bp HBV sequence (from nt 1402 to 1741) and the 198-bp SV40 early promoter was end labeled on either the coding or noncoding strand. Four major protected regions, I to IV, were seen on both strands (Fig. 4A). A protected region, IV, was identified as coming from the SV40 early promoter. The other three major protected regions, I to III, each bracketed by hypersensitive sites, were from the HBV-specific region. Regions II and III coincided with box-α and box-β, respectively, which are the minimal essential sequences for enhancer II function.

A fractionation experiment was initiated to examine the factors that bind to these regions. Crude nuclear extracts from HepG2 cells were loaded on a heparin-agarose column and step eluted with increasing concentrations of NaCl. Each fraction was then collected and tested for binding to



CAT activity (%)	Stimulating fold
1.4	1.0
101.0	73.6
40.0	28.4
41.8	29.4
92.5	65.1

FIG. 3. Delineation of the 3′ boundary of the enhancer II-B element. A schematic representation of enhancer II sequences in the pA3SVpCATABdl series of plasmids is shown on the left; normalized CAT activities and fold stimulations are summarized on the right. Experimental details and symbols are as described in the legend to Fig. 1.

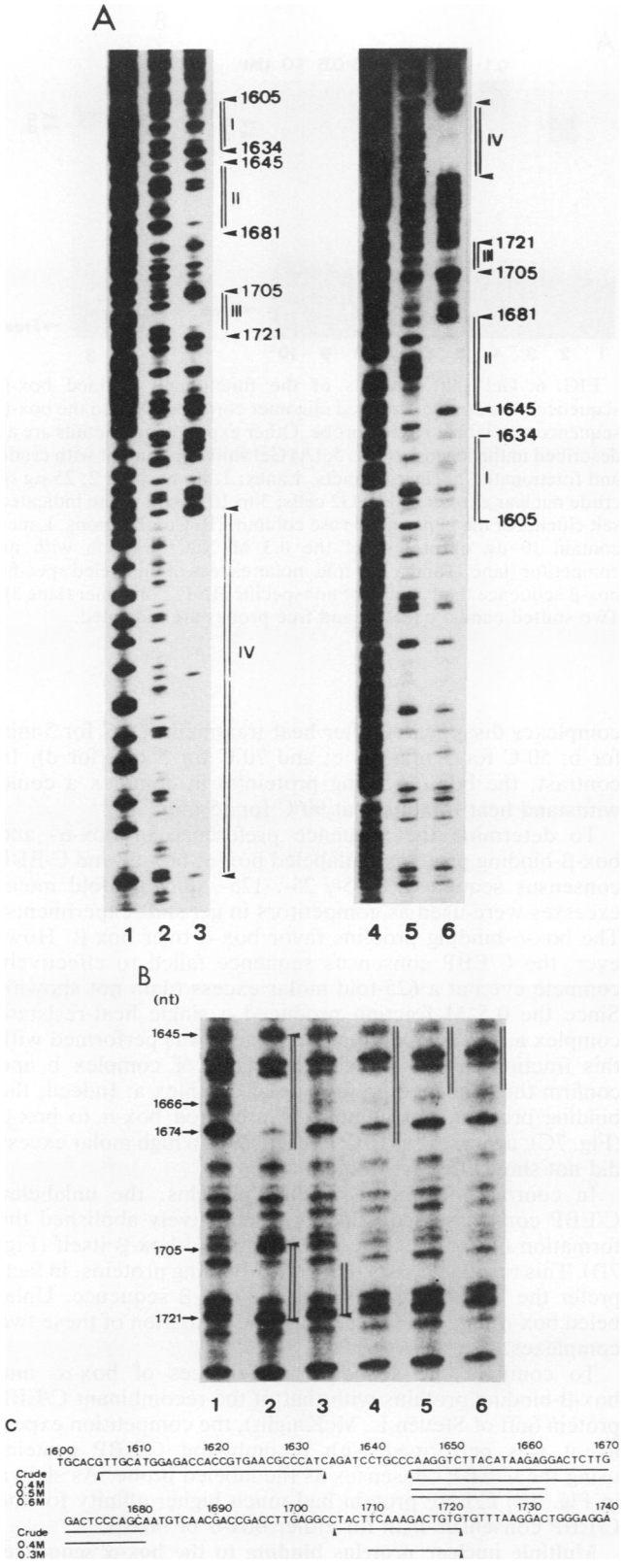


FIG. 4. DNase I footprinting analysis of the enhancer II region. The footprinting analysis was performed by using a fragment containing the HBV (nt 1402 to 1741) and SV40 early promoter sequences. The coding and noncoding strands were end labeled with the Klenow fragment of *Escherichia coli* DNA polymerase. The crude nuclear extract and partially purified fractions of a heparin-agarose column were prepared from HepG2 cells as described in Materials and Methods. The protected regions are shown on the right as two lines with numerals referring to the nucleotide number of HBV. The arrows show the DNase I-hypersensitive sites. Protected regions I, II, and III are in the enhancer II sequence of HBV, while region IV is in the SV40 early promoter. (A) DNase I footprint with crude nuclear extracts of HepG2 cells on the coding (lanes 1 to 3) and noncoding (lanes 4 to 6) strands. Lanes: 1 and 4, G+A sequence markers; 2 and 5, no protein; 3 and 6, 120 μg of protein from crude nuclear extracts. (B) DNase I footprint with crude and 0.3 to 0.6 M NaCl fractions of nuclear extracts on the coding strand. Lanes: 1, no protein; 2, 120 μg of crude nuclear extracts; 3 to 6, 40 μg each of 0.3, 0.4, 0.5, and 0.6 M NaCl fractions of the heparin-agarose column. (C) Summary of the protected sequences in HBV enhancer II. Lines indicate the sequences protected by crude and fractionated nuclear extracts; arrowheads indicate the DNase I-hypersensitive sites.

the HBV enhancer II sequence. As shown in Fig. 4B, region II (nt 1645 to 1681), covering box-α (nt 1646 to 1668) in its entirety, was mainly protected by a factor(s) that eluted at 0.4 M NaCl. A similar yet distinct footprint (nt 1645 to 1662) was seen with 0.5 and 0.6 M fractions. Footprinting to region III (nt 1707 to 1722), corresponding to the box-β sequence (nt 1704 to 1715), was eluted earlier in the 0.3 M NaCl fraction. Other fractions did not show clear footprinting activity to regions II and III (data not shown). These results are summarized in Fig. 4C.

