JOURNAL OF VIROLOGY, Dec. 1996, p. 8719–8726 Vol. 70, No. 12
0022-538X/96/$04.00+0

Promoters for Synthesis of the Pre-C and Pregenomic mRNAs of Human Hepatitis B Virus Are Genetically Distinct and Differentially Regulated

XIANMING YU AND JANET E. MERTZ*

McArdle Laboratory for Cancer Research, University of Wisconsin Medical School, Madison, Wisconsin 53706-1599

Received 17 June 1996/Accepted 19 August 1996

Two similar, yet functionally distinct genomic RNAs are transcribed from the DNA genome of the human hepatitis B virus. The pre-C RNAs encode the precore protein which is proteolytically processed to yield e antigen. The pregenomic RNAs encode both the nucleocapsid protein and reverse transcriptase and serve as the templates for viral DNA replication. To determine whether synthesis of these two RNAs is directed from a single or a closely spaced pair of promoters, we introduced point and insertion mutations into the basal elements of the promoter that directs their synthesis. Transcription from these mutants was examined both in cell-free transcription systems derived from hepatoma (HepG2) and nonliver (HeLa) cell lines and by transient transfection of hepatoma cell lines (Huh7 and HepG2). The data from these experiments indicated that synthesis of the pre-C and pregenomic RNAs is directed by two distinct promoters and that the basal elements of these two promoters partially overlap, yet are genetically separable, with each consisting of its own transcriptional initiator and a TATA box-like sequence situated approximately 25 to 30 bp upstream of its sites of initiation. A 15-bp insertion was found to be sufficient to physically separate these two promoters. Furthermore, these two promoters can be differentially regulated, with the transcriptional activator Sp1 specifically activating transcription from the pregenomic promoter and the hepatocyte nuclear factor 4 specifically repressing transcription from the pre-C promoter. Thus, we conclude that the promoters used in synthesis of the pre-C and pregenomic mRNAs are genetically distinct and separately regulated.

Human hepatitis B virus (HBV), a member of the hepadnavirus family, is a pathogen which causes viral hepatitis. Chronic hepatitis B can lead to liver cirrhosis and hepatocellular carcinoma (13, 36). The partially double-stranded DNA genome of HBV contains four open reading frames (ORFs) encoding the S proteins (envelope proteins), precore protein, nucleocapsid protein (C), reverse transcriptase (P), and X protein. The HBV genes are transcribed by the host RNA polymerase II. All of the viral transcripts utilize a common polyadenylation signal located within the C protein-coding region. The viral RNAs are subdivided into two classes, genomic and subgenomic, according to their sizes (reviewed in references 13, 33, and 36). The 5′ ends of the 3.5-kb genomic RNAs are situated within 30 bases of the beginning of the pre-C–C ORF and map to nucleotides (nt) 1785 to 1786, 1791 to 1793 (pre-C RNAs), and 1818 to 1820 (pregenomic RNAs) (42). These two species of genomic RNAs were also identified in (i) cultured hepatoma cells stably or transiently transfected with whole HBV genomic DNA or subclones of fragments of the viral DNA (18, 27, 31, 43), (ii) liver cells infected with HBV (40), and (iii) transgenic mice containing HBV (11).

The pre-C–C ORF contains two in-frame translation initiation signals separated by 28 codons, thus encoding both precore protein (212 amino acids, adr subtype), the precursor for the viral e antigen, and C protein (183 amino acids, adr subtype) (reviewed in reference 33). The pre-C RNAs have 5′ ends situated upstream of the first AUG codon (nt 1816 to 1818) (Fig. 1A). Microinjection of these RNAs into *Xenopus* oocytes results in the synthesis of precore protein, but little or no C protein (34). The pre-C RNAs also do not serve as templates for HBV DNA replication, presumably because translation of the precore ORF interferes with recognition of the encapsidation signal by the viral C and P proteins (25). On the other hand, the pregenomic RNAs have 5′ ends situated downstream of the first AUG codon (Fig. 1A). Thus, they can encode the C and P proteins, but not the precore protein. The pregenomic RNAs also serve as templates for viral DNA synthesis after encapsidation into core particles (25). Therefore, it is likely that the virus needs to differentially regulate synthesis of these two species of genomic RNAs.

Yuh et al. (43) showed by analysis of deletion mutants transiently transfected into hepatoma cells that the basic C promoter maps within an approximately 100-bp fragment (nt 1744 to 1851). This DNA fragment is sufficient for accurate initiation of both pre-C and pregenomic RNAs in vivo. Sequences located immediately 5′ of the basic C promoter can positively modulate this promoter in a liver-specific manner through *cis*-acting elements in enhancer II (Fig. 1A) (16, 43). Sequences located further upstream can negatively modulate this promoter through a negative regulatory element (15, 16, 23). The upstream regulatory region (nt 1403 to 1743), which contains both the positive and negative *cis*-acting elements, has been shown, as a whole, to stimulate transcription of the pre-C and pregenomic RNAs in hepatoma cells (42a). Moreover, transcription of the pre-C and pregenomic RNAs appears to be coordinately regulated by *cis*-acting elements in the enhancer II region (43). It was reported that sequences within enhancer I (Fig. 1A) can also activate the basic C promoter (44, 45). The strict liver cell specificity of this promoter is probably determined via the combined action of multiple *cis*-acting elements and *trans*-acting factors (1, 18).

We have been identifying *cis*-acting sequences and *trans*-acting factors involved in synthesis of the pre-C and pre-

* Corresponding author. Phone: (608) 262-2383. Fax: (608) 262-2824. Electronic mail address: mertz@oncology.wisc.edu.

