INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.

- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.

- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17"x 23" black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6"x 9" black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.

Zelent, Arthur Zyginant

EXPRESSION OF HEPATITIS B VIRUS GENES IN TISSUE CULTURE CELLS

City University of New York PH.D. 1987

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark __✓__.

1. Glossy photographs or pages ✓
2. Colored illustrations, paper or print _____
3. Photographs with dark background ✓
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages _____
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received _____
16. Other _____

_____

_____

University
Microfilms
International

EXPRESSION OF HEPATITIS B VIRUS GENES IN TISSUE CULTURE CELLS

by

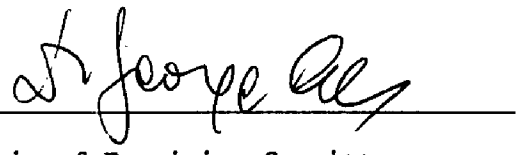
ARTHUR Z. ZELENT

A dissertation submitted to the Graduate Faculty in Biomedical Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

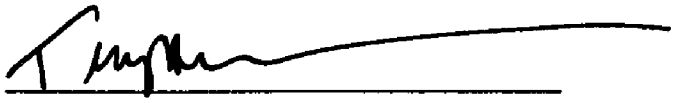
1987

This manuscript has been read and accepted for the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

12/04/1986
Date


Chair of Examining Committee

12/05/1986
Date


Executive Officer

Dr. Judith K. Christman

Dr. Beatrice Pogo

Dr. Peter M. Price

Dr. Norman Gabelman

Supervisory Committee

The City University of New York

Abstract

EXPRESSION OF HEPATITIS B VIRUS GENES IN TISSUE CULTURE CELLS

by

Arthur Z. Zelent

Adviser: Professor George Acs.

In order to study the life cycle of hepatitis B virus (HBV) our laboratory attempted to develop a tissue culture system which would support the viral replication. In the course of this study two cell lines were developed which markedly differed with respect to HBV gene expression.

The 4.10 cell line was established by cotransfection of mouse 3T3 cells with plasmid DNA containing a head-to-tail dimer of the HBV genome and DNA coding for methotrexate resistant dihydrofolate reductase. These cells contain at least 40 copies of intact HBV dimer per cell, and produce large amounts of 22 nm hepatitis B surface antigen (HBsAg) particles that include viral major and middle envelope proteins. They also synthesize viral "e" antigen (HBeAg). However, the presence of core antigen (HBcAg) was not detected.

Analysis of mRNAs isolated from 4.10 cells revealed the presence of two major HBV transcripts [the 2.1 kilobase (kb) mRNA and 3.5 kb longer-than-genome length RNA], although the amount of the latter species was approximately 0.1% of the former. In contrast to this cell line, high levels of the longer-than-genome length RNA and HBcAg expression were observed in D1113 cells. These cells were established by transfecting Moloney murine leukemia virus retroviral vector (carrying two head-to-tail HBV dimers and gene coding for neomycin

resistance) into a transformed mouse fibroblast cell line (ψAM22b). The data presented here represents the first example of a tissue culture system in which the two major HBV mRNAs are expressed in relatively equal amounts, similar to the proportions observed during in vivo infection. The initiation sites for the 2.1 kb and the 3.5 kb mRNAs were mapped 5' to and within the pre-S2 region and 5' to and within the pre-C region of HBV, respectively. These positions correspond very well with those from which the same mRNAs are transcribed in infected livers. Since the D1113 cells produce abundant amounts of the most important marker of HBV replication, the 3.5 kb "pre-genome" RNA, they should have a potential to support at least partial replication of the virus and assembly of "Dane-like" particles.

ACKNOWLEDGEMENTS

I would like to thank Dr. George Acs for his guidance and support during my graduate studies, as well as Dr. Peter M. Price whose continuous help was indispensible. I am also thankful to all our collaborators who made significant contributions to my education and scientific training, especially Drs Mary Ann Sells and Judith K. Christman. In addition, I also thank Dr. Mary Ann Sells for providing me with the 4.10 and D1113 cell lines.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS

aa - amino acid(s).