Gel shift analysis of the box-α and box-β-binding proteins. To study the nature of binding activities to box-α and box-β, gel shift assays were performed. When a double-stranded oligonucleotide (nt 1645 to 1669) corresponding to box-α was used as a probe in the binding reaction, formation of two DNA-protein complexes, a and b, was noted with the fraction eluted at 0.4 M NaCl (Fig. 5A). These complexes appeared to be specific in that they could be abolished with unlabeled box-α sequence but not by nonspecific HNF-1 oligomer (Fig. 5B). The binding proteins of complexes a and b are tentatively named proteins a and b, respectively. Fractions collected at 0.5 and 0.6 M NaCl fractions produced a single band that migrated to the same position as complex a. This complex was also competable only with the box-α sequence, not with the HNF-1 oligomer. The presence of shifting activity that produced complex b in the 0.4 M but not in the 0.5 or 0.6 M fraction may well explain the larger footprint seen with the former than the latter (Fig. 4B). We also observed complex formation with the 0.1 M NaCl eluent, but the shifting is more likely to be nonspecific since it could be abolished with the HNF-1 oligomer.

When similar experiments were done for double-stranded oligonucleotides (nt 1705 to 1721) corresponding to the box-β sequence, two complexes designated as c and d were detected with the 0.3 M NaCl fraction (Fig. 6A). Both complex c and complex d appeared to be specific, since they could not be abolished by the box-β sequence but not by the HNF-1 oligomer (Fig. 6B). The binding proteins of complexes c and d are tentatively designated proteins c and d, respectively.

The detection of footprints colocalized with box-α and box-β is consistent with the notion that these sequences are essential for enhancer II function and their effects are mediated via the interaction with some *trans*-acting factor(s). The formation of specific DNA-protein complexes with box-α and box-β sequences, as revealed in the gel shift assay, supports this point. Examination of the box-α (CAAGGTCTTACATAAGAGGACTCTT, from nt 1645 to

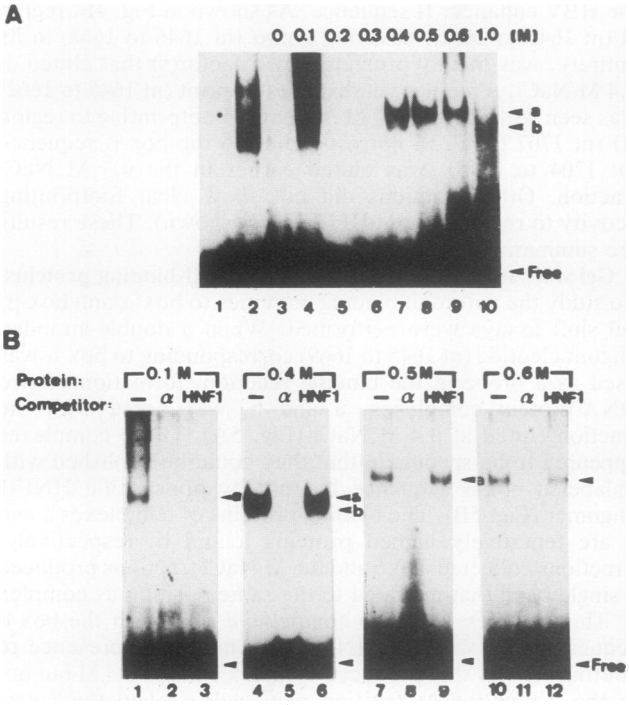


FIG. 5. Gel shift analysis of the functionally defined box-α sequence. The double-stranded oligomer corresponding to the box-α sequence was end labeled with polynucleotide kinase and incubated with proteins of crude and fractionated nuclear extracts of HepG2 cells at 30°C for 30 min. After incubation, the DNA-protein complexes were resolved on a native 4% polyacrylamide gel. (A) Gel shift experiment with crude and fractionated nuclear extracts. Lanes: 1, no protein; 2, 25 μg of crude nuclear extracts; 3 to 10, 10 μg of the indicated salt eluents of the heparin-agarose column. (B) Competitions. Lanes contain 10 μg of the indicated salt eluents with no competitor (−), a 125-fold molar excess of unlabeled specific box-α oligomer (α), or a 125-fold molar excess of unlabeled nonspecific HNF-1 oligomer (HNF1). Two specifically shifting bands, a and b, are indicated by arrows. Free denotes uncomplexed probe.

