JOURNAL OF VIROLOGY, Dec. 1992, p. 6841–6848
0022-538X/92/126841-08$02.00/0

Vol. 66, No. 12

Identification and Cloning of a Novel Heterogeneous Nuclear Ribonucleoprotein C-Like Protein That Functions as a Transcriptional Activator of the Hepatitis B Virus Enhancer II

NICHOLAS TAY, SOH-HA CHAN, AND EE-CHEE REN*

W.H.O. Collaborating Centre for Immunology and Department of Microbiology, Faculty of Medicine, National University of Singapore, Singapore 0511, Republic of Singapore

Received 6 July 1992/Accepted 24 August 1992

Liver specificity of hepatitis B virus (HBV) replication has been attributed to the action of its second enhancer (EII). We report here the characterization of EII and the subsequent isolation of a novel liver-specific DNA-binding protein which binds to and activates EII. The cDNA clone of the protein, designated E2BP, was isolated from a λgt11 expression library constructed from the hepatoma cell line HuH-6 which was screened with a binding site probe derived from EII. Sequence analysis of E2BP revealed 86.6% homology with a rat heterogeneous nuclear ribonucleoprotein C protein sequence, while conformational studies suggest a helix-loop-helix motif as a DNA-binding site. Cloned E2BP expressed in human fibroblasts by transient transfection displayed EII binding and activating characteristics similar to those of native E2BP in hepatocytes.

In hepatitis B virus (HBV)-infected hepatocytes, two major viral transcripts of 3.6- and 2.2-kb mRNAs have been observed during replication (3, 33). While the 2.2-kb mRNA is involved in the production of pre-S2 and S proteins, the 3.6-kb mRNA is involved in the expression of several protein products such as HBV core antigen, HBV envelope antigen, and viral polymerase. It also functions as the pregenome for virus replication (2). In DNA transfection studies, the 3.6-kb mRNA has been found in well-differentiated human hepatoma cell lines but is generally undetectable in nonhepatic cells (27, 30, 35). This finding suggests that either liver-specific factors, differentiation-related factors, or a combination of both is needed for correct transcription from the core promoter to produce the 3.6-kb pregenomic mRNA. Previous data have also shown that the HBV promoters are in turn regulated by HBV enhancer elements (1, 10, 36) which appear to be especially active in hepatocytes, resulting in high-level expression of HBV (25). This may be one of the reasons for the pronounced hepatotrophicity of the virus, i.e., the tissue specificity of HBV gene expression is determined by the combinatorial action of the enhancer and promoter. However, the actual mechanism and the interaction with cellular factors giving rise to tissue specificity are still obscure.

To date, two enhancers for transcription of the 3.6-kb mRNA have been identified. The first enhancer (EI), located between nucleotides (nt) 1070 and 1234, approximately 600 bp upstream from the transcription initiation sites of the 3.6-kb message (23), activates transcription in a relatively tissue independent manner (31). These results argue against a major role for this enhancer in directing tissue-specific control of HBV replication. The second enhancer (EII), located in the X gene within an 87-bp region at nt 1554 to 1645 (36), was by itself sufficient to direct efficient chloramphenicol acetyltransferase (CAT) expression in hepatoma cells, which indicated the possibility that EII was more tissue restricted than was EI. At the same time, a number of studies have alluded to the existence of liver-specific transcription factors that bind to and regulate the activity of EII (32, 34, 36). Furthermore, Yee (36) and Honigwachs et al. (10) showed that these factors were present in high amounts in well-differentiated hepatoma cell lines and were either undetectable or absent in dedifferentiated hepatoma cell lines and nonhepatic cell lines. In view of this finding, the identification of such DNA-binding proteins that regulate HBV replication would be important in helping us to understand the HBV replication cycle and possibly its role in the etiology of hepatocellular carcinoma. Therefore, we set out to determine the action of EII and its interaction and dependency on liver-specific factors. Here we report the characterization of EII, the identification of a novel positive regulatory *trans*-acting factor required for EII activity, and the cloning and characterization of this liver-specific factor, designated E2BP.

MATERIALS AND METHODS

Construction of an enhancer-CAT plasmid. CAT plasmids pCAT-Basic (an enhancerless and promoterless CAT plasmid) and pCAT-Promoter (an enhancerless CAT plasmid with a simian virus 40 promoter upstream from the CAT gene) were obtained from Promega Corp. (Madison, Wis.). Plasmid p87/26-CAT was constructed by ligation of the HBV DNA 87-bp *Hin*cII-*Rsa*I fragment (nt 1554 to 1645) excised from pADR4 (7) into a pCAT-Promoter vector which was digested with *Bgl*II and made blunt ended with Klenow polymerase.

Cell transfections and CAT assays. Cells were transfected with either p87/26-CAT, pCAT-Promoter, pCAT-Basic, or pSV-E2BP by the calcium phosphate coprecipitation method (9). Cell extracts were prepared and assayed for CAT activity by the method of Seed and Sheen (21). Quantitation of CAT activity, based on the enzymatic butyrylation of [^{14}C]chloramphenicol (NEN Research Products, Boston, Mass.), was determined by liquid scintillation counting. This highly sensitive assay is linear over 3 orders of magnitude of enzyme concentration and gave consistent and repeatable

* Corresponding author.

results. The relative CAT activity was calculated as a percentage fold of that of pCAT-Promoter, which was normalized to 1.0. The values reported are means of three to five independent experiments ± standard deviation.

Cell cultures. (i) Primary hepatocyte culture. Normal liver tissue was washed in RPMI 1640, minced and prepared as tissue explants, and maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum.

(ii) Cell lines. All hepatic cell lines were obtained from their laboratories of origin. All nonhepatic cell lines were obtained from the American Type Culture Collection. Adhesion cells were cultured as monolayers in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, while suspension cells were cultured in RPMI supplemented with 10% fetal calf serum.

Isolation of cell nuclei and preparation of nuclear extracts. Nuclear extracts were prepared as described by Dignam et al. (6). Protein concentration was determined with the Bio-Rad protein assay kit, with bovine serum albumin as the standard. Typical yield was 500 μg of nuclear protein isolated from 10^7 HuH-6 cells.