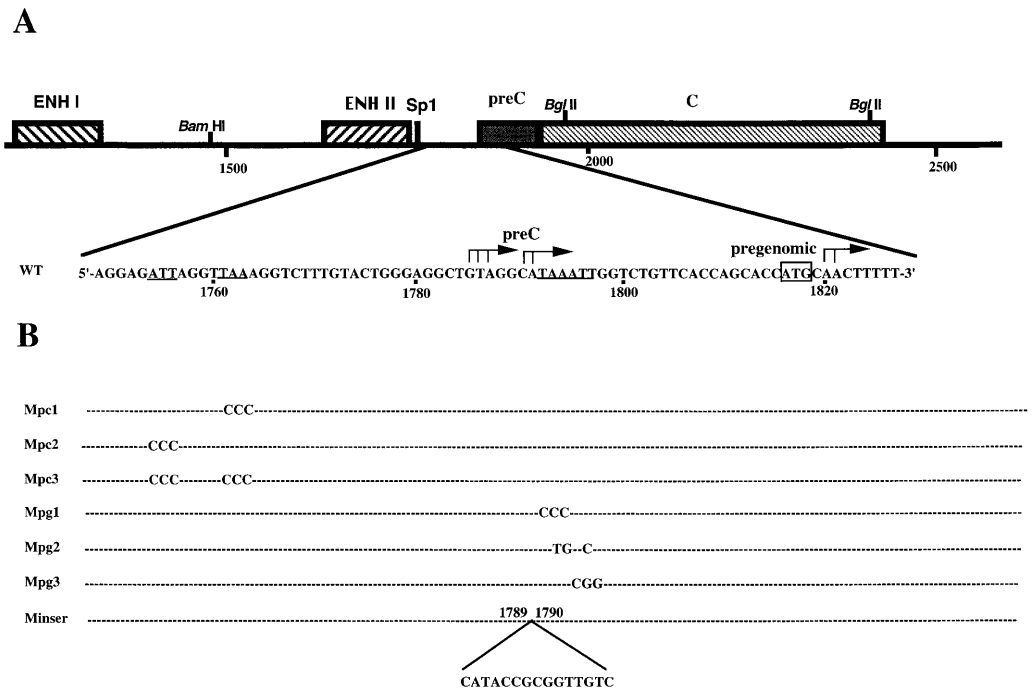

FIG. 1. (A) Schematic diagram of the HBV basic C promoter, its upstream regulatory region (nt 1403 to 1851), and nearby surrounding sequences. The enhancers (ENH I and ENH II), the pre-C and C ORFs, and the two Sp1 binding sites are indicated by labeled rectangular boxes. The nucleotide sequence of part of the basic C promoter is also shown. The arrows indicate the locations of the 5′ ends of the pre-C and pregenomic RNAs synthesized in Huh7 cells. The nucleotides at which point mutations were introduced are underlined. The AUG codon of the pre-C ORF is indicated by a box. (B) Sequence alterations in the mutants studied here. The dashes represent nucleotides identical to those of wild-type HBV (shown directly above in panel A).

genomic RNAs. In this report, we show that the basal elements of the promoters for the pre-C and pregenomic RNAs are distinct and genetically separable and that the synthesis of these RNAs can be differentially regulated in *trans* by positively and negatively acting factors.

MATERIALS AND METHODS

Plasmid constructs. Most of the plasmids used in the cell-free transcription and in vivo transfection assays are derivatives of plasmid pGL-2 (Promega), a promoterless luciferase expression vector. To construct plasmid pWT, the *Bam*HI-*Bgl*II fragment of HBV DNA (nt 1403 to 1991) was amplified by PCR from pADR1 (41), a plasmid in which a complete HBV genome (subtype adr) had been cloned into the *Bam*HI site of pBR322. This PCR-generated fragment was then inserted between the *Sma*I and *Bgl*II sites of pGL-2. Plasmids pMpc1-3, pMpg1-3, and pMinser are derivatives of pWT containing point and insertion mutations (Fig. 1); they were constructed by a two-step, PCR-based mutagenesis method (2). Plasmids pWT′, pMpc1′, pMpg1′, and pMpg2′ are derivatives of pWT, pMpc1, pMpg1, and pMpg2, respectively, from which the upstream regulatory region (nt 1403 to 1729) has been deleted; they were constructed by PCR-based amplification of the DNA fragments (nt 1730 to 1991) from the corresponding mutant plasmids and subcloning into pGL-2. All promoter mutations were confirmed by DNA sequence analysis.

Plasmid pADR2, containing a terminal redundancy of the HBV genome, was derived from pADR1 by insertion of the 589-bp *Bam*HI-*Bgl*II fragment of HBV DNA into one of the two *Bam*HI sites in pADR1. To ensure that infectious virus was not produced in transfected cells, a portion of the C protein-coding region was removed by deletion of the *Bgl*II fragment (nt 1987 to 2432). Plasmid pCDMHNF4, a gift from David Moore, contains a cDNA clone of rat hepatocyte nuclear factor 4 (HNF4) in the expression vector pCDM8. Plasmid pEQ176, a gift from Kim Mansky, contains the *Escherichia coli* β-galactosidase gene expressed from a cytomegalovirus immediate-early (CMV-IE) promoter (30).

Cell-free transcription assays. HeLa cell nuclear extract was kindly provided by Richard Kraus and Theresa Eisenbraun in our laboratory. The preparation of nuclear extract from suspension cultures of cells (18 to 20 mg of protein per ml) was performed as reported previously (8, 38, 39). HepG2 nuclear extracts were also prepared by this method, except the HepG2 cells were grown in 150-mm-diameter tissue culture dishes to 80% confluence, and the final protein concentration of the nuclear extract was 6 to 8 mg of protein per ml.

The cell-free transcription assays were performed as described previously (38, 39) with 0.1 μg of supercoiled plasmid DNA and 8 μl of nuclear extract in 25-μl reaction mixtures. The relative quantitation and locations of the 5′ ends of the viral pre-C and pregenomic RNAs were determined as described previously (38, 39) with a synthetic oligonucleotide corresponding to HBV nt 1953 to 1931 used as the primer. To study the effects of Sp1 on transcription from the pre-C and pregenomic promoters in cell-free systems, the template DNA was preincubated at 26°C for 15 min with 1 μl of preimmune serum, 1 or 2 μl of recombinant human Sp1 (Promega; 1 footprint unit/μl), or 2 μl of Sp1 plus 1 μl of anti-Sp1 antibody (Santa Cruz Biotech) prior to the addition of the nuclear extract.

In vivo transfection assays. Hepatoma cell lines Huh7 and HepG2 were grown in Dulbecco's modified Eagle's medium-F12 medium supplemented with 10% fetal bovine serum. Transfections were done according to the calcium phosphate precipitation method (28). Each 100-mm-diameter tissue culture dish of cells was transfected with a total of 20 μg of plasmid DNA. In the cotransfection experiments, 10 μg of pCDMHNF4 or control plasmid pCDM8 (Invitrogen Corp.) and 6 μg of plasmid pGL2, pWT, or pADR2 were used. Plasmid pEQ176 (2 μg) was included in each transfection mixture as an internal control. Forty-eight hours posttransfection, total cellular RNA was isolated by lysis with sodium dodecyl sulfate (SDS)-EDTA. In brief, the cells were harvested in a buffer containing 0.5 M NaCl, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1% SDS, and 200 μg of proteinase K per ml. The lysates were passed through a 26-gauge needle several times to disrupt the chromosomal DNA and incubated at 37°C for 1 h. To select for polyadenylated mRNA, oligo(dT) cellulose (Collaborative Biochemical Products) was added to the lysate, and the mixture was incubated at room temperature for 1 h with rotation. After thorough washing of the oligo(dT) cellulose with lysis buffer, the mRNA was eluted with Tris-EDTA buffer. Total cellular RNA was isolated from HepG2 cells by the acid guanidinium thiocyanate-phenol-chloroform extraction method (6) before oligo(dT) selection. The mRNA obtained from one-third of a 100-mm-diameter dish of cells was used for each primer extension reaction with the oligonucleotide 5′-GTTTTCCCAGTC ACGAC-3′ serving as the primer for the detection of β-galactosidase mRNA.

RESULTS

Accurate transcription from the pre-C and pregenomic promoters in a cell-free system. To help us better understand the functions of *cis*-acting elements in the basic C promoter region, we first developed a hepatoma cell-derived cell-free transcription system. Nuclear extract was prepared from HepG2 cells.