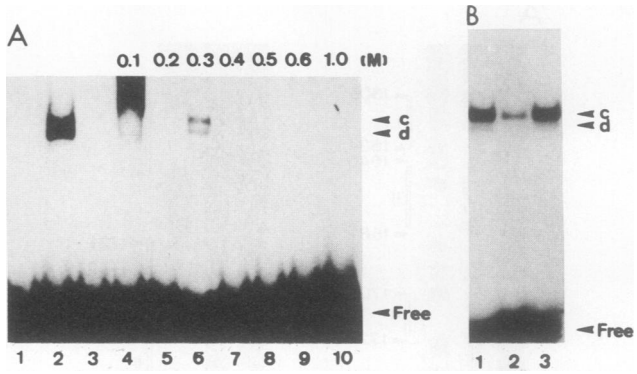


FIG. 6. Gel shift analysis of the functionally defined box-β sequence. The double-stranded oligomer corresponding to the box-β sequence was labeled as the probe. Other experimental details are as described in the legend to Fig. 5. (A) Gel shift experiment with crude and fractionated nuclear extracts. Lanes: 1, no protein; 2, 25 μg of crude nuclear extract of HepG2 cells; 3 to 10, 10 μg of the indicated salt eluents of the heparin-agarose column. (B) Competitions. Lanes contain 10 μg of protein of the 0.3 M NaCl fraction with no competitor (lane 1) and a 125-fold molar excess of unlabeled specific box-β sequence (lane 2) and of nonspecific HNF-1 oligomer (lane 3). Two shifted bands, c and d, and free probes are indicated.

1669) and the box-β (CCTACTTCAAAGACTGT, from nt 1704 to 1720) sequences reveals a very weak homology (underlined) to the consensus sequence ATTGCGCAAT of the binding site for C/EBP (CAAT/enhancer-binding protein), a transcription factor found abundantly in liver (5, 19–22).

Characteristics of box-α and box-β-binding proteins. The resemblance of the sequences of both box-α and box-β to the consensus of the C/EBP binding site raises the intriguing possibility that C/EBP or another member of the C/EBP multigene family such as DBP, Ig/EBP-1, or LAP (NF-IL6, IL-6DBP, and AGP/EBP designate the same protein) (3, 7, 11, 20, 27, 32, 33) may prove to be one of these binding activities. C/EBP and AGP/EBP have been shown to be resistant to heat treatment at 90°C for 15 min (7, 20). The heat sensitivity of other members of the C/EBP family has not been reported. To assess whether binding proteins of box-α and box-β are C/EBP or members of the C/EBP family, we set out to determine the heat sensitivity of box-α- and box-β-binding proteins and their relative binding affinities toward the box-α, box-β, and C/EBP consensus sequences. As shown in Fig. 7A and B, the box-α- and box-β-binding proteins that took part in the formation of complexes b to d appeared to be heat sensitive, since these complexes disappeared after heat treatment (70°C for 5 min for b; 50°C for 5 min for c; and 70°C for 5 min for d). In contrast, the box-α-binding protein(s) in complex a could withstand heat treatment at 90°C for 15 min.

To determine the sequence preference of box-α- and box-β-binding proteins, unlabeled box-α, box-β, and C/EBP consensus sequences at 5-, 25-, 125-, and 625-fold molar excesses were used as competitors in gel shift experiments. The box-α-binding proteins favor box-α over box-β. However, the C/EBP consensus sequence failed to effectively compete even at a 625-fold molar excess (data not shown). Since the 0.5 M fraction produced a single heat-resistant complex a, the competition experiment was performed with this fraction to avoid the interference of complex b and confirm the sequence preference of complex a. Indeed, the binding protein(s) of complex a preferred box-α to box-β (Fig. 7C); again, the C/EBP consensus at a high molar excess did not show effective competition.

In contrast, for box-β-binding proteins, the unlabeled C/EBP consensus sequence more effectively abolished the formation of complexes c and d than did box-β itself (Fig. 7D). This result suggests that box-β-binding proteins, in fact, prefer the C/EBP consensus to the box-β sequence. Unlabeled box-α failed to compete for the formation of these two complexes.

To compare the sequence preferences of box-α- and box-β-binding proteins with that of the recombinant C/EBP protein (gift of Steven L. McKnight), the competition experiment was performed with recombinant C/EBP protein, using the C/EBP consensus as the labeled probe. As shown in Fig. 7E, C/EBP protein had much higher affinity for the C/EBP consensus than for either box-α or box-β.

Multiple nuclear proteins binding to the box-α sequence. We have shown previously that box-α-binding proteins interact in a gel shift assay with the box-α sequence (nt 1645 to 1669) to form two complexes (a and b) that differ in heat lability. To determine how many binding factors interact with the box-α sequence and map their binding sites more