Gel retardation assays. Gel retardation assays were performed by the method of Rhodes (18). Binding reaction procedures were performed for 15 min in a total volume of 25 μl in a buffer containing 10 mM Tris-HCl (pH 7.8), 50 mM NaCl, 1 mM EDTA, and 1 mM dithiothreitol; 4 to 8 μg of nuclear extract was included in each reaction mixture along with nonspecific competitor DNA [poly(dI)-poly(dC); Boehringer GmbH, Mannheim, Germany]. End-labeled binding-site DNA was added last after a 10-min preincubation period. After binding, the reaction mixtures were loaded onto 4% nondenaturing polyacrylamide gels (30:1 acrylamide-to-bisacrylamide ratio); the mixtures were electrophoresed at 10 V/cm with recirculating buffer (6.6 mM Tris-HCl [pH 7.6], 3.3 mM sodium acetate, 1 mM EDTA) for 2 to 3 h. Gels were dried and autoradiographed with intensifying screens at −70°C.

Cloning of E2BP cDNA. The cDNA of E2BP was obtained by screening of a HuH-6 cDNA library, using the protocol of Singh et al. (24). The cDNA was synthesized and cloned into λgt11, using kits from Amersham. The library, which contained 2 million individual plaques, was screened as described previously (24), with the following modifications: the concentration of poly(dI-dC) was reduced to 2.5 μg/ml, and the washing period was increased to a total of 45 min. The probe used was the synthesized EIIb oligonucleotide, which is composed of the sequence

```
5'-GATCCTGGGAGGAGTTGGGGGAGGAGATTAGG     -3'
3'-     GACCCTCCTCAACCCCCTCCTCTAATCCCTAG-5'
```

The probe was end labeled with [^{32}P]dCTP and Klenow polymerase. The full-length cDNA clone isolated was subcloned into pUC18 and sequenced by using Pharmacia's T7 double-stranded sequencing kit.

Northern (RNA) hybridization analysis of E2BP mRNA. Total RNA from the cell lines HuH-6 and FS4 were extracted and purified by the method of Chomczynski and Sacchi (4). Total RNA (40 μg per lane) was electrophoresed through a 1.4% formaldehyde gel as described by Sambrook et al. (20) and transferred onto a Hybond-N membrane. The Northern blot was hybridized with ^{32}P-labeled E2BP cDNA and washed as described by Striebich et al. (26).

GenBank accession number. The GenBank accession number assigned for the E2BP cDNA sequence is M94630.

TABLE 1. CAT assay of p87/26-CAT against a cell panel

Cell line	Cell type	Relative CAT activity[a]
Liver		
Primary hepatocyte culture		15.6 (1.5)
HuH-6	Hepatoblastoma	51.4 (3.7)
HepG2	Hepatoblastoma	9.6 (1.9)
huH-4[b]	Hepatoma	7.5 (1.5)
Tong/HCC[b]	Hepatoma	5.5 (0.3)
HuH-7	Hepatoma	4.7 (0.4)
PLC/PRF/5[b]	Hepatoma	3.8 (0.6)
huH-1[b]	Hepatoma	3.6 (0.5)
HA22T/VGH[b]	Hepatoma	3.4 (<0.1)
Hep3B[b]	Hepatoma	3.1 (0.3)
KMCH-1	Hepatocholangiocarcinoma	2.5 (<0.1)
HCC-M	Hepatoma	2.0 (<0.1)
HCC-T	Hepatoma	1.6 (<0.1)
Mahlavu	Hepatoma	1.1 (<0.1)
SK-Hep1	Liver adenocarcinoma	1.1 (<0.1)
MH_1C_1	Rat hepatoma	3.3 (0.3)
Nonliver		
A431	Epidermoid carcinoma	1.0 (<0.1)
WISH	Transformed amnion cells	0.8 (<0.1)
FS4	Foreskin fibroblast	0.7 (<0.1)
MCF-7	Breast adenocarcinoma	0.6 (<0.1)
Tera-2	Embryonal carcinoma	0.5 (<0.1)
NIH 3T3	Mouse fibroblast	0.7 (<0.1)
Lymphoid		
Reh	ALL[c] (pre-B)	0.9 (<0.1)
Molt-4	ALL (mature T)	0.9 (<0.1)
U937	Histiocytic lymphoma	0.9 (<0.1)
RPMI 8226	Plasmacytoma	0.8 (<0.1)
YB2/0	Rat myeloma	0.7 (<0.1)

[a] Fold increase of p87/26-CAT activity over that of the control pCAT-Promoter vector (activity = 1.0). Values are means and standard deviations (in parentheses) of triplicate experiments.
[b] Contains integrated HBV in the host genome.
[c] ALL, acute lymphoblastic leukemia.

RESULTS

Characterization of EII. CAT assays performed by using the 87-bp EII ligated to the CAT reporter plasmid pCAT-Promoter demonstrated the liver specificity of EII when assayed against a panel of both hepatic and nonhepatic cell lines (Table 1). High CAT activity was seen with well-differentiated hepatoma cell lines such as HuH-6, HepG2, and HuH-7, while Mahlavu and SK-Hep1, which are poorly differentiated, did not show any significant CAT activity over the control level. These data agree well with similar observations by Yee (36) and Honigwachs et al. (10). A 7-day human primary hepatocyte culture was included to demonstrate that the enhanced CAT activity was not a consequence of transformed phenotype. The finding that the nonhepatic cell lines showed no CAT activity over the control level indicated that liver-specific factors were required for enhanced activity directed by EII. It was also observed that there was no particular correlation between the presence or absence of existing integrated HBV DNA in the cell lines (29) with the CAT activity levels. This finding suggests that HBV DNA integration per se does not modify the cellular environment to produce factors which bind to EII. However, this conclusion does not rule out the possibility that HBV integration can affect other cellular functions.