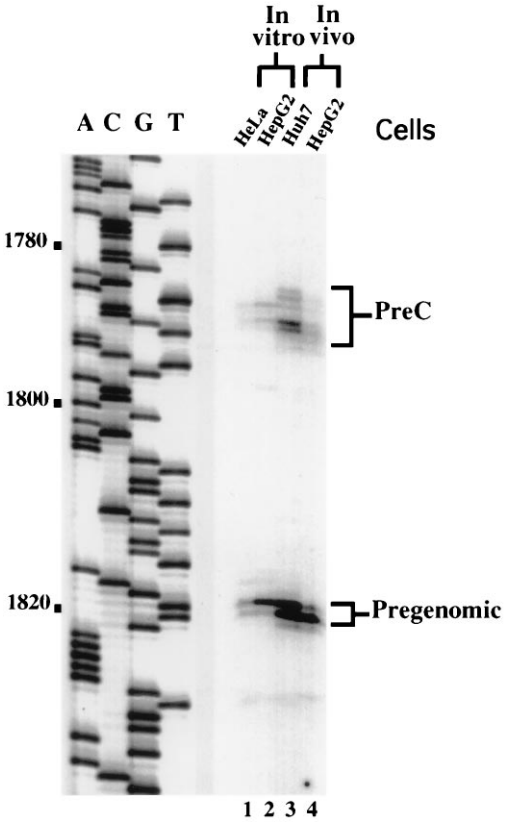

FIG. 2. High-resolution mapping of the 5′ ends of the pre-C and pregenomic RNAs synthesized in vitro and in vivo from pWT. Shown here are the products of primer extension analysis resolved by electrophoresis in a 7 M urea–6% polyacrylamide gel. The cDNAs corresponding to pre-C and pregenomic RNAs are indicated by brackets. Lanes: 1 and 2, cDNAs of RNAs synthesized in cell-free transcription systems made from HeLa and HepG2 nuclear extracts, respectively; 3 and 4, cDNAs of mRNAs accumulated by 48 h after transfection of pWT into Huh7 and HepG2 cells, respectively. The size markers in the left half of this figure consisted of dideoxy sequencing reactions performed with the same primer and pMpg2 DNA as the template.

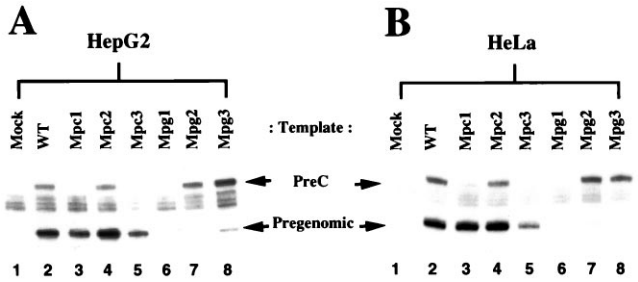

FIG. 3. Effects of promoter substitution mutations on synthesis of the pre-C and pregenomic RNAs in cell-free transcription systems made from HepG2 (A) and HeLa (B) nuclear extracts. The 5′ ends of the RNAs synthesized from the indicated templates were determined by primer extension analysis. Shown here is a 7 M urea–8% polyacrylamide gel with the cDNAs corresponding to the pre-C and pregenomic RNAs indicated by arrows. Lanes: 1, control plasmid pGL-2; 2 to 8, wild-type (WT) and mutant plasmids. The multiple bands present between the pre-C and pregenomic RNAs are not HBV specific and were observed with only some preparations of nuclear extracts.

As a control, we also prepared nuclear extract from HeLa cells. These extracts were used in a cell-free transcription system to synthesize pre-C and pregenomic transcripts from pWT, a plasmid which contains the nt 1403 to 1991 region of the wild-type HBV genome upstream of the luciferase gene. The precise locations of the 5′ ends of the RNAs were determined by primer extension analysis (Fig. 2, lanes 1 and 2). As a control, the hepatoma cell lines Huh7 and HepG2 were transiently transfected with pWT, and the polyadenylated RNAs synthesized in these cells were analyzed in parallel (lanes 3 and 4).

Consistent with the findings of Yaginuma et al. (42), the 5′ ends of the RNAs synthesized in transiently transfected Huh7 cells were found to map to nt 1785 to 1787, 1790 to 1791 (pre-C RNAs), and 1820 to 1821 (pregenomic RNAs) (Fig. 2, lane 3). Furthermore, the 5′ ends of the pre-C and pregenomic RNAs synthesized in HepG2 cells also mapped to these locations within the basic C promoter region (lane 4). The 5′ ends of the pre-C and pregenomic RNAs synthesized in the cell-free transcription systems derived from HepG2 and HeLa cells were found to map to most of the same sites, as did the in vivo-synthesized RNAs (lanes 1 and 2 versus lanes 3 and 4). Thus, we conclude that (i) our cell-free transcription system accurately recognizes the basic C promoter, and (ii) the selection of these transcription initiation sites is not liver cell specific.

Basic C promoter contains two sets of partially overlapping basal promoter elements. To determine if synthesis of the pre-C and pregenomic RNAs is under the regulation of distinct promoters, we attempted via mutagenesis to disrupt the transcription of one mRNA species while not affecting the other. The sequences changed in mutants Mpc1, Mpc2, and Mpc3 were selected because there are two AT-rich sequences, 5′-AGATTA-3′ (nt 1752 to 1757) and 5′-TTAAA-3′ (nt 1760 to 1764), which (i) are located approximately 30 to 35 bp 5′ of the pre-C RNA start sites (Fig. 1A) and (ii) can bind recombinant TATA-binding protein (TBP) on the basis of both footprint (4) and competitive gel mobility-shift assays (data not shown). The effects of these base substitution mutations on RNA transcription were analyzed with our HeLa- and HepG2-derived in vitro transcription systems (Fig. 3). Mutant Mpc1 (CCC in nt 1761 to 1763) made no detectable pre-C RNAs, yet made pregenomic RNAs at levels similar to those of the wild type (Fig. 3A and B, lane 3 versus lane 2). In contrast, the same base substitutions introduced into the 5′-most AT-rich sequence in mutant Mpc2 (CCC in nt 1754 to 1756) led to the synthesis of wild-type levels of pre-C and pregenomic RNAs (Fig. 3A and B, lane 4 versus lane 2). As expected, the mutant Mpc3, containing both triple substitution mutations, made no detectable pre-C RNAs (Fig. 3A and B, lane 5). Therefore, we conclude that the AT-rich sequence 5′-TTAAA-3′ at nt 1760 to 1764, situated approximately 25 to 30 bp 5′ of the pre-C RNA, is an essential *cis*-acting element for synthesis in vitro of the pre-C RNAs but not for synthesis of the pregenomic RNAs.