AMV - avian myeloblastosis virus.

bp - base pair(s).

BpB - brom phenol blue.

BSA - bovine serum albumin.

cpm - counts per minute.

cps - counts per second.

DEPC - diethylpyrocarbonate.

DHBV - duck hepatitis B virus.

DNase - deoxyribonuclease.

DR - direct repeat.

DTT - dithiotrietol.

EDTA - ethylenediaminetetraacetate, disodium salt.

GSHV - ground squirrel hepatitis virus.

HBV - hepatitis B virus.

HBcAg - hepatitis B core antigen.

HBeAg - hepatitis B e antigen.

HBsAg - hepatitis B surface antigen.

HCC - hepatocellular carcinoma.

kb - kilobase.

kd - kilodalton.

LTR - long terminal repeat.

MoMLV - Moloney murine leukemia virus.

MOPS - morpholinopropanesulfonic acid

nt - nucleotide(s).

LIST OF ABBREVIATIONS (continued)

ORF - open reading frame.

PBS - phosphate buffered saline.

pHSA - polymerized human serum albumin.

PIPES - piperazine-N,N'-bis(2-ethane-sulfonic) acid

poly(A)$^+$ - polyadenylated.

poly(A)$^-$ - non-polyadenylated.

PSSC - phosphate saline sodium citrate.

rm temp - room temperature.

RNase - ribonuclease.

rpm - revolution(s) per minute.

SDS - sodium dodecyl sulfate.

SSC - saline sodium citrate.

SV-40 - simian virus 40.

TEMED - N,N,N',N'-tetramethyl-ethylenediamine.

UV - ultra violet.

XCFF - xylene cyanol FF.

INTRODUCTION

Hepatitis B virus is an important human pathogen which persists in approximately 200 million people world-wide and produces a spectrum of clinical disease states. Primary contact of susceptible humans with HBV may result in an inapparent (subclinical) infection or in a symptomatic liver injury (acute hepatitis B). The latter is usually self-limited but often severe. Most individuals, however, recover from the primary infection. In 90 to 95% of cases recovery is associated with production of a neutralizing antibody to the virus, clearance of HBV from the liver and blood, and establishment of life-long immunity to reinfection (Redeker, 1975; Ganem, 1982). Unfortunately, in the remaining 5 to 10% of the individuals, viral infection of the liver is not eradicated, and viral antigens (and sometimes infectious virions) continue to circulate in the blood, usually for life (Ganem, 1982). Most such persistently infected individuals are clinically well (so called asymptomatic carriers), although many have abnormal liver biopsies of variable severity. However, an appreciable fraction of cases develop significant liver injury (chronic hepatitis B and/or cirrhosis) (Koretz et al., 1978; DeGroot et al., 1978). Importantly, certain group of patients with long standing chronic hepatitis B have been shown to be at greatly increased risk for the development of primary hepatocellular carcinoma (HCC), a devastating malignacy for which no effective therapy exists (Ganem, 1982). A recent Taiwanese study (Beasley et al., 1981) documents that the relative risk of HCC development in Asian chronic HBV carriers is 5 to 10 times greater than the relative risk of pack-a-day

cigarette smokers for the development of lung cancer.

Structure of HBV Particle.

Hepatitis B virus is a prototype member of a group of small DNA viruses known as hepadna viridae (Robinson et al., 1982). Members of this group include woodchuck hepatitis virus [WHV (Summers et al., 1978)], ground squirrel hepatitis virus [GSHV (Marion et al., 1980)] and Pekin duck hepatitis B virus [DHBV (Mason et al., 1980)]. These viruses are hepatotropic, have very similar structure and display extensive genomic homology (Mandart et al., 1984).