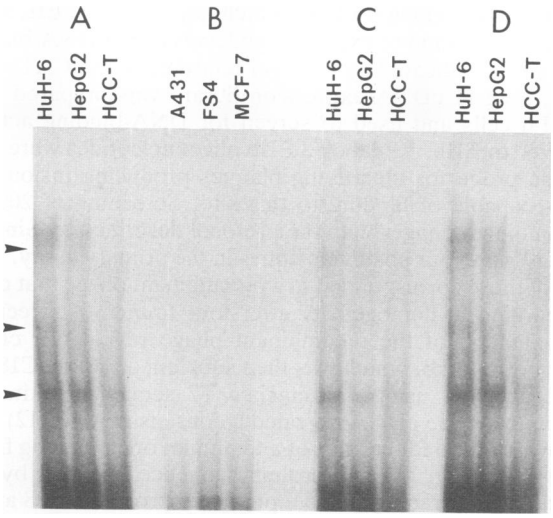

FIG. 1. Gel retardation assays of EII and EI. ^{32}P-labeled 87-bp EII DNA was incubated with nuclear extracts from hepatic and nonhepatic cell lines and electrophoresed through 4% nondenaturing polyacrylamide gels. The gels were then dried and autoradiographed with intensifying screens at −70°C. (A) Radiolabeled EII with hepatoma cell lines HuH-6, HepG2, and HCC-T; (B) radiolabeled EII with nonhepatic cell lines A431, MCF-7, and FS4; (C) radiolabeled EII with hepatoma cell lines HuH-6, HepG2, and HCC-T, with the addition of unlabeled competitor EII; (D) radiolabeled EII with hepatoma cell lines HuH-6, HepG2, and HCC-T, with the addition of unlabeled competitor EI. Addition of unlabeled EII DNA (C) resulted in a decrease in intensity of bands observed, while addition of unlabeled EI DNA (D) did not result in any visible change in intensity.

To identify the existence of factors interacting with EII, gel retardation assays were performed on nuclear extracts from HuH-6, HepG2, and HCC-T, using the 87-bp EII sequence. The results in Fig. 1 revealed the presence of several EII-specific DNA-binding proteins which were present in high amounts in HuH-6, followed by HepG2 and HCC-T, which paralleled the order seen in the CAT assay. The bound complexes did not appear when the assay was performed with nuclear extracts from the control nonhepatic cell lines A431, FS4, and MCF-7. Addition of unlabeled EII as a competitor resulted in a decrease in the intensities of the bound complexes, while no change was observed when an unlabeled *Acc*I-*Sph*I 164-bp competitor DNA from EI (nt 1070 to 1234) (23) was used, indicating the sequence specificity of the nuclear extract binding proteins from HuH-6, HepG2, and HCC-T. This result showed that the hepatospecific *trans*-activating factors which could bind to the 87-bp HBV enhancer are probably different from those that bind to the EI enhancer.

The 87-bp EII was shown previously by Yaginuma and Koike (34) and Yee (36) to have two regions of direct repeats which exhibit characteristics of protein-binding sites. Double-stranded oligonucleotides corresponding to the two sites, designated EIIa and EIIb (Fig. 2), were synthesized and used in gel retardation assays with nuclear extracts from HuH-6, HepG2, and HCC-T (Fig. 3). The results showed that the same pattern of three gel retardation bands seen in the assay using EII could essentially be reproduced by using EIIb alone (Fig. 3D). However, the lowest band that bound to EIIb and the single band that bound to EIIa appear to be the same, as a 10-fold molar excess of either unlabeled EIIa or EIIb could compete for the same binding. At the same time, the two other EIIb bands could be competed for only by EIIb, not by EIIa, implying that these factors were specific for EIIb DNA binding.

		HBV Sub-Type	GenBank Number
EIIa	CTTCAAAGACTGTGTGTTTAAAGACTGG		
EIIb	CTGGGAGGAGTTGGGGGAGGAGATTAGG		
EII	5'-AACGACCGACCTTGAGGCATACTTCAAAGACTGTGTGTTTAAAGACTGGGAGGAGTTGGGGGAGGAGATTAGGTTAAAGGTCTTTGT-3'	ADR	V00867
	----T----------CT----------C----------	ADR	M38636
	---A----------T----------	ADR	M12906
	----------T----------T----------	ADR	D00630
	----------T-------G----------	ADR1	M38454
	----------T-------G----------	ADR4	X01587
	----T----------CT----------C----------	ADRM	X14193
	----------C----------G----------	ADW	V00866
	-C----------CT--G----------C----------	ADW	M54923
	----------C----------	ADW	M57663
	----------T--G---------G----------	ADW1	D00329
	----------T--G----------G----------	ADW2	D00330
	----------C----------G----------C----------	ADW2	X02763
	----------C----------CT--G----------C----------	ADW3	D00331
	----------T----------	AYR	X04615
	----------T----------	AYW	V01460
	----------T----------A----------	AYW	X02496
	----------T-------G----------C----------A----------	A1	M32138
	----------C----------G----------C----------	991	X51970

FIG. 2. Comparison of EIIa and EIIb sequences in HBV strains. Direct repeat motifs are underlined.

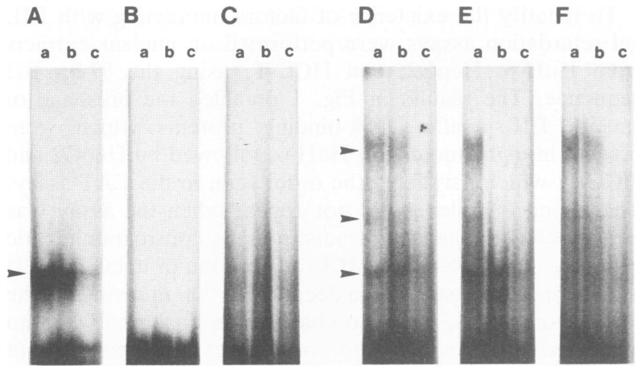

FIG. 3. Gel retardation assays of EIIa and EIIb. Nuclear extracts from hepatoma cell lines HuH-6 (lanes a), HepG2 (lanes b), and HCC-T (lanes c) were incubated with radiolabeled EIIa and EIIb oligonucleotides and electrophoresed through 4% nondenaturing polyacrylamide gels. The gels were then dried and autoradiographed with intensifying screens at −70°C. (A) With radiolabeled EIIa; (B) with radiolabeled EIIa and unlabeled EIIa; (C) with radiolabeled EIIa and unlabeled EIIb; (D) with radiolabeled EIIb; (E) with radiolabeled EIIb and unlabeled EIIb; (F) with radiolabeled EIIb and unlabeled EIIa. Three bound complexes were observed with EIIb, while only one was observed with EIIa. Addition of unlabeled EIIb (C) competed for EIIa binding, while addition of unlabeled EIIa (F) could only partially compete for the band which was common to both EIIa and EIIb.