The second region of the basic C promoter selected for mutagenesis was the AT-rich sequence (5′-CATAAATT-3′, nt 1790 to 1797), which we predicted from sequence analysis (21) might function as both (i) a TBP-binding site for initiation of transcription 25 to 30 bp downstream at the start sites of the pregenomic RNAs and (ii) the initiator (Inr) sequence for some of the pre-C RNAs (Fig. 1). Analysis with our cell-free transcription systems of the mutants Mpg1, Mpg2, and Mpg3, containing base substitution mutations in this second region (Fig. 1B), is also shown in Fig. 3. Whereas the mutation in mutant Mpg1 led to no detectable synthesis of either the pre-C or pregenomic RNAs, the mutations in mutants Mpg2 and Mpg3 only severely reduced synthesis of the pregenomic RNAs, leaving synthesis of the pre-C RNAs unaffected (Fig. 3A and B, lanes 6 to 8 versus lane 2).

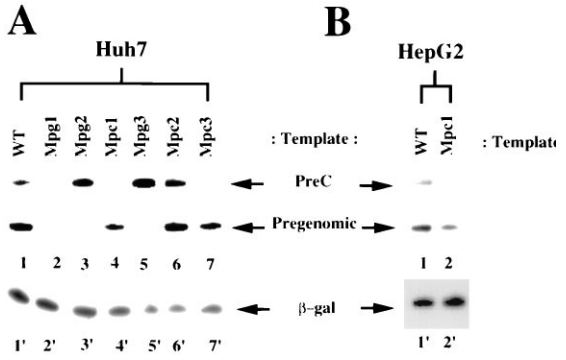

FIG. 4. Effects of promoter substitution mutations on synthesis of the pre-C and pregenomic RNAs in transfected hepatoma cell lines. Huh7 (A) and HepG2 (B) cells were transfected with the indicated plasmids. Forty-eight hours later, the cellular mRNA was isolated. The 5′ ends of the viral, polyadenylated RNAs were determined by primer extension followed by electrophoresis in a 7 M urea–8% polyacrylamide gel. Shown here are the resulting autoradiograms. Concurrent analysis of β-galactosidase (β-gal) RNA synthesized from the co-transfected plasmid pEQ176 served as an internal control.

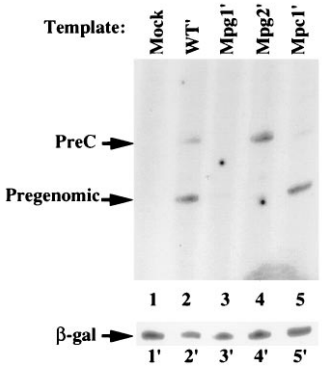

FIG. 5. Deletion of the upstream regulatory region does not affect the phenotypes of the pre-C and pregenomic promoter mutants. The experiment was performed as described in the legend to Fig. 4, except that the Huh7 cells were transfected with the plasmids pGL-2, pWT′, pMpg1′, pMpg2′, and pMpc1′, respectively.

Taken together, these data indicate that the basic C promoter contains *cis*-acting basal promoter elements that independently direct the synthesis of the pre-C and pregenomic RNAs. Nevertheless, the sequence altered in mutant Mpg1 is part of the elements of both promoters. Thus, we conclude that these promoters physically overlap in wild-type HBV.

Also noteworthy is our finding that the phenotypes of these mutants in both the HeLa and HepG2 cell-derived cell-free transcription systems were similar (Fig. 3A versus B). Thus, we further conclude that no liver-specific transcription factors are required to enable proper selection of transcription initiation sites from either of these two promoters in vitro.

To determine whether the pre-C and pregenomic promoters are functionally distinct in vivo as well, these six mutants were also introduced into the hepatoma cell line Huh7 by transfection. Forty-eight hours later, total cellular polyadenylated RNA was isolated from the cells and analyzed for HBV C promoter-specific transcripts by the primer extension technique (Fig. 4). Once again, the phenotypes of these mutants could be divided into four groups: those defective in pre-C RNA synthesis (mutants Mpc1 and Mpc3), those defective in pregenomic RNA synthesis (mutants Mpg2 and Mpg3), those defective in both (mutant Mpg1), and those competent for both (mutant Mpc2) (cf. Fig. 3 and 4A). To confirm that the nt 1760 to 1764 5′-TTAAA-3′ sequence is an essential *cis*-acting element for pre-C RNA synthesis in hepatoma cells, mutant Mpc1 was also introduced into HepG2 cells. This mutant was found to be defective in pre-C RNA synthesis in this cell line as well (Fig. 4B). Therefore, we conclude that synthesis of the pre-C and pregenomic RNAs is regulated by two separate promoters both in vivo and in vitro.

All six of these mutant promoters contain the upstream regulatory region (nt 1403 to 1743), which we have found to stimulate transcription of the pre-C and pregenomic RNAs. To examine whether this upstream region might also play a role in determining the start sites of these two RNA species in vivo, we deleted the upstream regulatory region from mutants in each of the phenotypic groups mentioned above. Analysis of the RNAs synthesized in vivo from each of these double mutants indicated, as expected, that while the overall levels of viral RNA synthesis were greatly reduced, promoter functions remained unchanged (Fig. 5). Thus, we conclude that the upstream regulatory region does not affect the selection of the initiation sites used in synthesis of the pre-C and pregenomic RNAs.

The pre-C and pregenomic promoters can be physically separated. If the pre-C and pregenomic promoters are truly independent promoters, it should be possible to physically separate them. This was accomplished by the construction of a mutant, Minser, which contains a 15-bp insertion between nt 1789 and 1790 that includes a second copy of the Inr sequence 5′-CATA-3′ (Fig. 1B). Analysis of the RNAs synthesized from pMinser in a cell-free transcription system indicated that most of the 5′ ends of the pre-C RNAs became approximately 15 nt longer than those of the wild type (preC′ in Fig. 6A), while those of the pregenomic RNAs remained unchanged (pregenomic in Fig. 6A). A similar result was obtained when Huh7 cells were transiently transfected with pMinser DNA (preC′ in

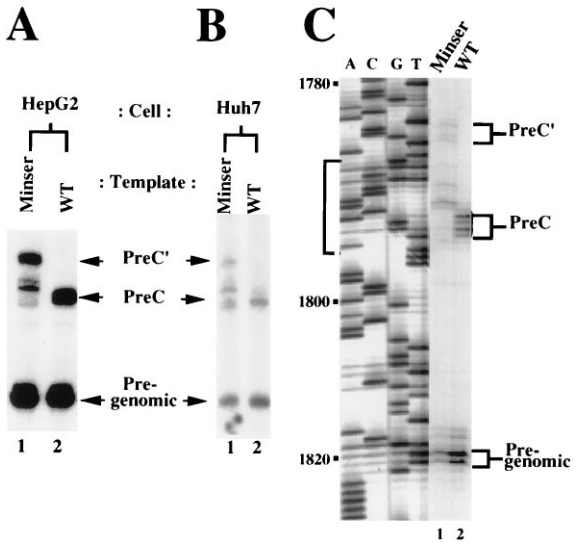

FIG. 6. Physical separation of the pre-C and pregenomic promoters by insertional mutagenesis. (A) Analysis of the 5′ ends of the RNAs synthesized from pMinser (lane 1) and pWT (lane 2) in a cell-free transcription system made from HepG2 nuclear extract. (B) Analysis of the 5′ ends of the mRNAs accumulated in Huh7 cells transfected with pMinser (lane 1) and pWT (lane 2). (C) High-resolution mapping of the 5′ ends of RNAs from panel A. A sequencing ladder generated from pMinser DNA was used as the size markers. The 15-bp insertion is indicated by the bracket on the left.