Most of our current knowledge about the structure of HBV derives from the study of virions purified from carrier serum (Robinson and Lutwick, 1976). Such sera contain four HBV related structures: 22 nm spheres, long filaments of similar diameter, core particles and 42 nm spherical double-shelled particles known as Dane particles. It has been documented (Robinson and Lutwick, 1976) that 22 nm spheres and filaments are devoid of nucleic acid and are aggregates of excess HBsAg. The Dane particle represents intact hepatitis B virion, and HBsAg is the major viral coat protein. All serotypes of HBsAg which have been characterized (Galibert et al., 1979; Valenzuela et al., 1980; Ono et al., 1983) carry a common group determinant a and two mutually exclusive sets of subtype determinants d/y or w/r. Treatment of virions with nonionic detergents (e.g. NP40) results in removal of the outer viral coat and liberation of a particulate core structure. This nucleocapsid-like structure contains the viral DNA polymerase (Robinson, 1975), a protein kinase activity (Gerlich et al., 1982), and bears a distinct antigenic determinant - the HBcAg, a protein of approximately 22,000 daltons (Pasek et al., 1979; Burrell et al., 1979). This protein

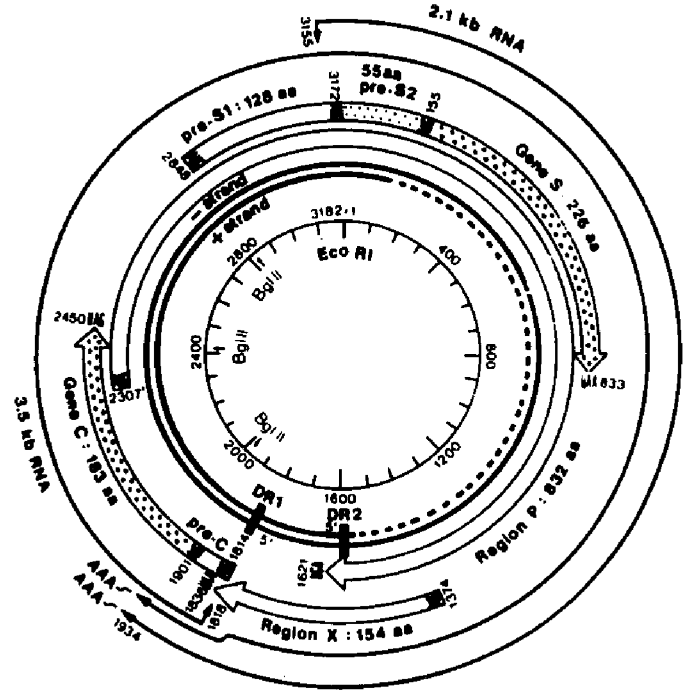
contains a characteristic, arginine rich, protamine-like structure at its carboxyl terminus, which most likely binds HBV DNA. An additional cryptic antigen, the hepatitis B "e" antigen, has also been observed in sera of infected individuals. It has a molecular weight of 15,000 daltons, and was shown to be derived from the HBcAg by proteolysis (Budkowska et al., 1978).

Recently, Gerlich and his co-workers (Stibbe and Gerlich, 1983; Heermann et al., 1984) redefined the HBsAg composition of the viral envelope. They described a major protein, constituting about 85% of the envelope proteins, and two larger proteins comprising the remaining 15%. The major protein, encoded by the S-gene (fig. 1), is 226 amino acids (aa) long and can exist in glycosylated (GP27) and nonglycosylated (P24) form. The middle molecular weight protein is 281 aa long and consists of the S-region plus 55 amino terminal amino acids which are encoded by the pre-S2 region. It exists in monoglycosylated (GP33) and diglycosylated (GP36) forms. The pre-S2 sequence codes for epitopes which are more immunogenic than the S-region alone, and contains a receptor for polymerized human serum albumin (pHSA) (Machida et al., 1984). Since liver cells also contain a receptor for pHSA, it is generally believed that the GP33 and pHSA function in the attachment of HBV to hepatocytes. The largest sAg protein can either be glycosylated (GP42) or nonglycosylated (P39), and is encoded by S, pre-S2 and pre-S1 regions. Recently Neurath et al. (1985, 1986) proposed that the pre-S1 region may also function in the attachment of HBV to liver cells.