Cloning of E2BP cDNA. Comparison of EIIa and EIIb sequences in 19 different HBV strains (Fig. 2) showed that the EIIb direct repeat motif GGGAGGA was conserved in all strains, whereas the EIIa direct repeat motif of AAA GACT was less conserved. Furthermore, EIIb showed binding to two factors in gel retardation which could not be competed for by EIIa. Therefore, EIIb was used as the probe for screening proteins which might bind to EII. Since HuH-6 cells seem to express high levels of EII DNA-binding factors, as indicated by both gel retardation and CAT assay data, a λgt11 cDNA expression library was prepared from HuH-6 cells and used to screen for DNA-binding activity specific to EIIb. ^{32}P-labeled EIIb oligonucleotides were used as the probe for identifying plaques producing fusion proteins capable of binding to this site. Screening of 200,000 recombinant phages with the protocol described by Singh et al. (24) gave four positive plaques in the primary assay. Only one of these corresponded to a recombinant phage that could be purified to homogeneity after four rounds of screening. Examination of the recombinant phage revealed a cDNA insert of 1.15 kb, which was then subcloned into pUC18 and sequenced by using Pharmacia's T7 sequencing kit. The sequence of the cDNA revealed a consensus ATG (12) at nt 22, which formed the start of a minimum open reading frame of 984 bp (Fig. 4). The protein sequence encoded by this open reading frame yielded a predicted protein of 328 amino acid residues with a molecular size of 35.9 kDa. GenBank searches showed 86.6% homology in a 201-amino-acid overlap with the published sequence of rat heterogeneous nuclear ribonucleoprotein C (hnRNP-C) (22) and 72.0% homology in a 214-amino-acid overlap with the human nuclear protein UP2 (16) (Fig. 5). A combined Chou-Fasman (5) and GOR (8) protein structure conformation prediction for E2BP, obtained by using the computer program SeqAidII, revealed a helix-loop-helix (HLH) motif (17) which was similar to that in rat hnRNP-C.

E2BP expression in hepatoma cells. To determine the size of intact E2BP and possible precursors, total RNAs from HuH-6 and FS4 were extracted, purified by the method of Chomczynski and Sacchi (4), and used for Northern blot hybridization with E2BP cDNA. High-stringency filter-washing conditions were used as described by Striebich et al. (26). A single transcript of 1.3 kb was observed with

```
        *       *       *       *       *       *       *       *       *       *       *       *

                                                                                      GAATTCGAGGATCCGGGTACC    21

ATGGCGGCGGGCGAGCAGGAGGGAGCCATGGTGGCGGCGACACAGGGGGCAGCGGCGGCGGCGGGAAGCGGACCGGGACCGGGGGCGGAACCGCGTCTGGAGGCACCGAAGGGCAGCGCC   141
 M  A  A  G  E  Q  E  G  A  M  V  A  A  T  Q  G  A  A  A  A  A  G  S  G  P  G  P  G  A  E  P  R  L  E  A  P  K  G  S  A     40

GAGTCGGAGGGGGCGAAGATTGACGCCAGTAAGAACGAGGAGGATGAAGGCCATTCAAACTCCTCCCCACGACACTCTGAAGCAGCGACGGCACAGCGGGAAGAATGGAAAATGTTTATA   261
 E  S  E  G  A  K  I  D  A  S  K  N  E  E  D  E  G  H  S  N  S  S  P  R  H  S  E  A  A  T  A  Q  R  E  E  W  K  M  F  I     80

GGAGGCCTTAGCTGGGACACTACAAAGAAAGATCTGAAGGACTACTTTTCCAAATTTGGTGAAGTTGTAGACTGCACTCTGAAGTTAGATCCTATCACAGGGCGATCAAGGGGTTTTGGC   381
 G  G  L  S  W  D  T  T  K  K  D  L  K  D  Y  F  S  K  F  G  E  V  V  D  C  T  L  K  L  D  P  I  T  G  R  S  R  G  F  G    120

TTTGTGCTATTTAAAGAATCGGAGAGTGTAGATAAGGTCATGGATCAAAAAGAACATAAATTGAATGGGAAGGTGATTGATCCTAAAAGGGCCAAAGCCATGAAAACAAAAGAGCCGGTT   501
 F  V  L  F  K  E  S  E  S  V  D  K  V  M  D  Q  K  E  H  K  L  N  G  K  V  I  D  P  K  R  A  K  A  M  K  T  K  E  P  V    160

AAAAAAATTTTTGTTGGTGGCCTTTCTCCAGATACACCTGAAGAGAAAATAAGGGAGTACTTTGGTGGTTTTGGTGAGGTGGAATCCATAGAGCTCCCCATGGACAACAAGACCAATAAG   621
 K  K  I  F  V  G  G  L  S  P  D  T  P  E  E  K  I  R  E  Y  F  G  G  F  G  E  V  E  S  I  E  L  P  M  D  N  K  T  N  K    200

AGGCGTGGGTTATGCTTTATTACCTTTAAGGAAGAAGAACCAGTGAAGAAGATAATGGAAAAGAAATACCACAATGTTGGTCTTAGTAAATGTGAAATAAAAGTAGCCATGTCGAAGGAA   741
 R  R  G  L  C  F  I  T  F  K  E  E  E  P  V  K  K  I  M  E  K  K  Y  H  N  V  G  L  S  K  C  E  I  K  V  A  M  S  K  E    240

CAATATCAGCAACAGCAACAGTGGGGATCTAGAGGAGGATTTGCAGGAAGAGCTCGTGGAAGAGGTGGTGACCAGCAGAGTGGTTATGGGAAGGTATCCAGGCGAGGTGGTCATCAAAAT   861
 Q  Y  Q  Q  Q  Q  Q  W  G  S  R  G  G  F  A  G  R  A  R  G  R  G  G  D  Q  Q  S  G  Y  G  K  V  S  R  R  G  G  H  Q  N    280

AGCTACAAACCATATCTAAATTATTCCATTTGCAACTTATCCCCAACAGGTGGTGAAGCAGTATTTTCCAATTTGAAGATTCATTTGAATGGCTCCTGCCACCTGCTAATAGCAGTTCAA   981
 S  Y  K  P  Y  L  N  Y  S  I  C  N  L  S  P  T  G  G  E  A  V  F  S  N  L  K  I  H  L  N  G  S  C  H  L  L  I  A  V  Q    320

ACTAAATTTTTTGTATCAAGTCCCTGAATGGAAGTATGACGTTGGGTCCCTCTGAAGTTTAATTCTGAGTTCTCATTAAAAGAAATTTGCTTTCATTGTTTTATTTCTTAAAAAAAAAAA  1101
 T  K  F  F  V  S  S  P  <        328

AAAAAAAAAAAAAAAAAAAAAAAAAAAAACCATGGTACCCGGATCCTCGAATTC      1154

        *       *       *       *       *       *       *       *       *       *       *       *
```

FIG. 4. Sequences of E2BP cDNA and the protein deduced from the open reading frame. The nucleic acid sequence of E2BP cDNA is shown, with numbering starting from the *Eco*RI linker. The single-letter code of the deduced protein sequence is displayed below, with numbering starting from the initiation methionine. Underlined residues form the RNP-binding consensus sequence (R/K)GF(G/A)FVX(F/Y).