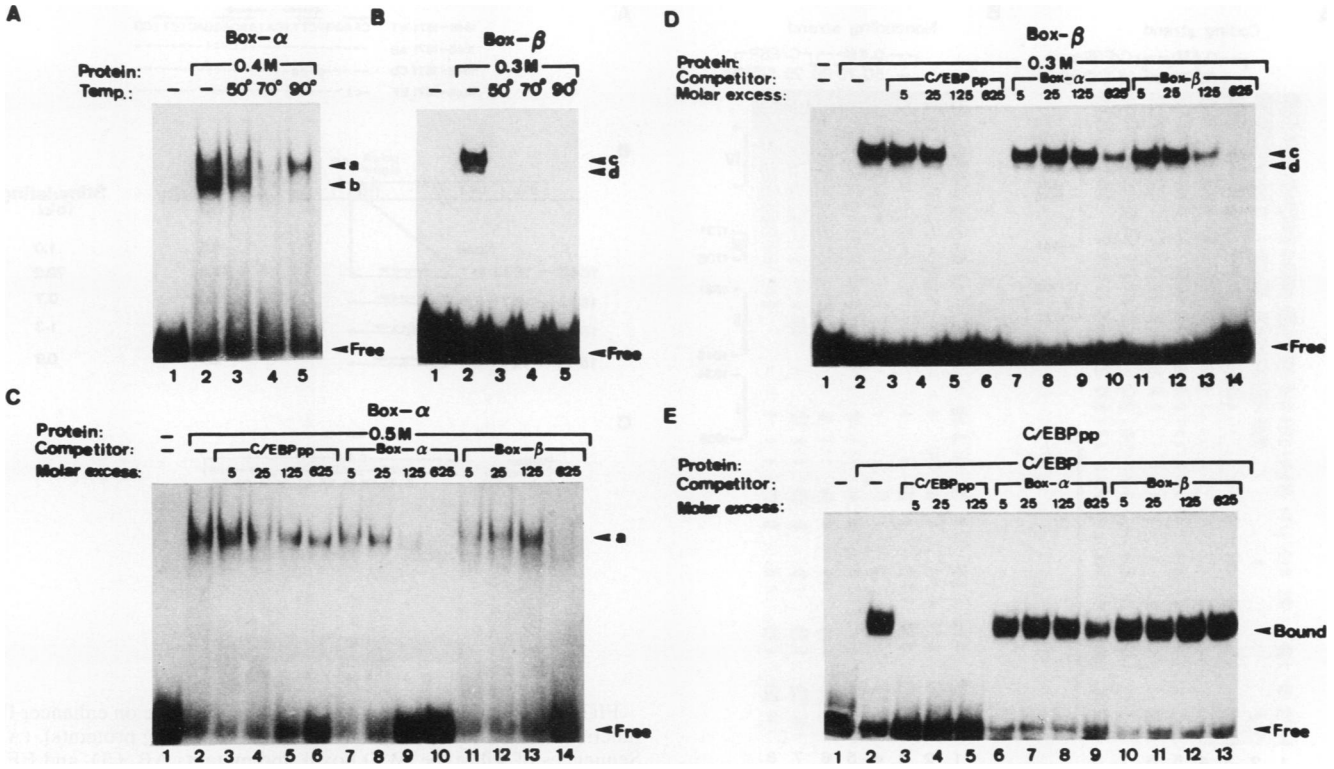


FIG. 7. Characteristics of box-α- and box-β-binding proteins. (A) Heat sensitivity of box-α-binding proteins. End-labeled box-α was incubated with 10 μg of protein of the 0.4 M NaCl fraction of HepG2 nuclear extracts without (lane 2) or with heat treatment at 50°C for 5 min (lane 3), at 70°C for 5 min (lane 4), and at 90°C for 15 min (lane 5). Lane 1 is a no-protein control. Other experimental details are as described in the legend to Fig. 5. (B) Heat sensitivity of box-β-binding proteins. End-labeled box-β was incubated with 10 μg of protein of the 0.3 M NaCl fraction. Other treatments and symbols are as for panel A. (C) Relative efficiency of competitions by box-α, box-β, and C/EBP consensus sequences for box-α-binding proteins. End-labeled box-α was incubated with 10 μg of the 0.5 M NaCl fraction in the presence of no competitor (lane 2) and of unlabeled C/EBP consensus sequence (lanes 3 to 6), box-α (lanes 7 to 10), and box-β (lanes 11 to 14) competitors (molar excess is indicated) for gel shifting. Lane 1 is a no-protein control. One specifically shifting band, a, is indicated; free denotes uncomplexed probe. (D) Relative efficiency of competitions by box-α, box-β, and C/EBP consensus sequences for box-β-binding proteins. End-labeled box-β was incubated with 10 μg of the 0.3 M NaCl fraction in the presence of no competitor (lane 2), unlabeled C/EBP consensus sequence (lanes 3 to 6), box-α (lanes 7 to 10), and box-β (lanes 11 to 14). Lane 1 is a no-protein control. Two shifted bands, c and d, are indicated. (E) Sequence preference for the recombinant C/EBP protein. End-labeled C/EBP consensus sequence was incubated with 5 ng of recombinant C/EBP in the presence of no competitor (lane 2) and of unlabeled C/EBP consensus sequence (lanes 3 to 5), box-α (lanes 6 to 9), and box-β (lanes 10 to 13) competitors (molar excess is indicated). Bound indicates the C/EBP-DNA complex.

precisely, a footprinting assay using the 0.4 M NaCl fraction before and after heat treatment was undertaken.

As shown in Fig. 8A, the protection (on the coding strand) by the unheated 0.4 M fraction was from nt 1645 to 1681, a region larger than that observed with the heated fraction (from nt 1645 to 1662). This finding indicates that the protections over nt 1645 to 1662 and nt 1662 to 1681 resulted from interactions with heat-resistant and heat-labile binding factors, respectively. As described for the gel shift analysis, the oligomer corresponding to the sequence from nt 1645 to 1669 produces one heat-resistant complex a and one heat-labile complex b, suggesting that these two proteins bind to this sequence. Therefore, there are at least three binding proteins that interact with the sequence from nt 1645 to 1681. One is heat-resistant protein a (nt 1645 to 1662), the boundary of which was clearly defined by the protection region of the heated 0.4 M fraction. Heat-labile protein b was evidenced by the presence of a second, heat-sensitive complex b in the gel shift assay, the boundary of which is difficult to define because of the possible overlapping protection with protein a. And another heat-labile protein, e, was revealed by the distinct protective pattern in the footprinting analysis. However, the possible overlapping protection with protein b make it difficult to precisely map the boundary of protein e.

Consistent with this conclusion, 0.5 and 0.6 M NaCl fractions, which gave rise to complex a, a heat-resistant activity binding solely to box-α (Fig. 4B), produced a binding pattern indistinguishable from that with the heated 0.4 M fraction (Fig. 5B).