Structure of Viral DNA.

The viral DNA is double stranded, approximately 3.2 kb, relaxed circular species with several novel features. First, the molecule is

Figure 1. Structure and genetic organization of the HBV genome. Four open reading frames on the L-strand are indicated by broad arrows. The thin arrows correspond to the two major HBV transcripts and the positions for their initiations and terminations are indicated. DR1 and DR2 are at positions 1824 and 1590, respectively.


2.1 kb RNA
3155
55aa
pre-S2
pre-S1 : 128 aa
3172
155
2848
Gene S : 226 aa
833
− strand
+ strand
3182/1
Eco RI
2800
400
Bgl II
2400
800
2450
2307
Gene C : 183 aa
3.5 kb RNA
2000
1200
1600
DR1
DR2
Pre-C
Region P : 832 aa
1621
1374
1901
1814
1836
1816
1934
AAA~
AAA~
Region X : 154 aa

not a complete duplex. The majority of encapsidated molecules harbor a single stranded region varying in size from 50 to 100% of the genome length (Landers et al., 1977). Although the size of this gap is variable, its polarity is fixed. Thus, the virion DNA consists of a unique long (L) or minus (-) strand of approximately 3.2 kb (the coding strand), annealed to a short (S) or plus (+) strand of less than unit length (the non-coding strand). Covalently linked to the 5' end of the L-strand is a protein which is otherwise not characterized (Gerlich and Robinson, 1980). The circularity of the genome is maintained by 5' cohesive termini of about 200 base pairs (bp) (Sattler and Robinson, 1979). At both sides of these cohesive ends is an 11 bp direct repeat (DR). Since similar DRs are conserved in other hepadna viruses, their function in the life cycle of HBV is most likely significant. Recent studies using animal viruses [DHBV (Molnar-Kimber et al., 1984) and GSHV (Seeger et al., 1986)] have shown that these DRs define the 5' ends of the minus (DR1) and plus (DR2) strands of viral DNA during replication by reverse transcriptase.

Organization of the Genome.

HBV DNA (Galibert et al., 1979; Valenzuela et al., 1980; Ono et al., 1983), as well as DNAs of three animal viruses (Galibert et al., 1982; Mandart et al., 1984; Seeger et al., 1984), were cloned and sequenced. The genomes of all mammalian viruses harbor four major open reading frames (ORF) (C,S,P and X) which are encoded by the L-strand. The DHBV lacks the "X" gene coding domain and contains long C ORF which extends into the region of the genome where the "X" gene is found in other hepadna viruses. The surface antigen coding region was identified by searching for the codons predicted by the known amino terminal aa

sequence of HBsAg (Peterson et al., 1977). As mentioned above, this whole ORF is subdivided into the pre-S1, pre-S2 and S regions. The core antigen region was identified by noting that certain plasmids, bearing only this region, directed HBcAg synthesis in E. coli (Burrell et al., 1979; Pasek et al., 1979). Similarly to the HBsAg coding region, this ORF was divided into the core, and pre-core region which comprises an additional 29 aa at the amino terminus of the core protein. Interestingly, although the HBeAg can be formed in vitro by proteolytic cleavage of HBcAg, translation of the pre-core region also gives rise to HBeAg, which is thought to attach the core to the viral envelope and is released upon disruption of the virion (Ohori et al., 1980). In addition, Ou et al. (1986) showed that expression of the pre-core region of HBcAg targets this protein into the endoplasmic reticular membranes and facilitates secretion of HBeAg. When the pre-C region was deleted from the expression vectors used in these studies, both antigens accumulated in the cytoplasm and neither one was secreted.

Recently, Toh et al. (1983) reported extensive aa sequence homology between the retroviral reverse transcriptase and the putative polymerase of HBV, which is assumed to be encoded by the longest ORF of the virus. Also, Moriarty et al. (1985) were able to raise antibodies to synthetic peptides derived from the "X" region, which reacted with a 28,000 dalton protein from HBV infected-livers, and Meyers et al. (1986) found antibodies in sera from HBV infected patients which reacted with the "X" region-ß-galactosidase fusion protein expressed in E. coli.