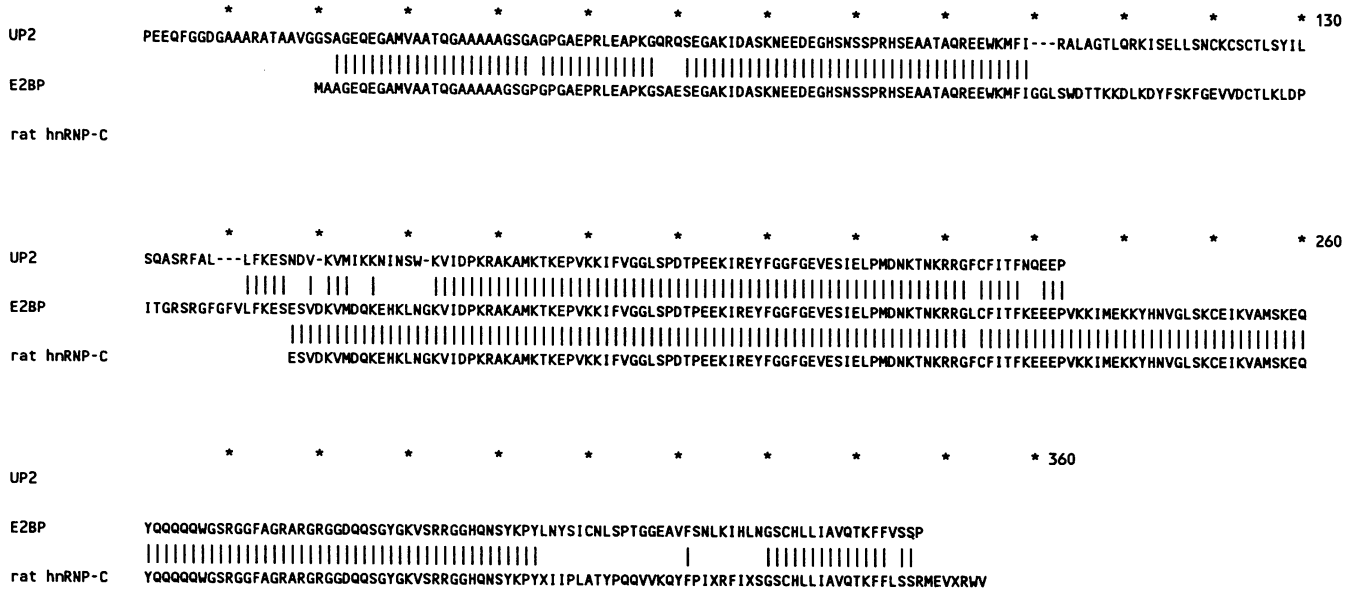

FIG. 5. Amino acid sequence comparison of E2BP with UP2 and rat hnRNP-C. Symbols: |, homology of amino acid residues; –, gap introduced into the UP2 sequence to align it with the E2BP sequence.

mRNA from HuH-6, which was absent from the human fibroblast line FS4 (Fig. 6) and in two other nonhepatic cell lines, A431 and MCF-7 (data not shown). This size indicated that the E2BP cDNA that we had isolated is a near-complete sequence, with approximately 150 to 200 bp missing from the 5′ end.

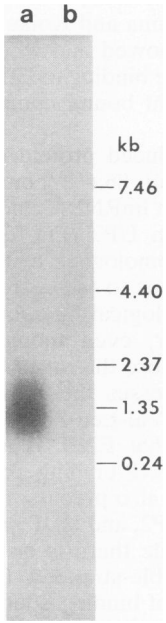

FIG. 6. Northern blot hybridization of HuH-6 and FS4 mRNAs. Total RNAs (40 μg of each) from HuH-6 (lane a) and FS4 (lane b) were electrophoresed through a 1.4% formaldehyde agarose gel and capillary transferred to a Hybond-N membrane. The blot was hybridized with ^{32}P-labeled E2BP cDNA and washed as described by Striebich et al. (26). A single transcript of 1.3 kb was observed with mRNA from HuH-6, which was absent from that of FS4.

E2BP activates transcription in vitro via EIIb. To examine whether E2BP had the ability to act as a transcriptional activator, transient DNA transfer experiments were performed in hepatic and nonhepatic cell lines. An E2BP expression vector, pE2BP-SVK/4, was constructed by joining the simian virus 40 enhancer/promoter region of the expression vector pSVK3 (Pharmacia-LKB) to the open reading frame of E2BP. pE2BP-SVK/4 was then cotransfected with p87/26-CAT and assayed for CAT activity. The presence of pE2BP-SVK/4 resulted in 1.9- to 3.2-fold increases of relative CAT activity in the nonhepatic cells FS4 and MCF-7 with respect to the activity of p87/26-CAT alone (Table 2). The hepatoma cell line HCC-T showed an increase of 2.9-fold, while HuH-6 gave only a 1.1-fold increase. The latter was most likely due to the already high endogenous expression of E2BP in HuH-6 cells, as was suggested by the data presented here as well as by Yee (36), and therefore further expression of E2BP would not result in a significant enhancement of CAT results. However, the highest level of E2BP-dependent activation observed was only 3.2-fold. This

TABLE 2. Effect of pE2BP-SVK/4 on p87/26-CAT activity

Cell line	Relative CAT activity[a]		Fold increase
	p87/26-CAT	p87/26-CAT with pE2BP-SVK/4	
Non-HCC			
FS4	0.8 (0.1)	1.5 (0.1)	1.9
MCF-7	0.6 (0.1)	1.9 (0.1)	3.2
HCC			
HCC-T	1.1 (0.1)	3.2 (0.1)	2.9
HuH-6	50.2 (1.7)	53.7 (1.9)	1.1

[a] The indicated cell lines were transfected with p87/26-CAT with or without pE2BP-SVK/4. Values are means and standard deviations (in parentheses) of triplicate experiments. Control pCAT-Promoter vector activity is 1.0.