Since the sequence protected by the heat-resistant complex a protein(s) shows homology with the C/EBP consensus sequence, the recombinant C/EBP protein was used for DNase I footprinting analysis. It was indeed able to bind to the box-α sequence, which is consistent with the result of López-Carbrera et al. (25). Furthermore, the recombinant C/EBP produced a protection pattern very similar to that of the heated 0.4 M NaCl fraction.

Indispensable requirement of the entire box-α sequence for enhancer II function. A closer analysis of the box-α sequence reveals some interesting features (Fig. 9A). In addition to the existence of a sequence weakly homologous to the extended C/EBP consensus (shown as thicker arrows in Fig. 9A), the

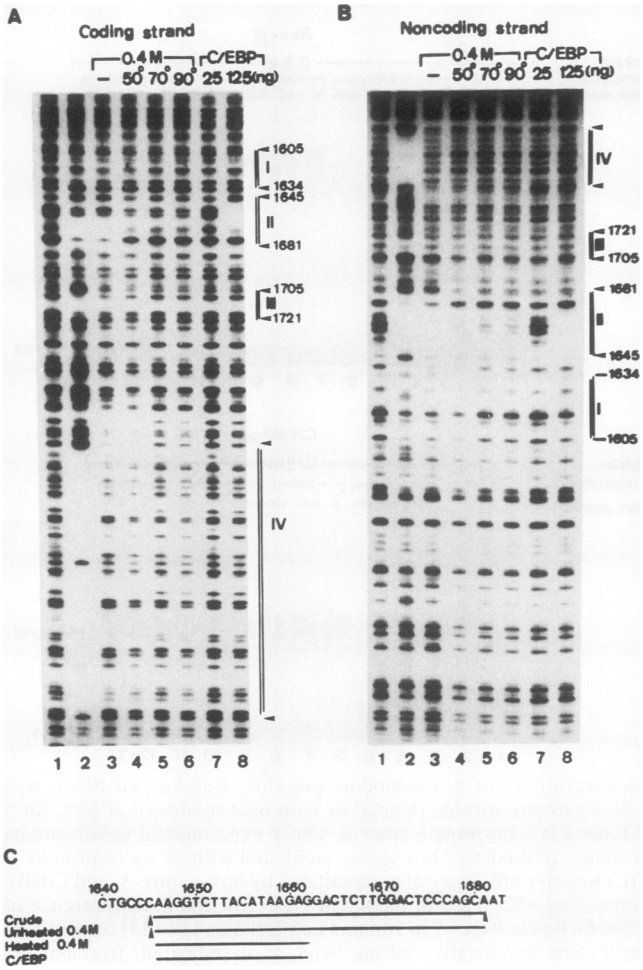


FIG. 8. Heat sensitivity of the 0.4 M NaCl fraction of HepG2 nuclear extracts in DNase I footprinting analysis. (A) Footprinting analysis on the coding strand. End-labeled probe was incubated with 120 μg of crude nuclear extracts (lane 2), 40 μg of the 0.4 M NaCl fraction unheated (lane 3) and heated at 50°C for 5 min (lane 4), 70°C for 5 min (lane 5), and 90°C for 15 min (lane 6), and 25 ng (lane 7) and 125 ng (lane 8) of recombinant C/EBP. Lane 1 is the no-protein control. The protected regions with crude nuclear extracts are shown on the right as two lines with numerals indicating the nucleotide number of HBV. Arrowheads indicate the DNase I-hypersensitive sites. (B) Footprinting analysis on the noncoding strand. Experimental details and symbols are as for panel A. (C) Summary of the protected regions in the HBV enhancer II sequence.

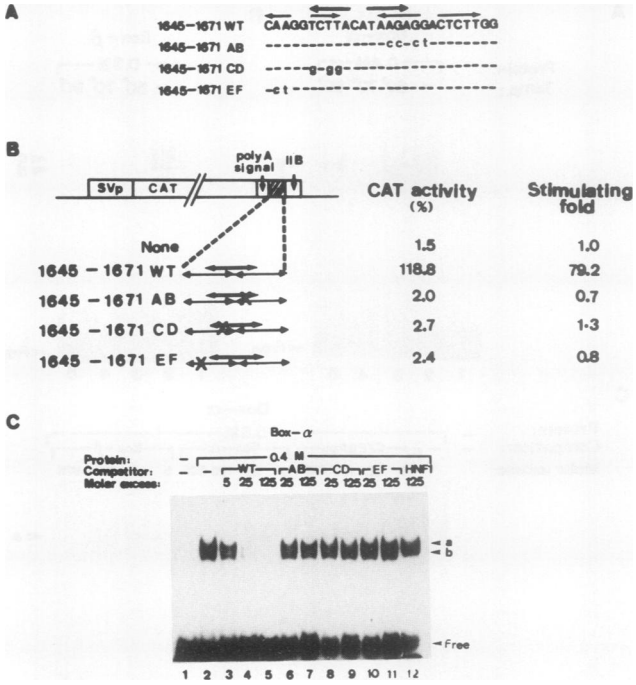


FIG. 9. Effect of mutations of the box-α sequence on enhancer II function and on the binding affinity to box-α-binding protein(s). (A) Sequences of wild-type (WT) box-α and mutants AB, CD, and EF. Mutated nucleotides are represented by lowercase letters. The upper arrows represent the inverted repeat weakly homologous to the extended C/EBP conserved sequence; the lower arrows represent two adjacent inverted repeats. (B) Summary of the enhancer function of the wild-type box-α sequence and its three mutants through cooperation with the box-β sequence. Experimental details are as described in the legend to Fig. 1. (C) Competition with box-α of three mutated box-α sequences in its binding activity to the box-α-binding protein(s) by gel shift assay. Lanes: 1, no protein; 2, 10 μg of protein of the 0.4 M NaCl eluent; 3 to 12, 10 μg of protein of the 0.4 M NaCl eluent and different unlabeled competitors at various molar excesses as indicated.