Viral Transcripts.

The S-strand has four regions capable of encoding 109, 115, 149 and 170 aa long polypeptides. Since these regions are not strictly

conserved in other hepadna viruses, it is assumed that they do not encode any proteins. To date only our group demonstrated the presence of transcription from the S-strand of HBV DNA (Price et al., 1986). It is not clear why other investigators were not able to detect any S-strand RNA transcripts in either transfected cell lines, or cells derived from liver tumors. It is possible that the methods used in other laboratories were not sensitive enough to detect transcripts present in very small amounts, or the cell lines assayed suppressed production of such an RNA. Interestingly, Standring et al. (1983) reported the presence of a 700 base long RNA polymerase III dependent transcript from the S-strand in a cell free transcription system. This finding indicates that the S-strand contains information which can be recognized by cellular transcriptional machinery, and possibly is transcribed in vivo. Thus far, no reports assessing the presence of S-strand RNA transcripts in the HBV infected chimpanzee livers, or livers of other animals infected with their species specific hepadna viruses have been published.

The transcripts originating from the L-strand of HBV DNA have been characterized for various in vivo and/or in vitro systems. Analysis of RNAs from animal livers infected with their corresponding hepadna viruses revealed the presence of two major transcripts which have the same 3' end and share a common polyadenylation signal (Buscher et al., 1985; Moroy et al., 1985; Enders et al., 1985). For HBV, one transcript is approximately 2.1 kb in length, has several initiation sites about 30 bases around the ATG defining the beginning of the pre-S2 region (Cattaneo et al., 1983 and 1984; Standring et al., 1984) and can serve as a messenger for the major and middle proteins of the envelope.

The second transcript, which is about 3.5 kb in length, probably originates within the pre-core region and can serve as a mRNA for all viral proteins (Cattaneo et al., 1984). In addition, experiments with cell free transcription systems (Rall et al., 1983) demonstrated the presence of two RNA polymerase II-dependent transcripts with origins at position 1680 and 2810 on the L-strand (the nucleotide positions are relative to the unique EcoRI site at position 1, see fig. 1). These mRNAs could be used for translation of the pre-core and pre-S1 regions, respectively. Recently, Siddiqui et al. (1986) also showed a functional initiation site at around nucleotide 2810 in mouse fibroblast cells transfected with a plasmid carrying the entire HBsAg gene under control of its own promoters; and Roossinck et al. (1986), working with a similar system, found an additional initiation site at position 1790 from which the "pre-genomic" RNA, capable of translating the pre-C region, can be transcribed.

Signals Regulating Transcription.

At present little information is available about the sequences and factors regulating HBV transcription. Cattaneo et al. (1983) described the presence of a sequence within the pre-S1 region resembling the late promoter of the Simian Virus 40 (SV40) and sequence analysis revealed a number of 'TATA'-like promoter sequences which could be recognized *in vivo*. Also the sequence TATAAA was found at nucleotide position 1916 (about 20 bp from the 3' end of the two major RNAs) and was shown by mutational analysis to be a functional polyadenylation signal (Simonsen and Levinson, 1983). Since the initiation of the "pre-genomic" RNA is approximately 100 bp upstream of this sequence, the RNA polymerase II - in order to synthesize the longer-than-genome length RNA - must first

bypass this signal and then recognize it at the second encounter. The mechanism which confers on the HBV polyadenylation signal this subtle degree of "leakiness" is not understood.