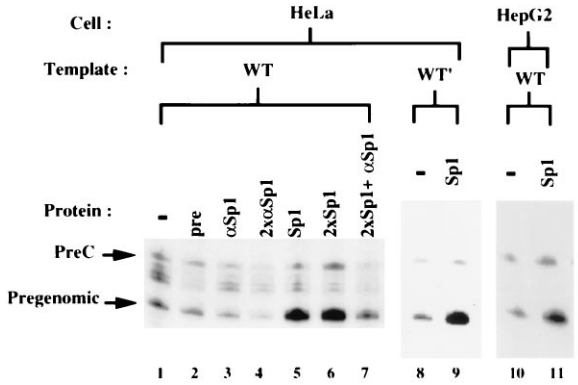

FIG. 7. Recombinant human Sp1 protein specifically transactivates transcription in vitro from the pregenomic RNA promoter. The indicated templates were preincubated with the indicated preimmune serum and/or recombinant protein at 26°C for 15 min prior to the addition of the nuclear extract made from the indicated cell line. Transcription and analysis of the 5′ ends of the viral RNAs were performed as described in Materials and Methods. WT, wild type; αSp1, anti-Sp1 antibody.

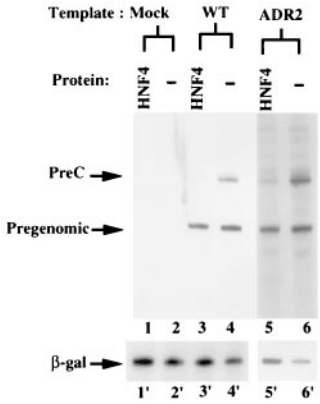

FIG. 8. Protein factor HNF4 specifically represses synthesis of the pre-C RNAs in Huh7 cells. Plasmids pGL-2 (lanes 1 and 2), pWT (lanes 3 and 4), and pADR2 (lanes 5 and 6) were cotransfected with pCDMHNF4 (lanes 1, 3, and 5) or pCDM8 (lanes 2, 4, and 6) into Huh7 cells. Forty-eight hours later, the cells were harvested and processed as described in Materials and Methods. Shown here is the resulting autoradiogram of the 7 M urea–8% polyacrylamide gel in which the cDNAs were electrophoresed. β-gal, β-galactosidase.

Fig. 6B). High-resolution mapping of the 5′ ends of the pre-C′ and pregenomic RNAs synthesized from mutant Minser revealed that the sites of initiation relative to the TATA box-like sequences remained unchanged (Fig. 6C). Primer extension products, smaller and larger than the ones synthesized from the wild-type pre-C RNAs, were also observed (Fig. 6A and B). Likely, they originated from aberrant transcriptional initiations caused by the insertion close to the start sites of the pre-C RNAs. Thus, we conclude that the basal elements of the pre-C and pregenomic promoters can be physically separated by an insertion mutation that includes a duplication of the sequence 5′-CATA-3′, which likely functions as an Inr element for synthesis of the pre-C RNAs.

Sp1 can selectively activate transcription of pregenomic RNAs. Zhang et al. (46) previously reported the existence of three binding sites (nt 1623 to 1632, 1733 to 1742, and 1744 to 1753) for the transcription factor Sp1. Two of these sites are situated immediately 5′ of the 5′-TTAAA-3′ sequence (1760 to 1764) that we have demonstrated to be involved in pre-C RNA synthesis (Fig. 1A). To investigate whether these Sp1 binding sites specifically modulate transcription from either the pre-C or pregenomic promoters, we included recombinant human Sp1 protein in our cell-free transcription reaction mixtures (Fig. 7). The presence of the exogenously added Sp1 increased synthesis of the pregenomic RNAs three- to fivefold, but had little, if any, effect on synthesis of the pre-C RNAs (lane 1 versus lane 5). This activation was due to the presence of Sp1, since no activation occurred when the Sp1 protein was preincubated with a polyclonal antibody to Sp1 (lane 7). The same level of activation was observed with the template pWT′, which lacks the upstream regulatory region (nt 1403 to 1729) (lane 9 versus lane 8). Thus, activation by Sp1 probably occurs via its interaction with the two Sp1 binding sites situated proximal to the basic C promoter. Sp1 did not increase synthesis of the pregenomic RNAs when a double-stranded oligonucleotide containing an Sp1 binding site was included in the reaction; however, activation still occurred when the competitor oligonucleotide contained a G-to-C mutation preventing Sp1 binding (data not shown). Thus, we conclude that transactivation of the pregenomic promoter by Sp1 is sequence specific. Sp1 also selectively activated pregenomic RNA synthesis in the HepG2-made cell-free system (lane 11 versus lane 10). This differential response of the pre-C and pregenomic promoters to activation by Sp1 clearly indicates that the synthesis of the pre-C and pregenomic RNAs occurs from separate promoters and that these promoters can be differentially regulated.

HNF4 selectively represses transcription of the pre-C RNAs in vivo. HNF4 is a member of the steroid-thyroid hormone receptor superfamily which is enriched in liver cells (32). An HNF4 binding site was identified in HBV enhancer I, and transactivation of promoters linked to a multimerized domain of enhancer I containing this HNF4 binding site was reported (14). HNF4 has also been shown in HeLa cells, but not in Huh7 cells, to activate the HBV C promoter via a binding site within enhancer II (16). To examine whether HNF4 might differentially regulate transcription of the pre-C and pregenomic RNAs in hepatoma cells, pWT and pCDMHNF4, a plasmid from which HNF4 can be expressed, were cotransfected into Huh7 cells. The viral, polyadenylated RNAs that accumulated in these cells were analyzed (Fig. 8). Overexpression of HNF4 led to an at least fivefold decrease in pre-C RNA synthesis while leaving pregenomic RNA synthesis unaffected (lane 3 versus lane 4). A similar result was obtained with pADR2 (lane 5 versus lane 6), a plasmid which contains all of the HBV genome (including two additional HNF4 binding sites [14, 16]) except for a segment of the coding region for C (nt 1987 to 2432). Thus, we conclude that the pre-C and pregenomic promoters can be differentially regulated by HNF4 as well.

DISCUSSION

In this report, we describe a cell-free transcription system derived from a hepatoma cell line, HepG2, which directs accurate transcription of the pre-C and pregenomic RNAs of HBV in vitro. Using high-resolution mapping, we showed that the 5′ ends of the pre-C and pregenomic RNAs synthesized in cell-free transcription systems made from HepG2 and HeLa cells are identical (Fig. 2), thus indicating that the selection of the transcription initiation sites of these two RNA species is not liver cell specific. We also showed that deletion of the upstream regulatory region (nt 1403 to 1729), which positively modulates the basic C promoter of HBV in a liver cell-specific manner, also does not affect the selection of the transcription initiation sites of the pre-C and pregenomic RNAs in vivo (Fig.

