JOURNAL OF VIROLOGY, Sept. 1989, p. 4011–4018
0022-538X/89/094011-08$02.00/0

Vol. 63, No. 9

Hepatitis B Virus Transcript Produced by RNA Splicing

TSUNG-SHENG SU,[1,2,3]* CHENG-JUNG LAI,[2] JAU-LING HUANG,[2] LING-HUANG LIN,[1] YAR-KHING YAUK,[1] CHUNGMING CHANG,[1,2] SZECHENG J. LO,[2,3] AND SHOU-HWA HAN[1,2]

Department of Medical Research, Veterans General Hospital, Taipei, Taiwan 11217,[1] and Graduate Institutes of Microbiology and Immunology[2] and Genetics,[3] National Yang-Ming Medical College, Taipei, Taiwan 11221, Republic of China

Received 6 March 1989/Accepted 6 June 1989

A new hepatitis B virus (HBV) transcript of about 2.2 kilobases was identified in HBV DNA-transfected human hepatoma cells. The 5′ terminus of this viral RNA appears to map at one or more of the precore initiation sites, contains a deletion of 1,223 bases corresponding to the last codon of the core gene to the middle of the surface antigen gene, and terminates at the 3′ polyadenylation site used by the other known HBV RNAs. The junction region of the deleted sequences showed the conserved splice donor and acceptor GT-AG sequences. Moreover, when a mutant HBV DNA in which the splice acceptor site was changed from AG to CG was transfected into human hepatoma cells, no 2.2-kilobase RNA was detected, further suggesting that this RNA represents a spliced transcript. The core gene, although an amino acid shorter, still encoded a functional viral core protein in complementation experiments. Sequence analysis of the cDNA of the 2.2-kilobase RNA suggests that this transcript can potentially encode a new protein that comprises the reverse transcriptase domain of HBV. However, genetic analysis using a transient DNA transfection system suggests that the gene product(s) of this transcript is not essential for viral replication. The function of this transcript remains to be studied.

Hepatitis B virus (HBV) infection is a public health problem of worldwide importance (9). Nucleotide sequencing studies of HBV genomes of different subtypes have shown four common open reading frames that code for presurface, surface, and core antigens, the putative polymerase, and an X protein (10, 31). Two major classes of viral RNAs have been reported: (i) the 2.4-kilobase (kb) and 2.1-kb subgenomic RNAs for expression of the presurface and surface proteins and (ii) the 3.5-kb pregenomic RNAs which serve as templates for both viral DNA replication and protein synthesis (10, 31). In addition, recent studies have suggested the existence of a minor transcript for HBV X protein (33). All of these RNAs are unspliced and terminate at a common site (10, 31). Other than these transcripts, some novel HBV RNAs, either spliced or unspliced, have also been detected in several heterologous systems (24, 27). More recently, an attempt to define additional HBV products that might play regulatory roles similar to those of proteins such as tat and rev in the human immunodeficiency virus system (8) has led to identification of a new viral RNA in HBV-infected human liver specimens (29). This 2.2-kb RNA species hybridized to an HBV core sequence and was polyadenylated. Since the distance between the HBV polyadenylation site and the core gene is greater than 2.2 kb (10, 31), the 2.2-kb RNA is probably a transcript that has undergone processing (29). More interestingly, this RNA coexisted with the 3.5-kb pregenomic RNA in two human livers analyzed in which the virions were actively replicating (29). Similarly, an RNA species of the same size has also been detected in an in vitro HBV transient expression system which supports viral DNA replication and produces Dane particlelike structures (4).

To elucidate the structure and function of the 2.2-kb transcript, a cDNA clone was derived from RNA prepared from the in vitro HBV transient expression system. The structure of this RNA was studied and confirmed to be that of a spliced product.

* Corresponding author.

MATERIALS AND METHODS

DNA and cell line. Plasmid pSHH2-1 contains a tandem repeat of a 3,182-base-pair (bp) *Eco*RI fragment which encodes the complete HBV genome (subtype ayw) inserted into vector pSV08 (3, 18). In this study, the unique *Eco*RI site in the HBV genome was defined as nucleotide 1. Plasmid pMH9/3091 contains an overlength HBV genome which encodes the viral pregenomic RNA which is under control of the human metallothionein promoter (15). The HuH-7 cell line (22), established from a primary hepatocellular carcinoma, was used for DNA transfection. Cells were maintained in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum.

DNA transfection. Plasmid DNA, 20 to 40 μg, was used to transfect 10^7 hepatoma cells per 150-mm-diameter culture dish by the calcium phosphate method as previously described (28). The culture medium was removed and replaced with fresh medium after incubation at 37°C for 24 h and at 2-day intervals.

Isolation, electrophoresis, and detection of HBV RNA. The total RNA of HBV DNA-transfected cells was prepared by the guanidinium-cesium chloride method (11) 7 days after DNA transfection.

RNA was purified by fractionation with oligo(dT)-cellulose chromatography as described by Aviv and Leder (1). Poly(A)$^+$ RNA from HBV DNA-transfected HuH-7 cells was denatured with glyoxal and applied to a 1.2% agarose gel for electrophoresis (30, 34). The RNA was transferred to nitrocellulose paper and hybridized with nick-translated ^{32}P-labeled HBV DNA probes under previously described conditions (34). The HBV subgenomic fragments used in this study were subcloned into plasmid pSP65 and reisolated before use as a probe. The blots were washed in 15 mM NaCl–1.5 mM sodium citrate plus 0.1% sodium dodecyl

sulfate at 50 to 55°C and autoradiographed with Fuji X-ray film at −70°C with a Kyokko intensifying screen.

Cloning and analysis of cDNA derived from the 2.2-kb RNA. Poly(A)$^+$ RNA isolated from the pSHH2-1-transfected HuH-7 cell line was used for cDNA library construction with vector λgt11 as described by Huynh et al. (13). The cDNA library was screened by plaque hybridization using ^{32}P-labeled nick-translated HBV DNA as a probe (13).