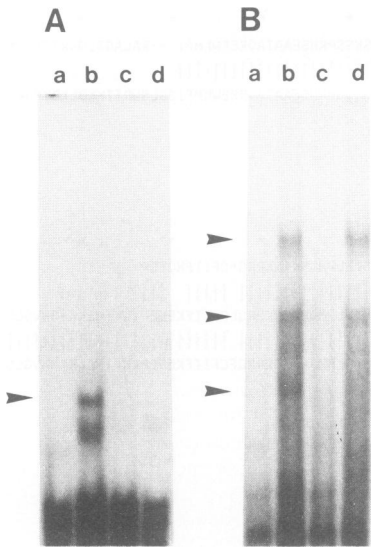

FIG. 7. Gel retardation of FS4 transfected with pE2BP-SVK/4. ^{32}P-labeled EIIa and EIIb DNAs were used for gel retardation studies with nuclear proteins from the fibroblast cell line FS4 which was transfected with pE2BP-SVK/4. (A) Lanes: a, control (untransfected FS4); b, transfected FS4 with EIIa; c, transfected FS4 with EIIa and competitor EIIa; d, transfected FS4 with EIIa and competitor EIIb. (B) Lanes: a, control lane (untransfected FS4); b, transfected FS4 with EIIb; c, transfected FS4 with EIIb and competitor EIIb; d, transfected FS4 with EIIb and competitor EIIa.

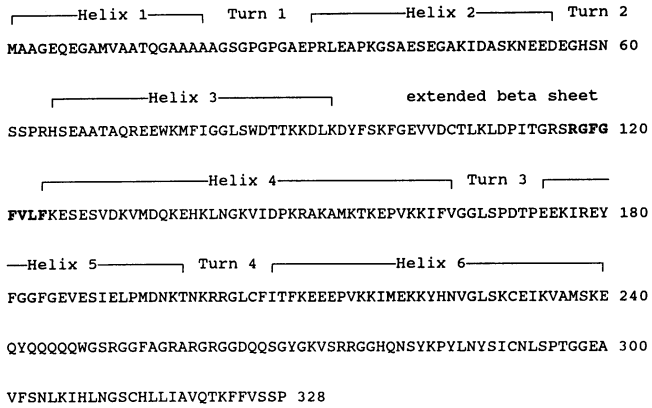

FIG. 8. Combined Chou-Fasman and GOR prediction of protein conformation for E2BP, obtained by using the computer program SeqAidII. The RNP consensus sequence is highlighted in bold.

finding implied that E2BP itself is not the sole activator of EII, but that the presence of the other two as yet unidentified factors (Fig. 3) together with E2BP is probably required for maximal activation.

From the gel retardation assays (Fig. 3), it was noted that three bands were seen with EIIb. To determine which factor was E2BP, gel retardation assays were conducted with nuclear proteins extracted from the fibroblast cell line FS4 which had been transfected with pE2BP-SVK/4. The result again showed three bands with EIIb (Fig. 7), all of which could be competed for by unlabeled EIIb and one of which could be competed for by unlabeled EIIa. This result showed that E2BP is probably present in all three bands along with other polypeptides. Untransfected FS4 did not show any binding to EIIb.

DISCUSSION

EII-binding proteins in differentiated hepatoma cell lines. From the CAT assays performed, it was determined that the DNA binding of the HBV EII sequence was highly tissue specific. Furthermore, it was observed that the CAT activity corresponded to the differentiation status of the hepatoma cell panel used. Well-differentiated hepatoma cell lines gave higher CAT enzyme activity, while poorly or dedifferentiated cell lines showed low activity levels. This difference was noted by Honigwachs et al. (10), who used a limited hepatoma cell panel and could not detect EII DNA-binding factors in the dedifferentiated cell line SK-Hep1. In our study, we noted that Mahlavu, a similarly dedifferentiated hepatoma cell line, also had minimal detectable levels of EII DNA-binding factor activity. In contrast, the human primary hepatocyte culture and the well-differentiated hepatoma lines HuH-6, HepG2, and HuH-7 had high CAT enzyme activity among the hepatoma panel, reflecting the high amounts of EII-binding factors present. HuH-6 in particular had CAT activity three- to fivefold higher than those of the human primary hepatocyte culture and HepG2, respectively. Why this is so is at present not known, but it is possible that in the transformed state of HuH-6, there is overexpression of one or more of the EII DNA-binding factors. Although HBV does not infect rodents, the enhanced CAT activity in the rat hepatoma cell line MH_1C_1 also suggests the existence of a rat homolog for this EII-binding protein.

Comparison between E2BP and related factors. Using ^{32}P-labeled EIIb DNA to screen a HuH-6 λgt11 cDNA expression library, we have isolated the cDNA for a liver-specific factor designated E2BP that binds to the EIIa and EIIb sites of HBV EII. The factors that bind to these sites have not been characterized, although their presence has been shown by Yaginuma and Koike (34) and Yee (36). Gel retardation analysis showed that while this factor bound to both EIIa and EIIb, the binding to EIIb resulted in two other higher-molecular-weight bound complexes which were not observed with EIIa.