Recently Shaul et al. (1985) documented the presence of an enhancer-like element in the HBV genome (positioned between HBsAg and "X" region ORFs), which they think could control transcription of the 3.5 kb (longer-than-genome length) RNA. In our laboratory, we showed (unpublished observation) that this element exhibits preferential activity in cells of liver origin. Tissue and host specificity of the HBV enhancer was also documented by Jameel and Siddiqui (1986), who showed that its activity is restricted to liver cells of human origin, and requires cellular trans-acting factor(s) which are independent of the viral gene products. If, indeed, transcription of the 2.1 kb mRNA is constitutive and 3.5 kb mRNA inducible, one can explain why rat or mouse cells transfected with the HBV DNA produce 10 to 100 times less of the 3.5 kb mRNA than 2.1 kb mRNA (Gough, 1983; Pourcell et al., 1982; Zelent et al., submitted).

Recently, Tur-Kaspa et al. (1986) found that the HBV genome contains a glucocorticoid responsive element, which they localized to a position upstream of the previously defined enhancer (around nucleotide 600). Although they did not show that this element can function independently of the enhancer, they ruled out the possibility of direct interaction between the enhancer and glucocorticoid-receptor complexes.

Replication of HBV.

Replication of HBV (Summers, 1984) is believed to proceed through a reverse transcriptase mediated mechanism involving the following steps: 1) the longer-than-genome length RNA is synthesized; 2) this mRNA

("pre-genome") is packaged with viral polymerase into immature cores; 3) the "pre-genome" is then reverse transcribed into the L-strand DNA - the 5' genome-bound protein is thought to function as a primer for this synthesis at the DR1 region; 4) the S-strand is synthesized. Recently, it was shown for the DHBV (Lien et al., 1986) and GSHV (Seeger et al., 1986) that approximately 20 bases at the 5' end of the S-strand consist of ribonucleotides; therefore, it is believed that a short RNA, probably originating from the DR1 region of the degraded "pre-genome", serves as a primer for S-strand synthesis at the DR2. The above described replication mechanism is believed to be applicable to all members of the hepadna viridae group and was developed using DHBV as a prototype (Summers and Mason, 1982; Mason et al., 1982).

Comparison of the hepadna viruses shows that mammalian viruses are more similar in structure and sequence to each other than to DHBV. However, the homologies which they share with the avian virus are also very significant. The most important is the finding that the longest reading frame (possibly encoding DNA polymerase) exists, and has a conserved aa sequence in all hepadna viruses. This supports the hypothesis that these viruses replicate via reverse transription of their "pre-genomes", and perhaps the duck virus evolved earlier than the mammalian viruses from a common ancestor.

Recently, similar results to those reported by Summers and Mason were obtained for HBV. Blum et al. (1984), by finding a heterogenous population of rapidly migrating single stranded HBV DNA species (L-strand polarity), demonstrated the asymmetric properties of HBV replication in human liver. At the same time Miller et al. (1984) reported findings of single stranded HBV DNA of (-)-polarity in virions

isolated from plasma and cytoplasm of liver cells. They argued that this DNA is duplexed with RNA, but the data which they presented is not very convincing. The only evidence showing presence of HBV DNA:RNA hybrids in their system is cesium sulfate density gradients of isolated nucleic acids, which could produce artifacts due to the phenomenons of salt precipitation, and aggregation of nucleic acids (Lozeron and Szybalski, 1966). They also reported presence of full genome length DNA:RNA hybrids, which is in disagreement with the results obtained by Summers and Mason who showed that degradation of the template is concomitant with L-strand synthesis.

The evidence discussed in these last paragraphs gives strong support for the authenticity of the proposed mechanism of replication of HBV. Additional work is necessary to show that the HBV L-strand DNA synthesis is RNA dependent and to isolate and characterize the viral reverse transcriptase.

Similarity of HBV to Retroviruses; Viral Integration and Hepatocellular Carcinoma.

During the last couple of years a great deal of attention has been devoted to similarities which exist between the hepadna viruses and retroviruses. Both viral families involve a reverse transcription step during their replication, however, the sequence of events which occurs during the HBV life cycle represents a permutation of the retroviral replication process. In analogy to retroviruses, the coding potential of HBV seems to be contained only in one DNA strand (the L-strand) of the virus, with similar genomic organization. Recently Miller and Robinson (1986) performed computer-assisted DNA and protein sequence analyses of all known hepadna viruses and found significant homologies

with several type C retroviruses in the corresponding regions of their genomes. They concluded that the hepadna viruses and retroviruses evolved from a common ancestor, and that the HBV arose through a process of deletion from a retrovirus.