The cDNA insert excised from vector λgt11 by *Eco*RI was subcloned into pUC9 to obtain pCH7-1. Plasmids pCH7-7, pCH7-8, and pCH7-10 were derived from pCH7-1 as follows. To construct pCH7-7, pCH7-1 was partially digested with *Bgl*II, followed by *Eco*RI, and the 1,412-bp HBV fragment (nucleotides 2423 to 1926) was cloned into *Eco*RI- and *Bam*HI-digested vector pGEM1. Plasmid pCH7-8 is an extension of the 5′ end of pCH7-1 to include the HBV sequence to the *Bam*HI site at nucleotide 1400. This plasmid contains the core promoter region. Plasmid pCH7-10 is an extension of the 5′ end of pCH7-1 to the *Fsp*I site at nucleotide 1800 to include the initiation codon of the precore gene (see Fig. 8). Both pCH7-8 and pCH7-10 contain 5′ sequences from pSHH2-1, and both were subcloned into the pGEM1 vector. Plasmid pCH7-12 was constructed by taking the *Bam*HI-*Xba*I fragment (nucleotides 488 to 1990) from pSHH2-1 and subcloning it into the *Bam*HI-*Xba*I1 site of pGEM1. The M13 *Bam*HI-*Bst*EII clone was constructed by *Bst*EII digestion of pSHH2-1. After Klenow filling in of the ends, the DNA was cleaved by *Bam*HI. The *Bam*HI-*Bst*EII fragment (nucleotides 1400 to 2660) was cloned into the *Bam*HI-*Hin*cII sites of M13 mp18. The HBV sequence cloned in the plasmid vector or in the single-stranded M13 phage was sequenced by the dideoxy-chain termination method (12, 25).

S1 nuclease mapping. Appropriate restriction fragments prepared from pSHH2-1 were labeled either at the 5′ end with T4 polynucleotide kinase and [γ-^{32}P]ATP or at the 3′ end with the DNA polymerase I Klenow fragment and [α-^{32}P]dATP (19). A uniformly labeled single-stranded probe was prepared by using M13 *Bam*HI-*Bst*EII single-stranded DNA as the template, and the ^{32}P-labeled DNA was synthesized by Klenow extension. The DNA obtained was digested with *Eco*RI, and the single-stranded labeled DNA was separated from the template DNA by alkaline agarose gel electrophoresis (23). S1 nuclease analysis was performed as previously described (6). The protected products were displayed in 8 M urea–polyacrylamide gels.

In vitro transcription and translation. Plasmids, pCH7-7, pCH7-8, pCH7-10, and pCH7-12 were linearized with the appropriate restriction enzymes. RNA transcripts were synthesized in vitro with either SP6 RNA polymerase or T7 RNA polymerase (Promega Biotec), depending on the orientation of the DNA insert. The conditions used were as specified by the supplier and included ^{7}mGpppG to ensure that the synthetic RNAs were capped. The synthesized RNA was used to program mRNA-dependent rabbit reticulocyte translation (Promega Biotec) in the presence of [^{35}S]methionine. The products were subjected to sodium dodecyl sulfate–10% polyacrylamide gel analysis (17).

Oligonucleotide-directed mutagenesis. Site-directed mutagenesis was performed by the Kunkel method of using a DNA template containing a small number of uracil residues in place of thymine (16). To mutate the proposed HBV splice acceptor site, a 1,234-bp *Eco*RI-*Sph*I fragment from HBV nucleotides 1 to 1234 was isolated from pSHH2-1 and cloned into M13 mp19. The mutagenic oligonucleotide covered the proposed splice site from nucleotide 478 to nucleotide 496 with a single mismatch sequence (underlined): TCTAATTC CCGGATCCTCA. The recombinant phage bearing the mutation was identified by DNA sequence analysis. The 1,234-bp *Eco*RI-*Sph*I fragment from the recombinant phage was sequenced and replaced the wild-type *Eco*RI-*Sph*I fragment of pMH9/3091. The resulting plasmid was designated pMH9/3091-ml. To introduce an early termination in the core protein, a *Sal*I-*Eco*RI fragment which covers the core region was isolated from pMH9/3091 and cloned into M13 mp18. The mutagenic oligonucleotide covered the core region from nucleotide 1927 to nucleotide 1949, with the mismatch sequence (underlined) replacing the original GTG sequence: TGGAGCTACTCGAGTTACTCT. The resulting mutation not only introduces a nonsense codon at amino acid 60 in the core reading frame but also creates a new *Xho*I cutting site. The recombinant phage bearing the mutation was identified by its *Xho*I restriction pattern. The *Sal*I-*Eco*RI fragment from the recombinant phage was sequenced and replaced the wild-type *Sal*I-*Eco*RI fragment of pMH9/3091. The resulting plasmid was designated pMH9/3091-m8.

Endogenous DNA polymerase assay. Culture fluid from HBV DNA-transfected HuH-7 cells at days 3, 5, and 7 posttransfection was pooled and clarified by centrifugation at 1,500 × *g* for 15 min. The supernatant was then centrifuged at 150,000 × *g* for 2 h. The resulting pellet was dissolved in polymerase buffer with one of the labeled deoxynucleotide triphosphates, and the endogenous DNA polymerase assay was performed as previously described (4). The repaired viral DNA was then digested with restriction endonucleases and subjected to agarose gel electrophoresis. The gel was subsequently dried and autoradiographed.

RESULTS

Detection of RNAs hybridizable to a core-specific probe in an HBV transient expression system. To examine whether the 2.2-kb RNA identified in HBV-infected human livers is also expressed in the HBV transient expression system, DNA of plasmid pSHH2-1 containing a tandem dimer of the HBV genome was transfected into human hepatoma cell line HuH-7. Poly(A)$^+$ RNA was purified 7 days after DNA transfection. Figure 1 shows the results of Northern (RNA) blot analysis of HBV RNA using either the entire HBV genome or HBV subgenomic fragments as probes. Our results showed that besides the 3.5-kb pregenomic RNA, a major population of RNA species of about 2.1 to 2.4 kb was also detected with the full-length HBV probe (Fig. 1B, lane 1). These RNAs encode mainly the presurface and surface proteins. In addition, a 2.2-kb RNA was detected with an HBV *Bgl*II fragment containing the core sequence mapping between nucleotides 1984 and 2423 (Fig. 1B, lane 2). Thus, the 2.1- to 2.4-kb RNA population is heterogeneous and includes the novel 2.2-kb species that were detected in HBV-infected human livers in our earlier study (29). By densitometer scanning, the ratio of the 2.2-kb RNA to the 3.5-kb RNA is about 1 to 4. The 2.2-kb RNA did not hybridize with a downstream *Bgl*II fragment probe from nucleotides 2424 to 2837 (Fig. 1B, lane 3), suggesting that a processing mechanism is involved in its production.