5). Analysis of base substitution and insertion mutants in the basic C promoter region enabled us to identify several *cis*-acting elements essential for the synthesis of the pre-C and pregenomic RNAs and to establish that synthesis of these two RNA species is directed from two distinct promoters, each consisting of a TATA box-like sequence and an Inr (Fig. 3 and 4). We concluded that these promoters slightly overlap, but are physically separable by a 15-bp insertion that includes a duplication of the Inr element (Fig. 6). We also found that the pre-C and pregenomic promoters can be differentially regulated by *trans*-acting factors: the transcription factor Sp1 can specifically increase synthesis of the pregenomic RNAs three- to fivefold, but has little effect on synthesis of the pre-C RNAs in cell-free transcription systems made from HepG2 and HeLa nuclear extracts (Fig. 7), and HNF4, a member of the steroid-thyroid receptor superfamily, can specifically repress the synthesis of the pre-C RNAs, leaving the pregenomic RNAs unaffected in vivo (Fig. 8).

A cell-free transcription system derived from the hepatoma cell line HepG2. Although it was reported that a HeLa whole-cell extract can direct transcription initiation from the HBV basic C promoter at nt 1792 ± 5 and 1817 ± 5 (37), this cell-free transcription system has not been widely used in HBV studies, possibly because of its nonliver origin. We report here a cell-free transcription system derived from HepG2 nuclear extract which accurately directs the initiation of the HBV pre-C and pregenomic RNAs (Fig. 2). The HBV C promoter mutants we tested have similar phenotypes when analyzed in this cell-free transcription system as when examined by transient transfection into Huh7 and HepG2 cells (Fig. 3 and 4). Therefore, this cell-free transcription system from HepG2 cells may prove to be useful for studying the pre-C and pregenomic promoters, other HBV promoters, and, possibly, even liver-specific cellular promoters. This nuclear extract might also be useful for purifying liver-specific factors and for assessing the effects of drugs on transcription from liver-specific promoters.

The pre-C and pregenomic RNAs of HBV are synthesized from two genetically distinct promoters. A typical RNA polymerase II promoter consists of a TATA box situated approximately 30 bp upstream of the initiation site of transcription and an Inr which overlaps the transcription start site (7, 21). Analysis in vitro and in vivo of the mutants of the HBV basic C promoter region constructed here supports the conclusion that two promoters exist which independently govern transcription of the pre-C and pregenomic RNAs.

For the pre-C promoter, the appropriately positioned TATA box-like sequence, 5′-TTAAA-3′ (nt 1760 to 1764), and the sequence overlapping the transcription initiation sites are both essential for its function as demonstrated by the phenotypes of mutants Mpc1 and Mpg1 (Fig. 3, 4, and 5). The importance of this TATA box-like sequence for the pre-C promoter was also suggested by the recent isolation from clinical patients of e-negative and e-suppressed HBV variants in which this sequence has been mutated (26, 29, 35). High-resolution mapping of the 5′ ends of the pre-C and pregenomic RNAs indicated that the 5′ ends of the pre-C RNAs are much more heterogeneous than are those of the pregenomic RNAs (Fig. 2). It is not clear whether this heterogeneity is due to the presence of a weak TATA box-like sequence or the presence of several weak Inr elements in the pre-C promoter. One stretch of nucleotides, 5′-CATA-3′ (nt 1790 to 1793), which overlaps some of the pre-C RNA initiation sites and resembles the optimal Inr sequence 5′-CA(T/G)T-3′ (21), appears to be part of the pre-C Inr. Mutant Mpg1, in which the sequence 5′-TAA-3′ (nt 1792 to 1794) has been changed to CCC, was found to be defective in synthesis of the pre-C RNAs, while mutant Mpg2, in which base substitutions were introduced in nt 1793 to 1794 and 1796, was not. Therefore, the T at nt 1792, probably as part of the Inr, is necessary for the pre-C promoter function. This is consistent with the report by Chen et al. (4) that base substitutions of AA at nt 1793 to 1794 with CC do not reduce transcription from the pre-C promoter. Mutations further downstream of this T site (Mpg2 and Mpg3) also did not affect pre-C RNA synthesis. Thus, we interpret our data to indicate that the 3′ boundary of the Inr element of the pre-C promoter is at nt 1792. This finding is further supported by the phenotype of mutant Minser in which duplication of the sequence 5′-CATA-3′ was found to be sufficient to physically separate these two promoters, with each promoter being fully competent in supporting viral RNA synthesis (Fig. 6).

The pregenomic promoter also appears to be defined by a TATA box-like sequence and an Inr element. Its approximate −30 region TATA box sequence (5′-CATAAAT-3′) contains two minor mismatches relative to the consensus sequence 5′-TATAAAA-3′. As expected, mutations in Mpg1, Mpg2, and Mpg3 were found to have detrimental effects on pregenomic RNA synthesis because they alter this TATA box sequence. The sequence 5′-CAACT-3′ overlaps the transcription initiation sites of the pregenomic RNAs. Two of its bases match the two most important bases of the optimal Inr sequence (21). Thus, this sequence is likely the Inr element. Attempts were not made to mutate this sequence because the level of transcription of the pregenomic RNAs may not be reduced dramatically in the presence of a strong TATA element, even if its Inr were to be replaced by a sequence which bears little resemblance to the optimal Inr sequence. The data we obtained with mutant Minser (Fig. 6) enabled us to define the 5′ boundary of the pregenomic promoter as being at nt 1790, the first nucleotide of its TATA box sequence.

Our results differ from those reported by Chen et al. (4) with respect to (i) the function of the TATA box-like sequence, 5′-TTAAA-3′, (ii) the heterogeneity in the 5′ ends of the pre-C RNAs, and (iii) the basic structures and the extent of overlap of the pre-C and pregenomic promoters. Most significantly, they found that a stretch of 15 nt (nt 1790 to 1804), which we define as containing the Inr of the pre-C promoter and the TATA element of the pregenomic promoter, was sufficient to direct precise initiation of both pre-C and pregenomic RNAs at nt 1794 ± 1 and 1823 ± 2, respectively. The apparent discrepancy between the findings by Chen et al. and our findings may be explained by differences in the mutations studied and the basic plasmid constructs used in the two sets of studies. For example, their vector sequences adjacent to the 15-bp promoter element may have fortuitously contained appropriately located sequences that could function as the TATA element for the pre-C promoter.

The pre-C and pregenomic promoters can be differentially regulated by *trans*-acting factors. Human Sp1 protein, one of the best-characterized transcription factors, plays a critical role in directing efficient transcription by binding to GC boxes in a variety of mammalian promoters (19). We showed here that recombinant human Sp1 protein can specifically increase synthesis of the pregenomic RNAs but not synthesis of the pre-C RNAs (Fig. 7). The failure of Sp1 to activate the pre-C promoter could be due to the fact that the Sp1 binding sites are situated only a few base pairs 5′ of the TATA box element of the pre-C promoter (46). Thus, the bound Sp1 may interfere with the binding of TBP to this TATA box-like element or the assembly of the transcription initiation complex on the pre-C promoter.