Although HBV can integrate into the human genome, in contrast to retroviruses this step is not essential for viral replication. In fact, since the longer-than-genome length RNA can not be synthesized from a single copy of integrated linear genome, the integration step should abolish the viral replication. Additionally, all of the integrated HBV sequences characterized to date represent partial genomes; and consequently, in tumors which only show integrated HBV sequences, viral replication is not observed. It has also not been possible to implicate similar processes of oncogenesis in the development of HCC that are thought to occur in tumors related to the retroviral integration. Since the HBV sequences do not hybridize with the human genome, it was concluded that the virus does not contain any exogenous genes, i. e. oncogenes; and in addition, previous analyses of hosts' flanking sequences did not reveal HBV integration sites in the proximity of any cellular oncogenes (Koch et al., 1984; Varmus, 1984). Recently, however, Dejean et al. (1986) found integrated HBV DNA fused in frame with a cellular coding region which resembles in sequence the oncogene v-erb-A and supposed DNA binding domain of the human glucocorticoid and estrogen receptor genes. Nevertheless, this result is descriptive of only one cell line and, although many other tumors have been examined, nobody else has reported similar findings. Consequently, this finding probably does not reflect a general mechanism of HBV carcinogenesis.

In many instances it has been shown that integration of the HBV DNA is associated with numerous rearrangements, deletions and duplications of the viral and host's sequences (Shaul et al., 1984; Koch et al., 1984, Koshy et al., 1984). It is possible that, by introducing a great degree of instability into the host's DNA, integrated HBV can act as a mutagen to activate a proto-oncogene or deactivate a gene which is involved in regulation of cell proliferation and differentiation. However, no evidence exists that HBV sequences have any mutagenic properties and this view remains very speculative. Another intriguing hypothesis proposes that the integration of HBV enhancer can activate the expression of a cellular oncogene to abnormally high levels, thus causing cell transformation. This hypothesis is also unsupported experimentally and even less convincing than the former, since no known oncogenes have been reported to be expressed at abnormally high levels in liver tumors which were found to contain HBV DNA sequences. At present, it is still not known how HBV can cause HCC and it remains possible that the viral integration plays an important but a less direct role in liver oncogenesis.

In Vitro Expression of HBV Genes.

Transfection of mammalian cells with a variety of plasmids containing all or part of the HBV genome has led to success in obtaining significant *in vitro* production of the viral envelope proteins (Christman et al., 1982; Pourcell et al. 1982; Liu et al., 1982; Michel et al., 1984). Production of viral "e" antigen and core proteins has also been detected in several transfected mammalian cell systems (Gough and Murray, 1982; Price et al., 1983; Will et al., 1984). However, in the absence of heterologous promoters of core gene transcription, the

amounts of 3.5 kb RNA and core protein produced were, in majority of cases, close to the limits of detection by available techniques (Price et al., 1983).

This study describes two HBV transfected murine cell lines which differ greatly, with respect to each other, in expression of the HBV genes. The 4.10 cell line (derived from NIH 3T3 cells) contains very little of the 3.5 kb RNA and produces virtually no HBcAg. In contrast, the D1113 cells - derived from ψAM22b cell line which contains integrated, replication defective Moloney murine leukemia virus (MoMLV) proviral DNA - express high levels of the longer-than-genome length RNA, and synthesize large amounts of HBcAg.

MATERIALS AND METHODS

Cloning of the HBV genome (subtype ayw), which was used in this study, has been previously described (Price et al., 1980). The size of the genome of this isolate is 3182 bp, and the numbering system used sets nucleotide 1 at the unique EcoR1 site (see fig. 1).

Cells and Culture Conditions.