Cloning and sequence analysis of cDNA derived from the 2.2-kb RNA. To study the structure of the novel 2.2-kb RNA transcript, a cDNA corresponding to this RNA was isolated from a λgt11 cDNA library constructed from HBV DNA (pSHH2-1)-transfected HuH-7 cells. With the HBV *Bgl*II fragment of nucleotides 1984 to 2423 as a probe, 2 × 10^5

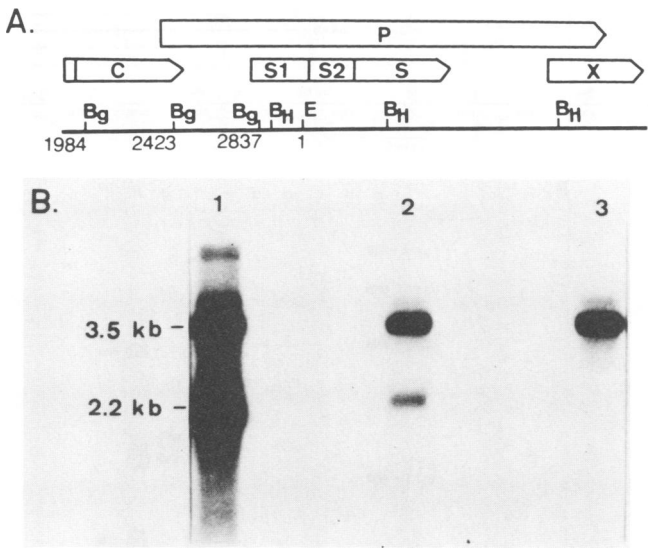

FIG. 1. (A) HBV map showing positions of the probe fragments used. The open arrows represent the four major open reading frames of HBV. Bg, *Bgl*II; B_H, *Bam*HI; E, *Eco*RI. (B) Northern blot analysis of RNA from pSHH2-1-transfected HuH-7 cells. The RNA filter was hybridized with an HBV probe prepared from the entire HBV insert of pSHH2-1 (lane 1); the *Bgl*II fragment, nucleotides 1984 to 2423 (lane 2); and the *Bgl*II fragment, nucleotides 2424 to 2837 (lane 3). Three micrograms of poly(A)$^+$ RNA was applied to each lane.

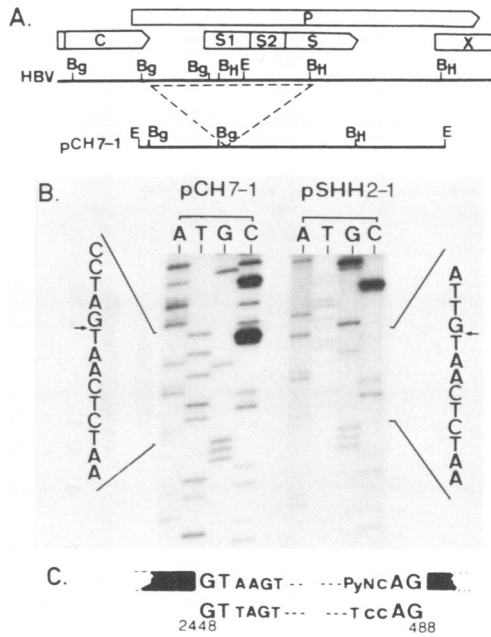

FIG. 2. (A) Restriction map of pCH7-1. The open arrows represent the four major open reading frames of wild-type HBV. The broken line beneath the top line indicates the deleted area in the cDNA. Bg, *Bgl*II; B_H, *Bam*HI; E, *Eco*RI. (B) DNA sequencing analysis of pCH7-1 and pSHH2-1 showing the point of deletion in pCH7-1. The arrows indicate the proposed deletion sites. (C) Comparison of the consensus sequence of the mRNA splice site (top) and the HBV sequences at the deletion junction in pCH7-1 (bottom). The black bars flanking the sequences represent exons.

plaques were screened and three positive clones were identified. These clones did not hybridize to the *Bgl*II fragment of nucleotides 2424 to 2837. However, by DNA sequence analysis, all three clones were shown to possess identical sequences at both the 5′ and the 3′ ends (data not shown). Since the library had been amplified before screening, these three clones were probably derived from a single clone. The cDNA insert was isolated by *Eco*RI digestion and cloned into the *Eco*RI site of plasmid pUC9. The resulting plasmid was designated as pCH7-1. Restriction mapping and sequence analysis of pCH7-1 revealed that the cDNA lacks a piece of the HBV genome from the 3′ end of the *C* gene (nucleotide position 2448) to the middle of the *S* gene (nucleotide position 488), thus deleting a 1,223-bp fragment (Fig. 2A and B). The junction of the deleted region contains the RNA splice donor and acceptor sites, that is, the GT-AG sequences (21; Fig. 2C). This result further supports the suggestion that this RNA was produced by a splicing event. The 5′ end of the cDNA was also determined to be at nucleotide 1904, and its 3′ end was at nucleotide 1926 (data not shown), which covers the HBV polyadenylation signal utilized by the other known HBV RNAs (10, 31). The entire length of this cDNA is 1,982 bp.

Mapping of the splice junction in the 2.2-kb transcript by S1 nuclease analysis. To rule out the possibility that the deletion detected in the cDNA was due to artifacts created during the cloning procedure, S1 nuclease mapping analysis of the 2.2-kb RNA was performed. A 577-bp *Acc*I-*Xba*I fragment labeled at the *Acc*I site was used as a probe to map the position of the acceptor site (Fig. 3). If the junction of the disjoint fragment was a true acceptor site, a fragment 339 nucleotides long should be protected by the 2.2-kb RNA after S1 nuclease digestion. A fragment about this long was detected (Fig. 3, lane c). In addition, a 577-base fragment was also protected. This fragment resulted mainly from protection by other HBV RNA species. Thus, S1 nuclease analysis confirmed the suggestion that the 2.2-kb RNA did arise from splicing and the acceptor site is located at the point deduced from the cDNA sequence. Similarly, the 5′ donor site was mapped by an 850-bp *Hpa*II-*Eco*RI fragment labeled at the *Hpa*II site. A 117-base fragment was protected, as expected (Fig. 3, lane f). In this case, a smearing signal was also observed. Since it also appeared in a reaction with RNA from nontransfected HuH-7 cells (Fig. 3, lane e), the smear would not be related to the HBV sequences. This study demonstrated that the structure of the cDNA clone reflects the structure of RNA and is not an artifact of the cloning procedure.