Analysis of the deduced protein sequence of E2BP revealed that E2BP had 86.6% homology with published partial sequences of rat hnRNP-C and 72.0% homology with human nuclear protein UP2 (Fig. 5). The N-terminal sequence of E2BP is homologous to that of UP2, while the latter half is homologous to the corresponding region of rat hnRNP-C. The physiological function of E2BP is therefore not immediately clear, even though our present cDNA sequence resembles more the rat hnRNP-C sequence. Furthermore, E2BP possesses the RNP consensus sequence Arg-Gly-Phe-Gly-Phe-Val-Leu-Phe (28), highlighted in bold in Fig. 8, indicating that E2BP is a member of the RNP family. Although hnRNP-binding proteins recognizes primarily RNA, they are also precursors for the helix-destabilizing proteins UP1, UP2, and HDP, which are DNA-binding proteins (19, 22). While there is no known precedence of hnRNPs binding double-stranded DNA, E2BP, like rat hnRNP-C, is capable of binding double-stranded DNA, and the ability to do so may be the result of a unique structural conformation apart from the RNA-binding domain. Amino acid composition analysis of E2BP showed that the two basic amino acids, Lys and Arg, comprise 10.7 and 4.6% of the residues, giving it a net positive charge. This structure is similar to that of the basic regions used by sequence-specific DNA-binding proteins such as mammalian CTF/NF1 (15). A

combined Chou-Fasman and GOR protein secondary structure analysis of E2BP predicted extensive helical domains separated by beta-turn regions (Fig. 8). Of interest is that the sequences of the first (Gly-Ser-Gly-Pro-Gly-Pro-Gly) and third (Gly-Gly-Leu-Ser-Pro-Asp) turn regions are consistent with a loop structure that frequently uses the amino acids Gly, Pro, Asn, Asp, and Ser (14). HLH structures are found in the DNA-binding domains of the daughterless, MyoD, Myc, Id, E12, and E47 sequence-specific DNA-binding proteins (13, 17). This finding suggests that the HLH first turn region of E2BP, at the N-terminal portion, may be involved in the binding to EII. The third turn region is situated proximal to the C-terminal portion of the E2BP sequence, suggesting that it may not be involved in nucleic acid binding but that it may be required for dimerization of the protein as was reported for MyoD, Myc, and the immunoglobulin enhancer-binding proteins (17). It is also possible that the missing 5′ sequence in our cDNA clone presents a basic motif for the binding of DNA, as suggested by Kadesch (11). However, CAT assays and gel retardation assays of cell lines which were transiently transfected with an E2BP expression vector (Table 2 and Fig. 7) showed that the protein product of the cloned E2BP sequence could bind to and activate EII, implying that the essential DNA-binding motif is retained in the sequence cloned. There was no apparent similarity of the nucleic acid-binding region of E2BP to other DNA-binding domains such as homeoboxes, zinc fingers, or leucine zippers.

Tissue specificity of E2BP and hepatotrophicity of HBV. A putative 1.3-kb E2BP mRNA was detected in the hepatoma cell line HuH-6 but was absent in the human fibroblast line FS4 and in two other nonhepatic cell lines, A431 and MCF-7 (data not shown). Although UP2 and hnRNP-C would be expected to be ubiquitous, E2BP was not observed to be so. The tissue-specific expression of E2BP suggests that it may be a variant of the hnRNP family which is localized to hepatocytes cells, which in turn would account for the liver specificity of HBV infections.

Results of the gel retardation experiments showed that at least three factors bind to the EIIb motif. A possible explanation is either that E2BP forms multimeric complexes with other proteins in each of these bands or that E2BP binding facilitates binding of other proteins to itself or to adjacent sites. As mentioned earlier, E2BP itself may not be the sole activator of EII, and complexing with the other factors is most likely required for maximal EII activation. The results shown in Fig. 3 suggest that the EIIa sequence favors monomeric binding, while the EIIb sequence was capable of multimeric binding. One possibility is that E2BP binds to the region of sequence overlap between EIIa and EIIb, and in the case of EIIb, the binding might enhance the binding of the other two factors to the EIIb sequence. EIIa, which does not have the EIIb sequence motifs, would not have these two other factors binding to it; hence, only monomeric binding is observed. Another possibility is that the proximity of the two EIIb direct repeat motifs enables dimerization and complex formation with the other two factors, which is stable to the extent that it cannot be competed for by EIIa. EIIa direct repeat motifs, on the other hand, are further apart and may not permit the formation of such a complex; hence, the binding can easily be competed for by EIIb. Furthermore, E2BP may have a helix-destabilizing function, given its homology with the helix-destabilizing protein UP2 (16). If, indeed, E2BP belongs to the group of developmentally relevant proteins (e.g., MyoD and Myc) having the HLH motif, then its binding pattern in the gel retardation is more easily understood, as these proteins can bind DNA either as monomers, homodimers, or heterodimers (11). Furthermore, it would explain the tissue specificity in that E2BP may function by forming heterodimers between tissue-specific (i.e., E2BP) and ubiquitously expressed HLH proteins. Further work will be done to determine the action of E2BP, to determine whether the tissue specificity of HBV infection is indeed a result of E2BP acting as part of a multimeric complex, and, if it is, to identify the other factors involved. Given the very strong sequence homology with rat hnRNP-C, it is very likely that E2BP is a member of the RNP family, but its restricted expression in hepatocytes suggest that it is possible for E2BP to have a distinct and different role in vivo, unrelated to that of the known hnRNP proteins.

To our knowledge, this is the first report of the interaction of an hnRNP-binding protein with a viral enhancer element. The idea that tissue-specific expression of such nuclear proteins may be one of the mechanisms for viral tropism is certainly intriguing.

ACKNOWLEDGMENTS

We thank the following for generous gifts of hepatoma cell lines: C. Chang, Veterans General Hospital, Taipei, Taiwan (HA22T/VGH); T. Morizane, Keio University (HCC-M and HCC-T); B. B. Knowles, Wistar Institute (Hep3B and HepG2); N. Huh, Cancer Institute, Tokyo, Japan, (huH-1 and huH-4); I. Doi, Okayama University (HuH-6); H. Nakabayashi, Okayama University (HuH-7); T. Murakami, Kurume University (KMCH-1); O. Prozeskey, South Africa Medical Research Council (Mahlavu); J. J. Alexander, Virus Cancer Research Unit, Johannesburg, South Africa (PLC/PRF/5); J. Fogh, Sloan-Kettering Institute (SK-Hep1); and M. J. Tong, Huntington Memorial Hospital, Pasadena, Calif. (Tong/HCC).

This work was supported by grant RP890308 from the National University of Singapore.