region appears to consist of two distinct inverted repeats (shown as thinner arrows). Deletion studies have so far shown that a 23-bp sequence (nt 1646 to 1668), but not a 20-bp sequence (nt 1646 to 1665), is sufficient for the box-α function (Fig. 2). It has not been determined, however, whether all of the 23 bp are required for this activity. Alterations of sequences in this region (Fig. 9A) were made and tested for their effects on enhancer function in the presence of intact II-B. It was found that nucleotide changes throughout this sequence (nt 1647 to 1662) in mutants AB, CD, and EF completely abolished the box-α function (Fig. 9B). Oligonucleotides bearing these changes were then examined for their propensity to disrupt specific complex formation in a gel shift assay, using labeled wild-type oligonucleotides. As shown in Fig. 9C, competition was by the wild-type oligomer at a fivefold molar excess was observed. However, even at a 125-fold molar excess, the mutant oligomers did not compete at all. This result reflects at least a 25-fold decrease in binding affinity incurred by these sequence variations toward their cognate binding factors. Such a diminution in their abilities to interact with relevant factor(s) may well explain the failure of these box-α mutations to implement enhancer function. Therefore, the entire sequence from nt 1646 to 1668 is essential for box-α to cooperate with box-β to achieve the enhancer function.

DISCUSSION

The second enhancer of HBV, mapping at nt 1636 to 1741, stimulates transcription from the homologous HBV promoter and such heterologous promoters as the SV40 early promoter in a differentiated human hepatoma cell-specific manner. In previous work, we have dissected this enhancer into two elements, II-A (nt 1636 to 1690) and II-B (nt 1704 to 1741), the cooperation of which is required for enhancer II function (39). In this study, two elements have been further delineated by detailed deletion mapping. A 23-bp sequence, designated box-α (nt 1646 to 1668), in element II-A and a

12-bp sequence, designated box-β (nt 1704 to 1715), in element II-B are found to be indispensable and sufficient for enhancer II function. Yee (38) has shown that the sequence from nt 1686 to 1775, which does not contain the box-α sequence, exerts an enhancer activity on the thymidine kinase promoter of herpes simplex virus. The reason for the discrepancy between his and our results is unclear.

The dissection of enhancer II into box-α and box-β and the obligate interaction between the two for enhancer function suggest an intriguing interplay between the two boxes and their binding factor(s). It would be interesting to determine whether an enhancer function would have a stringent requirement for spacing or orientation between the two boxes. In one example, in which we put a box-β sequence upstream of an inverted box-α, a 10-fold decrease in enhancer activity was observed (data not shown). On the other hand, when the spacing between box-α and box-β was varied and set at 1, 9, 23, and 35 nt (Fig. 2), the enhancer function was largely unaffected. These results indicate that the spacing between the two boxes is not under serious constraint. On the other hand, the positions or relative orientations of the two boxes seem to play an important role in their combined effects on transcription activation.

C/EBP binds to the enhancer sequence of several viral long terminal repeats and the CCAAT homology (16, 19). It is also a transcriptional activator of the albumin gene in hepatocytes and of several adipocyte-specific genes in preadipocytes (10, 17, 24). The binding sites for C/EBP have the consensus sequence T(T/G)NNG(A/C/G)AA(T/G). Several proteins such as DBP (27), Ig/EBP-1 (32), and LAP (11) (NF-IL6, AGP/EBP, and IL-6DBP designate the same protein) (3, 7, 32) also bind to C/EBP binding sites. These proteins, which are highly homologous in their basic and leucine zipper regions (22), are considered to form a multigene family, the C/EBP family. However, DBP does not have the leucine zipper region. Among them, C/EBP has been shown to form heterodimers with IL-6DBP (LAP) or Ig/EBP-1 (11, 32, 33). Therefore, members of this family are able to form many different transcription factors, probably with different properties.

Box-α (CAAGGTCTTACATAAGAGGACTCTT, nt 1645 to 1669) contains in its center a sequence (TTACATAAG, nt 1652 to 1660) that loosely fits the extended C/EBP consensus T(T/G)NNG(A/C/G)AA(T/G) (34). Alternatively, box-α appears to comprise two adjacent but not identical inverted repeats (CAAGGTCTTAC, nt 1645 to 1655; and ATAAG AGGACTCTT, nt 1656 to 1669). Footprinting and gel shift analyses reveal that two proteins, a heat-resistant protein a and a heat-labile protein b, bind to the box-α sequence. While the binding region of protein a is from nt 1645 to 1662, that of protein b is more difficult to define. Mutants AB, CD, and EF of the box-α sequence, which have 2- to 4-nt mutations at 1659 to 1663, 1650 to 1653, and 1646 to 1647, respectively, cannot bind either protein a or protein b. This finding indicates that the binding of proteins a and b requires the sequence from at least nt 1647 to 1662.