Mapping of the 5′ ends of the 2.2-kb transcript by S1 nuclease analysis. DNA sequence analysis of the cDNA clone showed that its 5′ end is at nucleotide 1904 while the 3′ end is at nucleotide 1926. The termini of this cDNA are close to the termini of the 3.5-kb transcript, suggesting that the 3.5-kb RNA is a potential candidate as the precursor of the 2.2-kb transcript. Since the 5′ end of the 3.5-kb RNAs is heterogeneous (10), the RNA start sites of the 2.2-kb transcript were determined by S1 nuclease mapping analysis. A uniformly labeled single-strand probe covering the HBV *Bam*HI-*Bst*EII region (nucleotides 1400 to 2660) was hybridized to RNAs from transfected and nontransfected HuH-7 cells (Fig. 4). Three transcriptional initiation sites have been mapped for the 3.5-kb pregenomic RNA at +2, −26, and −32 with respect to the precore AUG (35). Thus, the expected sizes of the protected species from the pregenomic RNA would be 847, 875, and 881 bases, respectively, if pregenomic RNA in this study did start from these sites. With RNA from pSHH2.1 DNA-transfected cells, S1 nucle-

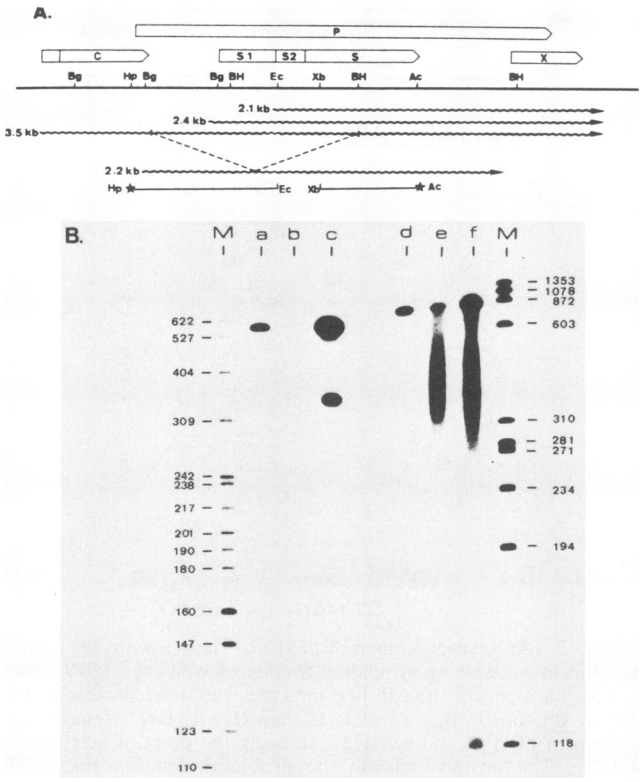

FIG. 3. S1 nuclease mapping of the proposed splice sites. (A) Schematic representation of the HBV genome (subtype ayw). The open arrows represent the four major open reading frames of wild-type HBV. The structures of the HBV transcripts are indicated by wavy lines. The probes used for S1 analysis are depicted, and the ^{32}P-labeled ends are indicated by stars. (B) Polyacrylamide gel electrophoresis of the protected products. The 577-bp *Xba*I-*Acc*I probe was used in lanes a, b, and c, while the 850-bp *Hpa*II-*Eco*RI probe was used in lanes d, e, and f. Lanes: a and d, probe only; b and e, S1 nuclease analysis using RNA from nontransfected HuH-7 cells; c and f, S1 nuclease analysis using RNA from HBV DNA-transfected HuH-7 cells. The products were electrophoresed through a 6% polyacrylamide gel containing 8 M urea. The markers (leftmost lane M) whose sizes are shown on the left (in bases) were end-labeled *Msp*I fragments of pBR322 DNA, while the markers on the right were end-labeled ϕX174 *Hae*III fragments. Ac, *Acc*I; Bg, *Bgl*II; B_H, *Bam*HI; Ec, *Eco*RI; Hp, *Hpa*II; Xb, *Xba*I.

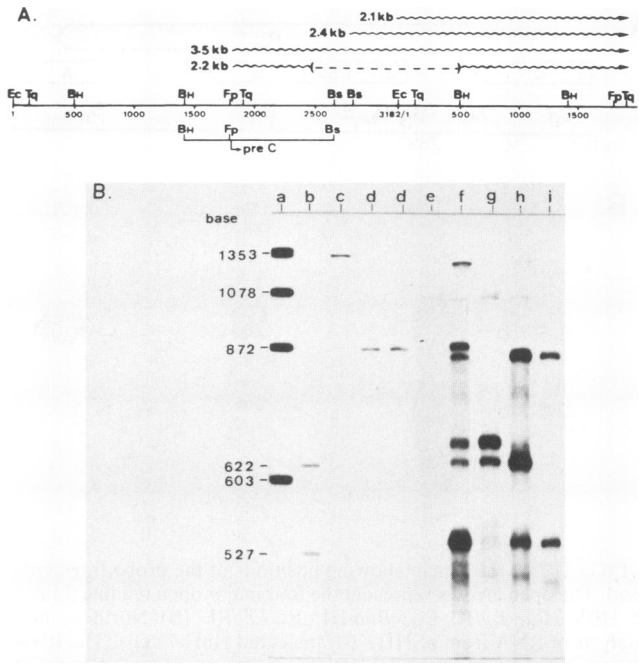

FIG. 4. S1 nuclease mapping of the 5′ end of the 2.2-kb transcript. (A) Schematic representation of HBV transcriptional map. The restriction maps of wild-type HBV (top line) and labeled HBV DNAs used for S1 nuclease analysis (bottom line) are also presented. The precore (pre C) initiation codon is marked by an arrow. HBV RNAs are indicated by wavy lines. The dashed line indicates the spliced region. (B) Polyacrylamide gel electrophoresis of protected products. Lanes: a, end-labeled ϕX174 *Hae*III fragments; b, end-labeled *Msp*I fragments of pBR322; c, uniformly labeled DNA probe used in S1 nuclease analysis; d, end-labeled *Fsp*I-*Bst*EII fragment; e, S1 nuclease analysis using RNA from nontransfected HuH-7 cells; f to i, results of S1 nuclease analysis using RNA from HuH-7 cells transfected with pSHH2-1 (lane f), pCH7-8 (lane g), pMH9/3091 (lane h), and pMH9/3091-ml (lane i). The *Taq*I fragment blocker DNA (nucleotides 128 to 1908) was included to prevent RNA snapback during hybridization. The protected products were analyzed on a 4% polyacrylamide gel containing 8 M urea. B_H, *Bam*HI; Bs, *Bst*EII; Ec, *Eco*RI; Fp, *Fsp*I; Tq, *Taq*I.