HNF4 was originally identified from a rat liver nuclear extract (32). Human HNF4 isoforms have also been cloned (9).

HNF4 plays an important role in the early embryogenesis of mammals (5, 10). Although HNF4 is found in many organs, binding sites for HNF4 have been found predominantly in genes expressed in liver and kidney. In almost all of the reported studies, HNF4 serves as a transactivator in the expression of liver-specific genes (3, 12, 17, 20, 22, 24). It has also been shown to enhance the expression of a reporter gene linked to the HNF4-binding sites identified in enhancers I and II of HBV (14, 16). We have shown here that HNF4 specifically represses the pre-C promoter in cotransfection experiments (Fig. 8). Moreover, this repression persists in the presence of the HBV enhancers I and II and in the context of the HBV genome (Fig. 8). Recombinant HNF4 purified from HeLa cells (a gift from Todd Gulick and David Moore) can also repress transcription of the pre-C promoter in a cell-free transcription system via an HNF4-binding site overlapping the TATA box-like sequence of the pre-C promoter (42a). The binding of HNF4 to this site could specifically block the formation of preinitiation complexes on the pre-C promoter, thereby repressing transcription initiation from this promoter.

Differential regulation of these two promoters may serve biological functions in the HBV life cycle. The HBV pregenomic RNAs play a central role in the life cycle of the virus by both encoding the C and P proteins needed for nucleocapsid formation and providing RNA templates for viral DNA replication. The importance of the pre-C RNAs is not clear because the biological function of HBV e antigen remains to be determined. However, the conservation of the pre-C ORF during viral evolution strongly suggests that e antigen serves functions advantageous to the virus. Our finding that the pre-C and pregenomic promoters can be differentially regulated lends credence to the notion that the pre-C and pregenomic RNAs probably serve different functions. This differential regulation probably occurs via interactions of *cis*-acting elements in these two promoters with *trans*-acting factors present in various tissues and cell types and during different stages of liver cell differentiation, thereby facilitating the propagation and maintenance of HBV in its host. Farza et al. (11) reported that similar amounts of pre-C and pregenomic RNAs were present in the kidney and heart in transgenic mice, whereas pregenomic RNAs were much more abundant than pre-C RNAs in liver. The latter is also the case in the liver in HBV-infected chimpanzees (11). Taken together with our findings, we speculate that the pre-C promoter is probably partially repressed in HBV-infected liver cells to minimize the host immune response to HBV infection. This selective repression of synthesis of the pre-C RNAs might be especially important during the early stages of HBV infection: by releasing fewer e antigen molecules into the bloodstream, the virus might delay the onset of the host immune response against HBV so that a full-blown infection can be successfully established.

In summary, we conclude that the synthesis of the pre-C and pregenomic RNAs of HBV is under the control of two discrete promoters. These two promoters slightly overlap, yet are genetically separable. Differential regulation of these two promoters by various transcription factors may be crucial for HBV to be maintained and to replicate in its host.

ACKNOWLEDGMENTS

We thank David Moore, Jing-Hsiung Ou, Wyeth Wasserman, and Kim Mansky for providing cell lines and plasmids. We also thank members of our laboratory, Paul Lambert, and Dan Loeb for helpful discussions and comments on the manuscript.

This work was supported by Public Health Science research grants CA22443 and CA07175 from the National Cancer Institute.

REFERENCES

1. **Billet, O., G. Grimber, M. Levrero, K. A. Seye, P. Briand, and V. Joulin.** 1995. In vivo activity of the hepatitis B virus core promoter: tissue specificity and temporal regulation. J. Virol. **69:**5912–5916.
2. **Butz, K., and F. Hoppe-Seyler.** 1993. Transcriptional control of human papillomavirus (HPV) oncogene expression: composition of the HPV type 18 upstream regulatory region. J. Virol. **67:**6476–6486.
3. **Carter, M. E., T. Gulick, B. D. Raisher, T. Caira, J. A. A. Ladias, D. D. Moore, and D. P. Kelly.** 1993. Hepatocyte nuclear factor-4 activates medium chain acyl-CoA dehydrogenase gene transcription by interacting with a complex regulatory element. J. Biol. Chem. **268:**13805–13810.
4. **Chen, I.-H., C.-J. Huang, and L.-P. Ting.** 1995. Overlapping initiator and TATA box functions in the basal core promoter of hepatitis B virus. J. Virol. **69:**3647–3657.
5. **Chen, W. S., K. Manova, D. C. Weinstein, S. A. Duncan, A. S. Plump, V. R. Prezioso, R. F. Bachvarova, and J. E. Darnell, Jr.** 1994. Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. Genes Dev. **8:**2466–2477.
6. **Chomczynski, P.** 1994. Current protocols in molecular biology, p. 4.2.4–4.2.8. John Wiley & Sons, Inc., New York.
7. **Colgan, J., and J. L. Manley.** 1995. Cooperation between core promoter elements influences transcriptional activity in vivo. Proc. Natl. Acad. Sci. USA **92:**1955–1959.
8. **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.
9. **Drewes, T., S. Senkel, B. Holewa, and G. U. Ryffel.** 1996. Human hepatocyte nuclear factor 4 isoforms are encoded by distinct and differentially expressed genes. Mol. Cell. Biol. **16:**925–931.
10. **Duncan, S. A., K. Manova, W. S. Chen, P. Hoodless, D. C. Weinstein, R. F. Bachvarova, and J. E. Darnell, Jr.** 1994. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. Proc. Natl. Acad. Sci. USA **91:**7598–7602.
11. **Farza, H., M. Hadchouel, J. Scotto, P. Tiollais, C. Babinet, and C. Pourcel.** 1988. Replication and gene expression of hepatitis B virus in a transgenic mouse that contains the complete viral genome. J. Virol. **62:**4144–4152.
12. **Galson, D. L., T. Tsuchiya, D. S. Tendler, L. E. Huang, Y. Ren, T. Ogura, and H. F. Bunn.** 1995. The orphan receptor hepatic nuclear factor 4 functions as a transcriptional activator for tissue-specific and hypoxia-specific erythropoietin gene expression and is antagonized by EAR3/COUP-TF1. Mol. Cell. Biol. **15:**2135–2144.
13. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu. Rev. Biochem. **56:**651–693.
14. **Garcia, A. D., P. Ostapchuk, and P. Hearing.** 1993. Functional interaction of nuclear factors EF-C, HNF-4, and RXRα with hepatitis B virus enhancer I. J. Virol. **67:**3940–3950.
15. **Gerlach, K. K., and R. H. Schloemer.** 1992. Hepatitis B virus C gene promoter is under negative regulation. Virology **189:**59–66.
16. **Guo, W., M. Chen, T. S. B. Yen, and J.-H. Ou.** 1993. Hepatocyte-specific expression of the hepatitis B virus core promoter depends on both positive and negative regulation. Mol. Cell. Biol. **13:**443–448.
17. **Hall, R. K., F. M. Sladek, and D. K. Granner.** 1995. The orphan receptors COUP-TF and HNF-4 serve as accessory factors required for induction of phospho*enol*pyruvate carboxykinase gene transcription by glucocorticoids. Proc. Natl. Acad. Sci. USA **92:**412–416.
18. **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.
19. **Kadonaga, J. T., K. A. Jones, and R. Tjian.** 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem. Sci. **11:**20–23.
20. **Kimura, A., A. Nishiyori, T. Murakami, T. Tsukamoto, S. Hata, T. Osumi, R. Okamura, M. Mori, and M. Takiguchi.** 1993. Chicken ovalbumin upstream promoter-transcription factor (COUP-TF) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-4 (HNF-4). J. Biol. Chem. **268:**11125–11133.
21. **Kraus, R. J., E. E. Murray, S. R. Wiley, N. M. Zink, K. Loritz, G. W. Gelembiuk, and J. E. Mertz.** 1996. Experimentally determined weight matrix definitions of the initiator and TBP binding site elements of promoters. Nucleic Acids Res. **24:**1531–1539.
22. **Ladias, J. A. A., M. Hadzopoulou-Cladaras, D. Kardassis, P. Cardot, J. Cheng, V. Zannis, and C. Cladaras.** 1992. Transcriptional regulation of human apolipoprotein genes ApoB, ApoCIII, and ApoAII by members of the steroid hormone receptor superfamily HNF-4, ARP-1, EAR-2, and EAR-3. J. Biol. Chem. **267:**15849–15860.
23. **Lo, W.-Y., and L.-P. Ting.** 1994. Repression of enhancer II activity by a negative regulatory element in the hepatitis B virus genome. J. Virol. **68:**1758–1764.
24. **Mietus-Snyder, M., F. M. Sladek, G. S. Ginsburg, C. F. Kuo, J. A. A. Ladias,**