Protein a is a heat-resistant protein whose binding sequence is bound by the recombinant C/EBP protein. There are indications, however, that protein a is probably a modified form of C/EBP, a heterodimer containing C/EBP or other members of the C/EBP family but not identical to C/EBP. For example, C/EBP and protein a are noted to have different sequence preferences for binding. The recombinant C/EBP protein has a much higher binding affinity for its own binding site than for box-α, while the reverse is true for protein a. On the other hand, protein b, being heat labile, may be a member of a C/EBP-like family or of unrelated nature. Neither protein a nor protein b is identical to C/EBP, which is consistent with the observation that C/EBP is expressed at only a trace level in HepG2 cells (12). Interestingly, protein a can form complex a in the absence of complex b, suggesting it can interact directly with box-α (Fig. 5B). In contrast, the nature of protein b binding remains to be elucidated. In addition, it is worth noting that the signal of complex a, which is not affected by heat treatment at 90°C for 15 min, is reproducibly observed to be lower when heated at either 50 or 70°C for 5 min. It is possible, although unlikely, that the binding protein in complex a is sensitive to heat treatment and that another heat-resistant protein with a similar molecular weight binds to box-α and migrates to the same position at 90°C.

Box-β (CCTACTTCAAAGACTGT, nt 1704 to 1720) also remotely resembles the C/EBP consensus, in which the TT and AA in the middle TTcaAA sequence are reminiscent of the most conserved nucleotide residues TT and AA in the extended C/EBP consensus T(T/G)NNG(A/C/G)AA(T/G). It is worth pointing out, however, that only two nucleotides separate the TT and AA pairs in box-β, rather than four as seen in the extended C/EBP consensus. So far, the binding of a nucleotide sequence with this feature by members of the C/EBP family has not been reported. Two relatively heat sensitive binding proteins, c and d, are found to bind to box-β. Both proteins have better affinity for the C/EBP consensus sequence than for box-β itself. However, both proteins c and d are relatively heat labile, in contrast to the heat resistance of C/EBP. Thus, they are likely to be molecules other than C/EBP, such as a modified form of C/EBP, a C/EBP family member, or a heterodimer formed between C/EBP family members.

Proteins a and b bind to the indispensable box-α sequence and are good candidates for *trans*-acting factors that activate transcription via interaction with the candidates for box-β-binding proteins, proteins c and d. Mutations through the box-α sequence that abolish the binding of proteins a and b also render box-α functionless as an enhancer. However, results so far do not indicate whether proteins a and b are functionally redundant or are both required for enhancer function.

C/EBP family members function as transcriptional activators in the expression of liver-specific and adipocyte-specific genes (2, 10, 17). They are probably also involved in the signal transduction of interleukin-1 and interleukin-6 and some other functions (23, 31, 37). In this study, the interaction between box-α-binding proteins and box-β-binding proteins is essential to achieve the enhancer function. If proteins a, c, and d are indeed members of the C/EBP family, the second enhancer of HBV provides a unique model for investigations of how their interactions lead to transcriptional activation in an orientation- and position-independent manner.

ACKNOWLEDGMENTS

We thank Steven L. McKnight for providing the rat recombinant C/EBP protein, Sheng-Chung Lee for providing the C/EBP consensus sequence oligomer, Kong-Bung Choo and Shiuh-Wen Luoh for critically reviewing the manuscript, and Ming-Muiew Tsai for preparing the manuscript.

This study was supported by research grant NSC-80-0419-B010-21 from the National Science Council, Republic of China.

REFERENCES

1. **Aden, D. P., A. Fogel, S. S. Plotkin, I. Damjanov, and B. B. Knowles.** 1979. Controlled synthesis of HBsAg in a differenti-