ase analysis showed that protected species of about 876 and 847 bases were detected, suggesting that these fragments were derived by protection by the 3.5-kb RNA (Fig. 4, lane f). Furthermore, when transfecting a plasmid DNA (pMH9/3091) which encodes an HBV pregenomic RNA starting from nucleotide −4 with respect to the precore AUG (15), no 876-base fragment was protected (Fig. 4, lane h). Therefore, the 876-base fragment corresponds to the fragment protected by the long species of 3.5-kb RNA, while the 847-base fragment was derived by protection by the short RNA species. To map the 2.2-kb RNA start site, a control plasmid, pCH7-8, was constructed. This plasmid is a derivative of the cDNA clone, pCH7-1, in which the 5′ end of the HBV sequence was extended to nucleotide 1400 to cover the core promoter region. In pCH7-8 DNA-transfected HuH-7 cells, only the 2.2-kb RNA, not the 3.5-kb species, is synthesized. Indeed, RNA from this transfection protected only 657- and 628-base fragments, not the 876- and 847-base fragments (Fig. 4, lane g). The sizes of the protected fragments were as predicted if the same RNA initiation sites for pregenomic RNA were also used in the 2.2-kb HBV RNA in pCH7-8 DNA-transfected cells. RNA from pSHH2-1-transfected HuH-7 cells also protected the same 657- and 628-base fragments, suggesting that the 2.2-kb splicing transcript has the same start sites as the pregenomic RNA and the HBV RNA in pCH7-8 DNA-transfected cells (Fig. 4, lane f). Furthermore, if a plasmid DNA (pMH9/3091) that encodes only the shortest species of 3.5-kb RNA was used in the transfection, RNA from these cells protected only the 628-base fragment (Fig. 4, lane h). When RNA from pMH9/3091-ml, which did not produce the 2.2-kb species on transfection (see Fig. 8), was used, both the 657- and 628-base fragments were absent (Fig. 4, lane i). Clearly, all of the data support the notion that the 5′ end of the spliced transcript is heterogeneous and both the 2.2-kb and pregenomic RNAs have the same initiation site. In this study, a 535-base fragment that appeared in most of the S1 nuclease experiments was a consequence of protection by the 3′ sequences of HBV RNA. Since all known HBV RNA species are expected to terminate at the same site (10, 31), the weak signal of this fragment was due to the fact a *Taq*I fragment (nucleotides 128 to 1908) was included as a blocker in the

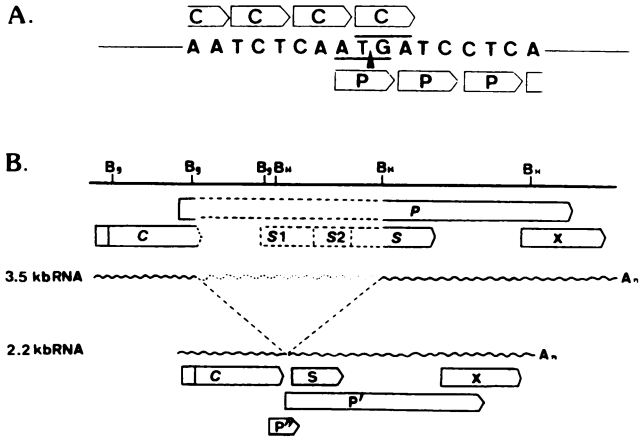

FIG. 5. (A) Assignment of HBV C and P open reading frames at the splice junction (arrowhead). (B) Open reading frame assignments for the 2.2-kb RNA. The broken wavy line indicates the spliced region. C, S, X, P', and P'' represent potential open reading frames in the 2.2-kb transcript. Bg, *Bgl*II; B_H, *Bam*HI.

hybridization reaction to prevent RNA snapback during S1 nuclease analysis (6). Consequently, most of the 3' HBV RNA sequences hybridized to this blocker RNA.

Potential open reading frames in the 2.2-kb transcript. In the sequence of the 2.2-kb RNA-derived cDNA, a termination codon, TGA, is created in the core reading frame right before the authentic core termination codon because of splicing (Fig. 5A). Therefore, the core protein translated from the spliced RNA would be one amino acid (cysteine) shorter. On the other hand, the splicing also created a new initiation codon in the polymerase reading frame which would potentially result in a 42-kilodalton (kDa) protein. Figure 5B summarizes the open reading frames identified in this transcript. Besides the core frame (C) that can encode the precore and core proteins just described, a P' frame that maintains the original carboxyl terminus of the original HBV polymerase reading frame was also observed. The P'' frame that initiates at the putative polymerase AUG encodes 46 amino acids in the N terminus of the polymerase open reading frame and 14 amino acids in a new frame at the carboxyl terminus. The surface frame (S) includes the C-terminal 102 amino acids of the surface gene. The entire X open reading frame is conserved in this transcript.

To study the product of this transcript, an in vitro transcription-translation system was used. Three plasmids, pCH7-7, pCH7-8, and pCH7-10, which were derived from the cDNA and encode different lengths of RNA (Fig. 6A), were transcribed in vitro. The resulting RNAs were translated in the rabbit reticulocyte translation system in the presence of the [^{35}S]methionine. Figure 6B shows the autoradiogram of the in vitro-translated products. The RNA from pCH7-7, which includes the P' frame, was found to be translated into two polypeptides of about 39 and 38 kDa (Fig. 6B, lane 1). To determine which of these proteins corresponds to the protein in the P' frame, the AUG codon created by the splicing event was deleted in plasmid pCH7-12 and the translation product was analyzed. In the absence of this AUG codon, only a 38-kDa protein was detected (Fig. 6C). Therefore, the 39-kDa band was probably the protein derived from the P' frame, while the 38-kDa protein was initiated from an internal AUG codon. Plasmids pCH7-8 and pCH7-10, besides having the P' frame, also include the core frame. The transcription-translation products from these two