J. E. Darnell, Jr., and S. K. Karathanasis. 1992. Antagonism between apolipoprotein AI regulatory protein 1, Ear3/COUP-TF, and hepatocyte nuclear factor 4 modulates apolipoprotein CIII gene expression in liver and intestinal cells. Mol. Cell. Biol. **12:**1708–1718.

25. **Nassal, M., M. Junker-Niepmann, and H. Schaller.** 1990. Translational inactivation of RNA function: discrimination against a subset of genomic transcripts during HBV nucleocapsid assembly. Cell **63:**1357–1363.
26. **Okamoto, H., F. Tsuda, Y. Akahane, Y. Sugai, M. Yoshiba, K. Moriyama, T. Tanaka, Y. Miyakawa, and M. Mayumi.** 1994. Hepatitis B virus with mutations in the core promoter for an e antigen-negative phenotype in carriers with antibody to e antigen. J. Virol. **68:**8102–8110.
27. **Roossinck, M. J., S. Jameel, S. H. Loukin, and A. Siddiqui.** 1986. Expression of hepatitis B viral core region in mammalian cells. Mol. Cell. Biol. **6:**1393–1400.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. p. 16.39. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
29. **Sato, S., K. Suzuki, Y. Akahane, K. Akamatsu, K. Akiyama, K. Yunomura, F. Tsuda, T. Tanaka, H. Okamoto, Y. Miyakawa, and M. Mayumi.** 1995. Hepatitis B virus strains with mutations in the core promoter in patients with fulminant hepatitis. Ann. Intern. Med. **122:**241–248.
30. **Schleiss, M. R., C. R. Degnin, and A. P. Geballe.** 1991. Translational control of human cytomegalovirus gp48 expression. J. Virol. **65:**6782–6789.
31. **Sells, M. A., A. Z. Zelent, M. Shvartsman, and G. Acs.** 1988. Replicative intermediates of hepatitis B virus in HepG2 cells that produce infectious virions. J. Virol. **62:**2836–2844.
32. **Sladek, F. M., W. Zhong, E. Lai, and J. E. Darnell, Jr.** 1990. Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. Genes Dev. **4:**2353–2365.
33. **Standring, D. N.** 1991. The molecular biology of the hepatitis B virus core protein, p. 145–169. *In* A. McLachlan (ed.), Molecular biology of the hepatitis B virus. CRC Press, Boca Raton, Fla.
34. **Standring, D. N., J.-H. Ou, F. R. Masiarz, and W. J. Rutter.** 1988. A signal peptide encoded within the precore region of hepatitis B virus directs the secretion of a heterogeneous population of e antigens in *Xenopus* oocytes. Proc. Natl. Acad. Sci. USA **85:**8405–8409.
35. **Takahashi, K., K. Aoyama, N. Ohno, K. Iwata, Y. Akahane, K. Baba, H. Yoshizawa, and S. Mishiro.** 1995. The precore/core promoter mutant ($T^{1762}A^{1764}$) of hepatitis B virus: clinical significance and an easy method for detection. J. Gen. Virol. **76:**3159–3164.
36. **Tiollais, P., C. Pourcel, and A. Dejean.** 1985. The hepatitis B virus. Nature (London) **317:**489–495.
37. **Waisman, A., Y. Aloni, and O. Laub.** 1990. In vitro regulation of human hepatitis B virus core gene transcription. Virology **177:**737–744.
38. **Wiley, S. R., R. J. Kraus, and J. E. Mertz.** 1992. Functional binding of the "TATA" box binding component of transcription factor TFIID to the −30 region of TATA-less promoters. Proc. Natl. Acad. Sci. USA **89:**5814–5818.
39. **Wiley, S. R., R. J. Kraus, F. Zuo, E. E. Murray, K. Loritz, and J. E. Mertz.** 1993. SV40 early-to-late switch involves titration of cellular transcriptional repressors. Genes Dev. **7:**2206–2219.
40. **Will, H., W. Reiser, T. Weimer, E. Pfaff, M. Büscher, R. Sprengel, R. Cattaneo, and H. Schaller.** 1987. Replication strategy of human hepatitis B virus. J. Virol. **61:**904–911.
41. **Wu, S. F., Y. Z. Zhou, Z. M. Feng, Z. P. Li, and S. Y. Sha.** 1983. Molecular cloning of human hepatitis B (adr) from a blood sample of a Chinese carrier. Sci. Sin. Ser. B **26:**954–960.
42. **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.

42a. **Yu, X.** Unpublished data.

43. **Yuh, C.-H., Y.-L. Chang, and L.-P. Ting.** 1992. Transcriptional regulation of precore and pregenomic RNAs of hepatitis B virus. J. Virol. **66:**4073–4084.
44. **Zhang, P., and A. McLachlan.** 1994. Differentiation-specific transcriptional regulation of the hepatitis B virus nucleocapsid gene in human hepatoma cell lines. Virology **202:**430–440.
45. **Zhang, P., A. K. Raney, and A. McLachlan.** 1992. Characterization of the hepatitis B virus X- and nucleocapsid gene transcriptional regulatory elements. Virology **191:**31–41.
46. **Zhang, P., A. K. Raney, and A. McLachlan.** 1993. Characterization of functional Sp1 transcription factor binding sites in the hepatitis B virus nucleocapsid promoter. J. Virol. **67:**1472–1481.