REFERENCES

1. **Antonucci, T. K., and W. J. Rutter.** 1989. Hepatitis B virus (HBV) promoters are regulated by the HBV enhancer in a tissue-specific manner. J. Virol. **63**:579–583.
2. **Blum, H. E., W. Gerok, and G. N. Vyas.** 1989. The molecular biology of hepatitis B virus. Trends Genet. **5**:154–158.
3. **Cattaneo, R., H. Will, and H. Schaller.** 1984. Hepatitis B virus transcription in the infected liver. EMBO J. **3**:2191–2196.
4. **Chomczynski, P., and N. Sacchi.** 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**:156–159.
5. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of protein conformation. Annu. Rev. Biochem. **47**:251–276.
6. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
7. **Fujiyama, A., A. Miyanohara, C. Nozaki, T. Yoneyama, N. Ohtomo, and K. Matsubara.** 1983. Cloning and structural analyses of hepatitis B virus DNAs, subtype adr. Nucleic Acids Res. **11**:4601–4610.
8. **Garnier, J., D. J. Osguthorpe, and B. Robson.** 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120**:97–120.
9. **Gorman, C.** 1985. High efficiency gene transfer into mammalian cells, p. 321–325. *In* D. M. Glover (ed.), DNA cloning—a practical approach, vol. 2. IRL Press, Oxford.
10. **Honigwachs, J., O. Faktor, R. Dikstein, Y. Shaul, and O. Laub.** 1989. Liver-specific expression of hepatitis B virus is determined by the combined action of the core gene promoter and the enhancer. J. Virol. **63**:919–924.
11. **Kadesch, T.** 1992. Helix-loop-helix proteins in the regulation of immunoglobulin gene transcription. Immunol. Today **13**:31–36.

12. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
13. **Kreider, B. L., R. Benezra, G. Rovera, and T. Kadesch.** 1992. Inhibition of myeloid differentiation by the helix-loop-helix protein Id. Science **255**:1700–1702.
14. **Lesczynski, F. D., and G. D. Rose.** 1986. Loops in globular proteins: a novel category of secondary structure. Science **234**:849–855.
15. **Mermod, N., E. A. O'Neill, T. J. Kelly, and R. Tjian.** 1989. The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell **58**:741–753.
16. **Merrill, B. M., M. B. LoPresti, K. L. Stone, K. L., and K. R. Williams.** 1986. High pressure liquid chromatography purification of UP1 and UP2, two related single-stranded nucleic acid-binding proteins from calf thymus. J. Biol. Chem. **261**:878–883.
17. **Murre, C., P. Schonleber-McCaw, and D. Baltimore.** 1989. A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD*, and *myc* proteins. Cell **56**:777–783.
18. **Rhodes, D.** 1989. Analysis of sequence-specific DNA-binding proteins, p. 177–198. *In* T. E. Creighton (ed.), Protein function—a practical approach. IRL Press, Oxford.
19. **Riva, S., C. Morandi, P. Tsoulfas, M. Pandolfo, G. Biamonti, B. Merrill, K. R. Williams, G. Multhaup, K. Beyreuther, H. Werr, B. Henrich, and K. P. Schafer.** 1986. Mammalian single-stranded DNA binding protein UP1 is derived from the hnRNP core protein A1. EMBO J. **5**:2267–2273.
20. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
21. **Seed, B., and J. Y. Sheen.** 1988. A simple phase-extraction assay for chloramphenicol acetyltransferase activity. Gene **67**:271–277.
22. **Sharp, Z. D., K. P. Smith, Z. Cao, and S. Helsel.** 1990. Cloning of the nucleic acid-binding domain of the rat hnRNP C-type protein. Biochim. Biophys. Acta **1048**:306–309.
23. **Shaul, Y., W. J. Rutter, and O. Laub.** 1985. A human hepatitis B viral enhancer element. EMBO J. **4**:427–430.
24. **Singh, H., J. H. LeBowitz, A. S. Baldwin, Jr., and P. A. Sharp.** 1988. Molecular cloning of an enhancer binding protein: isolation by screening of an expression library with a recognition site DNA. Cell **52**:415–423.
25. **Spandau, D. F., and C.-H. Lee.** 1988. Trans-activation of viral enhancers by the hepatitis B virus X protein. J. Virol. **62**:427–434.
26. **Striebich, C. C., R. M. Miceli, D. H. Schulze, G. Kelsoe, and J. Cerny.** 1990. Antigen-binding repertoire and IgH chain gene usage among B cell hybridomas from normal and autoimmune mice. J. Immunol. **144**:1857–1865.
27. **Sureau, C., J.-L. Romet-Lemonne, J. I. Mullins, and M. Essex.** 1986. Production of hepatitis B virus by a differentiated human hepatoma cell line after transfection with cloned circular HBV DNA. Cell **47**:37–47.
28. **Swanson, M. S., T. Y. Nakagawa, K. LeVan, and G. Dreyfuss.** 1987. Primary structure of human nuclear ribonucleoprotein particle C proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. Mol. Cell. Biol. **7**:1731–1738.
29. **Tay, N., S. H. Chan, and E. C. Ren.** 1990. Detection of integrated hepatitis B virus DNA in hepatocellular carcinoma cell lines by nonradioactive in situ hybridization. J. Med. Virol. **30**:266–271.
30. **Tsurimoto, T., A. Fujiyama, and K. Matsubara.** 1987. Stable expression and replication of hepatitis B virus genome in an integrated state in a human hepatoma cell line transfected with the cloned viral DNA. Proc. Natl. Acad. Sci. USA **84**:444–448.
31. **Vannice, J. L., and A. D. Levinson.** 1988. Properties of the human hepatitis B virus enhancer: position effects and cell-type nonspecificity. J. Virol. **62**:1305–1313.
32. **Wang, Y., P. Chen, X. Wu, A.-L. Sun, H. Wang, Y.-A. Zhu, and Z.-P. Li.** 1990. A new enhancer element, ENII, identified in the X gene of hepatitis B virus. J. Virol. **64**:3977–3981.
33. **Will, H., W. Reiser, E. Weimer, M. Pfaff, R. Buscher, R. Sprengel, R. Cattaneo, and H. Schaller.** 1987. Replication strategy of human hepatitis B virus. J. Virol. **61**:904–911.
34. **Yaginuma, K., and K. Koike.** 1989. Identification of a promoter region for 3.6-kilobase mRNA of hepatitis B virus and specific cellular binding protein. J. Virol. **63**:2914–2920.
35. **Yaginuma, K., Y. Shirakata, M. Kobayashi, and K. Koike.** 1987. Hepatitis B virus (HBV) particles are produced in a cell culture system by transient expression of transfected HBV DNA. Proc. Natl. Acad. Sci. USA **84**:2678–2682.
36. **Yee, J. K.** 1989. A liver-specific enhancer in the core promoter region of human hepatitis B virus. Science **246**:658–661.