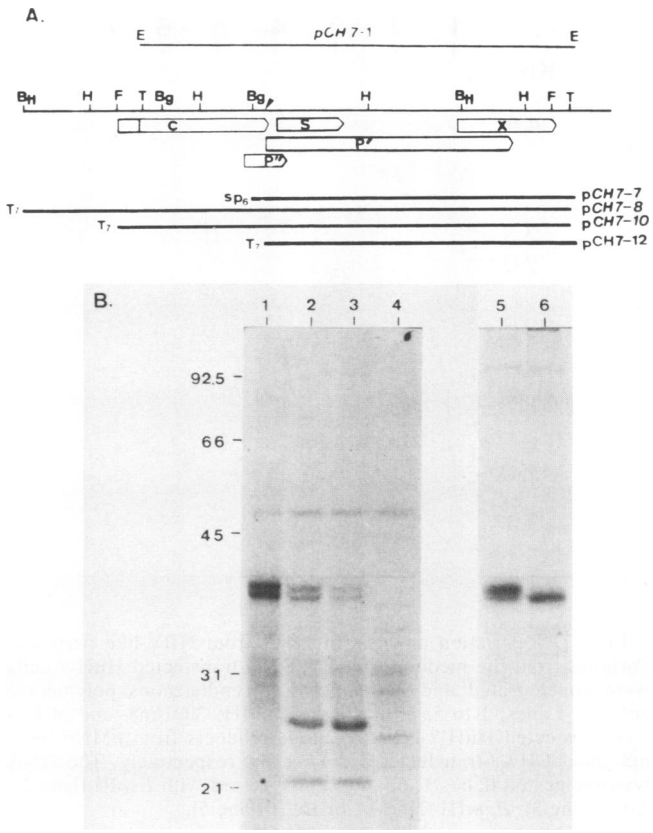

FIG. 6. (A) Structures of expression plasmids pCH7-7, pCH7-8, pCH7-10, and pCH7-12. The arrowhead indicates the splice site. B_H, *Bam*HI; Bg, *Bgl*II; F, *Fsp*I; H, *Hinc*II; T, *Taq*I. (B) Autoradiograph of in vitro-translated products. RNAs transcribed in vitro from plasmids pCH7-7 (lanes 1 and 5), pCH7-8 (lane 2), pCH7-10 (lane 3), and pCH7-12 (lane 6) were translated in a rabbit reticulocyte system supplemented with [^{35}S]methionine. The translation products were run on a sodium dodecyl sulfate–10% polyacrylamide gel. Lane 4 represents the translation reaction in the absence of exogenous RNA. The sizes of the relative molecular mass standards are shown in kilodaltons.

plasmids revealed two more polypeptides of 26 and 21 kDa. The sizes of the product and its RNA template suggest that the 39-kDa protein is the product of the P' frame, while the 26- and the 21-kDa proteins correspond to the precore and core proteins, respectively.

Functional analysis of the 2.2-kb transcript. Since the spliced transcript can potentially encode a core protein which lacks an amino acid, it would be important to determine whether this protein can still serve as a viral core. A complementation test was performed by cotransfecting DNA of pMH9/3091-m8, which is defective in core protein synthesis, and DNA of pCH7-8, which is able to encode a core protein which is an amino acid shorter. The result of complementation was analyzed by examination of the production of HBV-like particles as assayed by their endogenous polymerase activity. When either pMH9/3091-m8 or pCH7-8 was transfected alone, no HBV-like particle was detected (Fig. 7, lanes 6 and 7). However, when these two DNAs were cotransfected into HuH-7 cells, HBV-like particles were produced. Since the mutation in the core gene of pMH9/3091-m8 introduced a new *Xho*I site, two *Xho*I fragments about 1.8 and 1.4 kb long were expected and observed (Fig. 7, lane 3). Thus, the complementation achieved was not

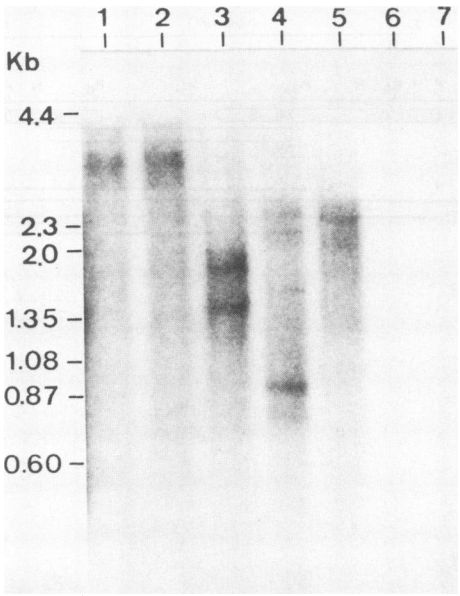

FIG. 7. Restriction analysis of DNA from HBV-like particles. Particles from the medium of HBV DNA-transfected HuH-7 cells were concentrated and radiolabeled by endogenous polymerase activity. Lanes: 1 to 5, products from pMH9/3091-m8- and pCH7-8-cotransfected HuH-7 cells; 6 and 7, products from pMH9/3091-m8- and pCH7-8-transfected HuH-7 cells, respectively. The DNA was undigested (lanes 1, 6, and 7) or digested with *Eco*RI (lane 2), *Xho*I (lane 3), *Bam*HI (lane 4), or *Bgl*II (lane 5).

due to recombination of the cotransfected DNAs. Therefore, the spliced transcript encodes a functional core protein.

Since the P′ frame in the 2.2-kb RNA can potentially encode a protein that reserved the sequence of the C terminus of the P gene, it is of interest to see whether this protein is required for virion production in the in vitro transfection system. A silent mutation at the splice acceptor site from AG to CG was introduced into plasmid pMH9/3091 to obtain pMH9/3091-ml. Plasmid pMH9/3091 was transfected into HuH-7 cells, and production of HBV-like particles was assayed. Figure 8 shows Northern blot analysis of RNA isolated from HBV DNA-transfected HuH-7 cells using the HBV core sequence as a probe. HuH-7 cells, after having been transfected with wild-type plasmid pMH9/3091, produced both the 3.5- and the 2.2-kb transcripts (Fig. 8, lane 1), while pCH7-8-transfected cells produced only the 2.2-kb transcript (Fig. 8, lane 4). In contrast, in cells transfected with pMH9/3091-ml, most of the 2.2-kb transcript disappeared (Fig. 8, lane 2). HBV-like particles isolated from the media of the cell cultures from which the RNAs were prepared for analysis in Fig. 8 were subjected to an endogenous polymerase assay and restriction mapping (Fig. 9). Although pMH9/3091-ml DNA-transfected HuH-7 cells no longer produced the 2.2-kb transcript, the HBV-like particles were still detectable (Fig. 9B). The levels of HBV-like particles, as judged by the intensity of the labeled HBV DNA, varied in the transfection experiments, probably because of different efficiencies of transfection. The structure of the viral genome contained in the HBV-like particles was analyzed further by restriction analysis of the endogenously labeled HBV DNA (Fig. 9). The patterns of the fragments obtained are as expected from the map of the HBV genome.