ated human liver carcinoma-derived cell line. Nature (London) **282**:615–617.
2. **Agre, P., P. F. Johnson, and S. L. McKnight.** 1989. Cognate DNA binding specificity retained after leucine zipper exchange between GCN4 and C/EBP. Science **246**:922–926.
3. **Akira, S., H. Isshiki, T. Sugita, O. Tanabe, S. Kinoshita, Y. Nishio, T. Nakajima, T. Hirano, and T. Kishimoto.** 1990. A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP family. EMBO J. **9**:1897–1906.
4. **Beasley, R. P., L. Y. Hwang, C. C. Lin, and C. S. Chien.** 1981. Hepatocellular carcinoma and hepatitis B virus: a prospective study of 22,707 men in Taiwan. Lancet **ii**:1129–1136.
5. **Birkenmeier, E. H., B. Gwynn, S. Howard, J. Jerry, J. I. Gordon, W. H. Landschulz, and S. L. McKnight.** 1989. Tissue-specific expression, developmental regulation and genetic mapping of the gene encoding C/EBP. Genes Dev. **3**:1146–1156.
6. **Bosch, V., C. Kuhn, and H. Schaller.** 1988. Hepatitis B virus replication, p. 43–58. *In* E. Domingo, J. J. Holland, and P. Ahiquist (ed.), Retroviruses, viroids, and RNA recombination, RNA genetics, vol. 2. CRC Press, Cleveland, Ohio.
7. **Chang, C. J., T. T. Chen, H. Y. Lei, D. S. Chen, and S. C. Lee.** 1990. Molecular cloning of a transcription factor, AGP/EBP, that belongs to members of the C/EBP family. Mol. Cell. Biol. **10**:6642–6653.
8. **Chang, H. K., and L. P. Ting.** 1989. The surface gene promoter of the human hepatitis B virus displays a preference for differentiated hepatocytes. Virology **170**:176–183.
9. **Chang, H. K., B. Y. Wang, C. H. Yuh, C. L. Wei, and L. P. Ting.** 1989. A liver-specific nuclear factor interacts with the promoter region of the large surface protein gene of human hepatitis B virus. Mol. Cell. Biol. **9**:5189–5197.
10. **Christy, R. J., V. W. Yang, J. M. Ntambi, D. E. Geiman, W. H. Landschulz, A. D. Friedman, Y. Nakabeppu, T. J. Kelly, and M. D. Lane.** 1989. Differentiation-induced gene expression in 3T3-L1 preadipocytes: CCAAT/enhancer binding protein interacts with and activates the promoters of two-adipocyte-specific genes. Genes Dev. **3**:1323–1335.
11. **Descombes, P., M. Chojkier, S. Lichtsteiner, E. Falvey, and U. Schibler.** 1990. LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. Genes Dev. **4**:1541–1551.
12. **Friedman, A. D., W. H. Landschulz, and S. L. McKnight.** 1989. CCAAT/enhancer binding protein activates the promoter of the serum albumin gene in cultured hepatoma cells. Genes Dev. **3**:1314–1322.
13. **Fromental, C., M. Kanno, H. Nomiyama, and P. Chambon.** 1988. Cooperativity and hierarchical levels of functional organization in the SV40 enhancer. Cell **54**:943–953.
14. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu. Rev. Biochem. **56**:651–693.
15. **Gorman, C., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
16. **Graves, B., P. F. Johnson, and S. L. McKnight.** 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell **44**:565–576.
17. **Herrera, R., H. S. Ro, G. S. Robinson, K. G. Xanthopoulos, and B. M. Spiegelman.** 1989. A direct role for C/EBP and the AP-I-binding site in gene expression linked to adipocyte differentiation. Mol. Cell. Biol. **9**:5331–5339.
18. **Hones, N. C., P. W. J. Rigby, and E. B. Ziff.** 1988. Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. Genes Dev. **2**:267–281.
19. **Johnson, P. F., W. H. Landschulz, B. J. Graves, and S. L. McKnight.** 1987. Identification of a rat liver nuclear protein that binds to the enhancer core element of three animal viruses. Genes Dev. **1**:133–146.
20. **Landschulz, W. H., P. F. Johnson, E. Y. Adashi, B. J. Graves, and S. L. McKnight.** 1988. Isolation of a recombinant copy of the gene encoding C/EBP. Genes Dev. **2**:786–800.
21. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240**:1759–1764.
22. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1989. The DNA binding domain of the rat liver nuclear protein C/EBP is bipartite. Science **243**:1681–1688.
23. **Li, X., J. H. Huang, H. Y. Rienhoff, Jr., and W. S. L. Liao.** 1990. Two adjacent C/EBP-binding sequences that participate in the cell-specific expression of the mouse serum amyloid A3 gene. Mol. Cell. Biol. **10**:6624–6631.
24. **Lichtsteiner, S., J. Wuarin, and U. Schibler.** 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. Cell **51**:963–973.
25. **López-Cabrera, M., J. Letovsky, K. Q. Hu, and A. Siddiqui.** 1990. Multiple liver-specific factors bind to the hepatitis B virus core/pregenomic promoter: trans-activation and repression by CCAAT/enhancer binding protein. Proc. Natl. Acad. Sci. USA **87**:5069–5073.
26. **Meyer, K. B., and M. S. Neuberger.** 1989. The immunoglobulin κ locus contains a second, stronger B-cell specific enhancer which is located downstream of the constant region. EMBO J. **8**:1959–1964.
27. **Mueller, C. R., P. Maire, and U. Schibler.** 1990. DBP, a liver-enriched transcriptional activator, is expressed late in ontogeny and its tissue specificity is determined posttranscriptionally. Cell **61**:279–291.
28. **Nomiyama, H., C. Fromental, J. H. Xiao, and P. Chambon.** 1987. Cell-specific activity of the constituent elements of the simian virus 40 enhancer. Proc. Natl. Acad. Sci. USA **84**:7881–7885.
29. **Ondek, B., L. Gloss, and W. Herr.** 1988. The SV40 enhancer contains two distinct levels of organization. Nature (London) **333**:40–45.
30. **Ondek, B., A. Shepard, and W. Herr.** 1987. Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J. **6**:1017–1025.
31. **Poli, V., and R. Cortese.** 1989. Interleukin 6 induces a liver-specific nuclear protein that binds to the promoter of acute-phase genes. Proc. Natl. Acad. Sci. USA **86**:8202–8206.
32. **Poli, V., F. P. Mancini, and R. Cortese.** 1990. IL-6DBP, a nuclear protein involved in interleukin-6 signal transduction, defines a new family of leucine zipper proteins related to C/EBP. Cell **63**:643–653.
33. **Roman, C., J. S. Platero, J. Shuman, and K. Calame.** 1990. Ig/EBP-1: a ubiquitously expressed immunoglobulin enhancer binding protein that is similar to C/EBP and heterodimerizes with C/EBP. Genes Dev. **4**:1404–1415.
34. **Ryden, T. A., and K. Beemon.** 1989. Avian retroviral long terminal repeats bind CCAAT/enhancer-binding protein. Mol. Cell. Biol. **9**:1155–1164.
35. **Shaul, Y., W. J. Rutter, and O. Laub.** 1985. A human hepatitis B virus enhancer element. EMBO J. **4**:427–430.
36. **Wang, Y., P. Chen, X. Wu, A. L. Sun, H. Wang, Y. A. Zhu, and Z. P. Lee.** 1990. A new enhancer element, ENII, identified in the X gene of hepatitis B virus. J. Virol. **64**:3977–3981.
37. **Wilson, D. R., T. S. C. Juan, M. D. Wilde, G. H. Fey, and G. J. Darlington.** 1990. A 58-base-pair region of the human C3 gene confers synergistic inducibility by interleukin-1 and interleukin-6. Mol. Cell. Biol. **10**:6181–6191.
38. **Yee, J. K.** 1989. A liver-specific enhancer in the core promoter region of human hepatitis B virus. Science **246**:658–661.
39. **Yuh, C. H., and L. P. Ting.** 1990. The genome of hepatitis B virus contains a second enhancer: cooperation of two elements within this enhancer is required for its function. J. Virol. **64**:4281–4287.