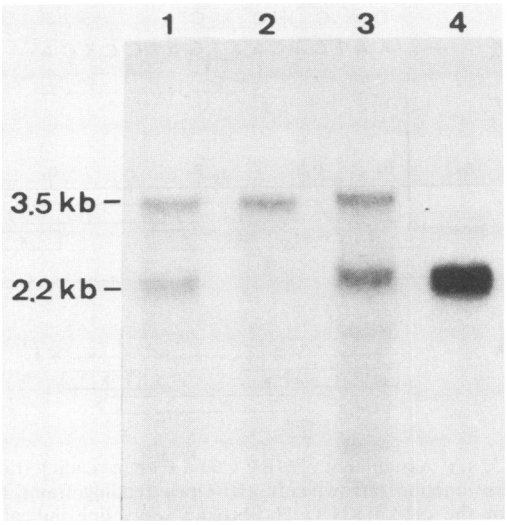

FIG. 8. Northern blot analysis of HBV RNA using a core-specific probe. Fifteen micrograms of total RNA isolated from HBV DNA-transfected HuH-7 cells was used for this analysis. The RNA filter was hybridized with an HBV probe prepared from the *Bgl*II fragment, nucleotides 1984 to 2423. Twenty micrograms of DNA was used in each transfection experiment. Lanes 1 to 4 contained RNAs from HuH-7 cells transfected with pMH9/3091 (lane 1), pMH9/3091-ml (lane 2), pMH9/3091-ml and pCH7-8 (lane 3), or pCH7-8 (lane 4).

DISCUSSION

A new species of HBV transcript was identified in HBV DNA-transfected human hepatoma cells. Analysis of the structure of its cDNA showed that this RNA has a deletion starting from the end of the core gene at nucleotide 2448 to

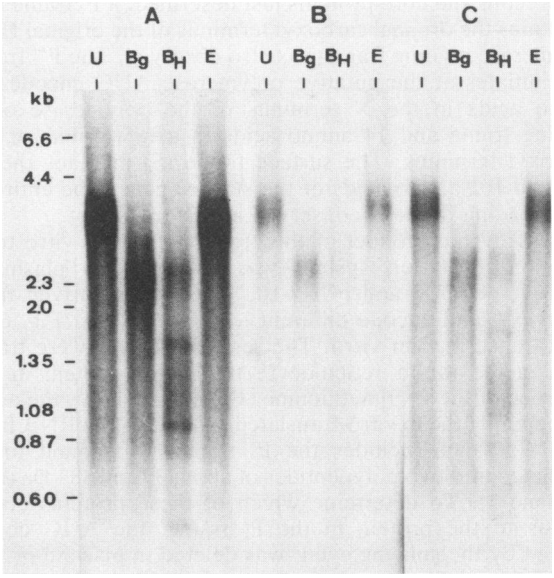

FIG. 9. Restriction analysis of DNA from HBV-like particles. Particles from the medium of HBV DNA-transfected HuH-7 cells were concentrated and radiolabeled by endogenous polymerase activity. The restriction enzymes used were *Bgl*II (Bg), *Bam*HI (B_H), and *Eco*RI (E). U represents undigested DNA. (A) Cells transfected with pMH9/3091; (B) cells transfected with pMH9/3091-ml; (C) cells cotransfected with pMH9/3091-ml and pCH7-8.

the middle of the surface antigen gene at nucleotide 488. Two pieces of evidence show that production of this RNA has undergone a mechanism of RNA splicing. (i) The junction region of the deleted sequences showed the conserved GT-AG sequence, as observed at the junctions of other eucaryotic introns (21). (ii) Mutation at the proposed splice acceptor site resulted in disappearance of this RNA species.

On the basis of S1 nuclease analysis, the 5′ end of this transcript appears to be the same as that of the 3.5-kb species. The initiation sites are mapped downstream of the precore AUG at about the +6 position and upstream of the precore AUG at about the −23 position. Yaginuma et al. (35) have analyzed the HBV transcripts in HBV DNA-transfected HuH-7 cells and found three species of 3.5-kb RNA. In the present study, because of the resolution of the gel system, only two major species were clearly resolved, although the minor but largest species could also be seen. More recently, Enders et al. (7) have shown that in ground squirrel hepatitis virus, the longer 3.5-kb RNA species is engaged essentially in viral protein synthesis while the shortest one is the template for synthesis of the viral genome. S1 nuclease analysis of the 2.2-kb spliced RNA suggests that both classes of the 3.5-kb RNA can serve as precursors for the 2.2-kb transcript.

Sequence analysis of the 2.2-kb RNA suggests that it contains several open reading frames. The core open reading frame can potentially encode a polypeptide that lacks an amino acid at the C-terminal end. By complementation test, this protein was shown to be able to serve as the viral core protein. Whether it also has any regulatory role remains to be determined. One of the open reading frames in this transcript is the P′ frame that can potentially encode a new protein of about 42 kDa. Furthermore, a protein of this size has actually been detected in an in vitro system. Toh et al. (32) have demonstrated amino acid sequence homology between the C-terminal end of the P gene and the reverse transcriptase of retroviruses. The P′ frame in the 2.2-kb transcript appears to comprise all of the homology sequences in the reverse transcriptase domain. It is of interest to determine whether this protein can be detected in vivo and whether this protein possesses reverse transcriptase activity. If this protein could be translated and shown to possess reverse transcriptase activity, it would not be the only viral protein to have this activity, since on mutation of the splice acceptor site which led to disappearance of the 2.2-kb transcript, the viruslike particles could still be produced. Mutations at both the splice donor and acceptor sites did not prevent production of viruslike particles (Su, unpublished data). Thus, the gene products of the 2.2-kb transcript might not be essential for viral replication in the transient DNA transfection system. Recent work on duck HBV has demonstrated that HBV DNA polymerase biosynthesis requires initiation at the first AUG of the P gene (5, 26). However, if this AUG is used in the spliced transcript, only a 60-amino-acid polypeptide could be encoded. On the other hand, recent studies of human T-cell leukemia virus and human immunodeficiency virus have identified several gene products produced by spliced RNA species which are involved in regulation of expression of the virion genes (8). Whether the products of the 2.2-kb transcript also play a regulatory role during the viral life cycle or are involved in some viral transformation processes remains to be determined.

Our previous study indicated that in human liver samples shown to contain the 3.5-kb transcript, the 2.2-kb transcript could also be detected (29). Therefore, this transcript is not only produced in a DNA transfection system but is also a product during natural viral infection. However, other studies of human chronic liver disease and hepatoma have thus far failed to detect this HBV transcript (14, 36). Furthermore, no such RNA species has been found in several HBV animal model systems (2, 6, 20). In light of these results, one would hypothesize that this splicing event involves a certain host factor(s). Only in cells that possess the factor(s) could the 2.2-kb RNA be produced. The factor(s) may not only be species and cell type specific, it may also be differentiation stage specific. It is interesting that although the roles of the long and short species of 3.5-kb RNA are different and they may be present in separate subcellular compartments, both species can be spliced. In contrast to cellular mRNAs, the introns of which are rapidly spliced out, we found that HBV RNA, which contains authentic splice sites, was maintained in a stable and unspliced form in the cells. Apparently, host or viral factors specific for the splicing event may have participated in this regulation, as has been reported for human retroviruses (8).

ACKNOWLEDGMENTS

We thank H. Schaller for plasmids pSHH2-1 and pMH9/3091, J. Sato for the HuH-7 cell line, and K.-B. Choo for critical reading of the manuscript.

This work was supported by research grant NSC77-0610-B075-72 from the National Science Council, Taiwan, Republic of China.

LITERATURE CITED

1. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
2. **Buscher, M., W. Reiser, H. Will, and H. Schaller.** 1985. Transcripts and the putative RNA pregenome of duck hepatitis B virus: implications for reverse transcription. Cell **40**:717–724.
3. **Cattaneo, R., H. Will, G. Darai, E. Pfaff, and H. Schaller.** 1983. Detection of an element of the SV40 late promoter in vectors used for expression studies in COS cells. EMBO J. **2**:511–514.
4. **Chang, C., K.-S. Jeng, C.-P. Hu, S. J. Lo, T.-S. Su, L.-P. Ting, C.-K. Chou, S.-H. Han, E. Pfaff, J. Salfeld, and H. Schaller.** 1987. Production of hepatitis B virus *in vitro* by transient expression of cloned HBV DNA in a hepatoma cell line. EMBO J. **6**:675–680.
5. **Chang, L.-J., P. Pryciak, D. Ganem, and H. E. Varmus.** 1989. Biosynthesis of the reverse transcriptase of hepatitis B virus involves *de novo* translational initiation not ribosomal frameshifting. Nature (London) **337**:364–368.
6. **Enders, G. H., D. Ganem, and H. Varmus.** 1985. Mapping the major transcripts of ground squirrel hepatitis virus: the presumptive template for reverse transcriptase is terminally redundant. Cell **42**:297–308.
7. **Enders, G. H., D. Ganem, and H. Varmus.** 1987. 5′-Terminal sequences influence the segregation of ground squirrel hepatitis virus RNAs into polyribosomes and viral core particles. J. Virol. **61**:35–41.
8. **Franza, B. R., B. R. Cullen, and F. Wong-Staal (ed.).** 1988. The control of human retrovirus gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **Ganem, D.** 1982. Persistent infection of humans with hepatitis B virus: mechanisms and consequences. Rev. Infect. Dis. **4**: 1026–1047.
10. **Ganem, D., and H. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu. Rev. Biochem. **56**:651–693.
11. **Glisin, V., R. Crkvenjakov, and C. Byus.** 1974. Ribonucleic acid isolated by cesium chloride centrifugation. Biochemistry **13**: 2633–2637.
12. **Hattori, M., and Y. Sakaki.** 1986. Dideoxy sequencing method using denatured plasmid templates. Anal. Biochem. **152**:232–238.

13. **Huynh, T. V., R. A. Young, and R. W. Davis.** 1985. Construction and screening cDNA libraries in λ*gt*10 and λ*gt*11, p. 49–78. *In* G. M. Glover (ed.), DNA cloning, vol. I: a practical approach. IRL Press, Oxford.
14. **Imazeki, F., K. Yaginuma, M. Omata, K. Okuda, M. Kobayashi, and K. Koike.** 1987. RNA transcripts of hepatitis B virus in hepatocellular carcinoma. Hepatology **7**:753–757.
15. **Junker, M., P. Galle, and H. Schaller.** 1987. Expression and replication of the hepatitis B virus genome under foreign promoter control. Nucleic Acids Res. **15**:10117–10132.
16. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**:488–492.
17. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
18. **Learned, R. M., R. M. Myers, and T. Tijan.** 1981. *In* D. S. Ray (ed.), The initiation of DNA replication, p. 555–566. Academic Press, Inc., New York.
19. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
20. **Moroy, T., J. Etiemble, C. Trepo, P. Tiollais, and M.-A. Buendia.** 1985. Transcription of woodchuck hepatitis virus in the chronically infected liver. EMBO J. **4**:1507–1514.
21. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
22. **Nakabayashi, H., K. Taketa, K. Miyano, T. Yamane, and J. Sato.** 1982. Growth of human hepatoma cell lines with differentiated functions in chemically defined medium. Cancer Res. **42**:3858–3863.
23. **O'Hare, K., R. Levis, and G. M. Rubin.** 1983. Transcription of the white locus in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **80**:6917–6921.
24. **Saito, I., Y. Oya, and H. Shimojo, H.** 1986. Novel RNA family structure of hepatitis B virus expressed in human cells using a helper-free adenovirus vector. J. Virol. **58**:554–560.
25. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
26. **Schlicht, H.-J., G. Radziwill, and H. Schaller.** 1989. Synthesis and encapsidation of duck hepatitis B virus reverse transcriptase do not require formation of core-polymerase fusion proteins. Cell **56**:85–92.
27. **Simonsen, C. C., and A. D. Levinson.** 1983. Analysis of processing and polyadenylation signals of the hepatitis B virus surface antigen gene by using simian virus 40-hepatitis B virus chimeric plasmids. Mol. Cell. Biol. **3**:2250–2258.
28. **Spandidos, D. A., and N. M. Wilkie.** 1984. Expression of exogenous DNA in mammalian cells, p. 1–48. *In* B. D. Hames and S. J. Higgins (ed.), Transcription and translation. IRL Press, Oxford.
29. **Su, T.-S., W.-Y. Lui, L.-H. Lin, S.-H. Han, and F.-K. P'eng.** 1989. Analysis of hepatitis B virus transcripts in infected human livers. Hepatology **9**:180–185.
30. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
31. **Tiollais, P., C. Pourcel, and A. Dejean.** 1985. The hepatitis B virus. Nature (London) **317**:489–495.
32. **Toh, H., H. Hayashida, and T. Miyata.** 1983. Sequence homology between retroviral reverse transcriptase and putative polymerases of hepatitis B virus and cauliflower mosaic virus. Nature (London) **305**:827–829.
33. **Treinin, M., and O. Laub.** 1987. Identification of a promoter element located upstream from the hepatitis B virus X gene. Mol. Cell. Biol. **7**:545–548.
34. **Wahl, G. M., R. A. Padgett, and G. R. Stark.** 1979. Gene amplification causes over-production of the first three enzymes of UMP synthesis in *N*-(phosphonacetyl)-L-aspartate-resistant hamster cells. J. Biol. Chem. **254**:8679–8689.
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. **Yokosuka, O., M. Omata, F. Imazeki, Y. Ito, and K. Okuda.** 1986. Hepatitis B virus RNA transcripts and DNA in chronic liver disease. N. Engl. J. Med. **315**:1187–1192.