The 4.10 cell line was established by transfecting NIH 3T3 cells (see Christman et al., 1982) with pTHBV-1 (pBR322 carrying a head-to-tail tandem repeat of the HBV genome cloned into the EcoRI site) and sheared genomic DNA from A29 hamster cells [a line with multiple copies of genes coding for a methotrexate-resistant dihydrofolate reductase (Flintoff et al., 1976)]. Cells from colonies arising after selection for methotrexate resistance at 0.2 µg/ml were picked with a sterile pipet tip, transferred to individual wells in a 24 well plate and, when sufficiently grown, tested for production of HBsAg. Cells from wells with medium containing HBsAg (S/N>5) were seeded into medium containing 40 µg/ml methotrexate, cloned and tested for amplification of HBsAg production. Clone 4.10 was one of a group of clones isolated by this procedure. It has been maintained for more than 3 years in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 250 units/ml penicillin, 0.2 µm/ml streptomycin and 40 µg/ml methotrexate. All cultures were maintained at 37°C in a moist atmosphere containing 5% CO_2.

The methods used in the development of D1113 cell line were essentially the same as those used in the establishment of 4.10 cells. In short, the ψAM22b cell line was transfected with 10 µg of pDo1THBV-1 (fig. 3) plasmid DNA in a calcium phosphate precipitate (Graham and van

Figure 2. Map of pTHBV-1 indicating the position of cleavage sites for restriction enzymes used in preparation of the probes specific for transcripts from the (-)-strand of HBV genome. Also illustrated are the restriction fragments subcloned into SP64 and/or SP65 vectors for preparation of probes that detect transcripts from the pre-S-S region (----), probe 1; the "X" region (——), probe 2; and the core region (——), probe 3, of the HBV genome. The HBV sequences are indicated with a heavier line and the coding regions - HBsAg, HBcAg, polymerase and "X" - are indicated by symbols S, C, P, and X, respectively.

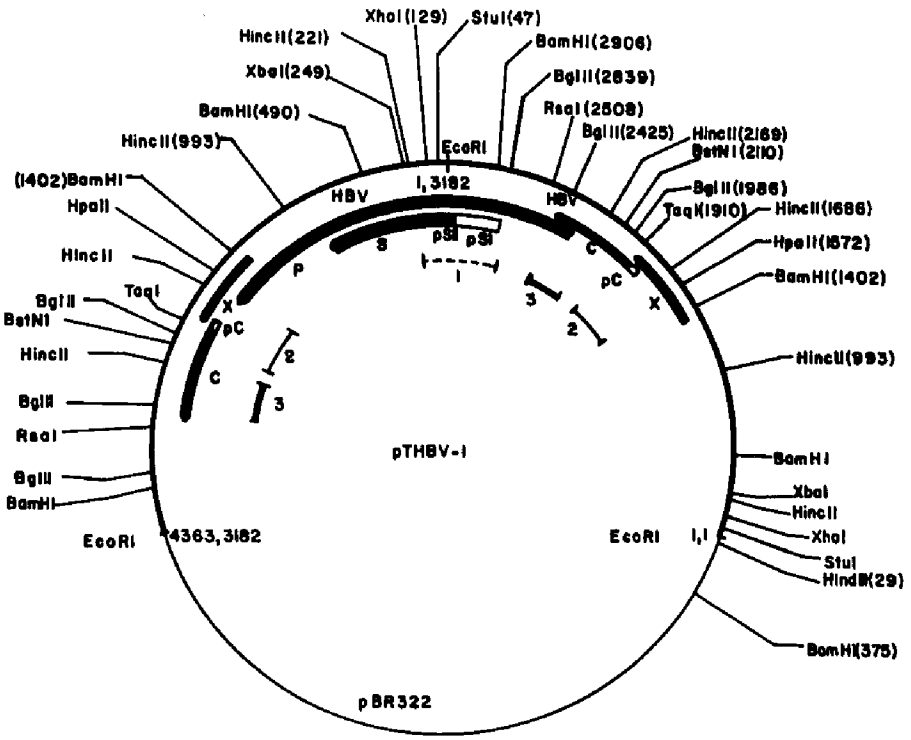
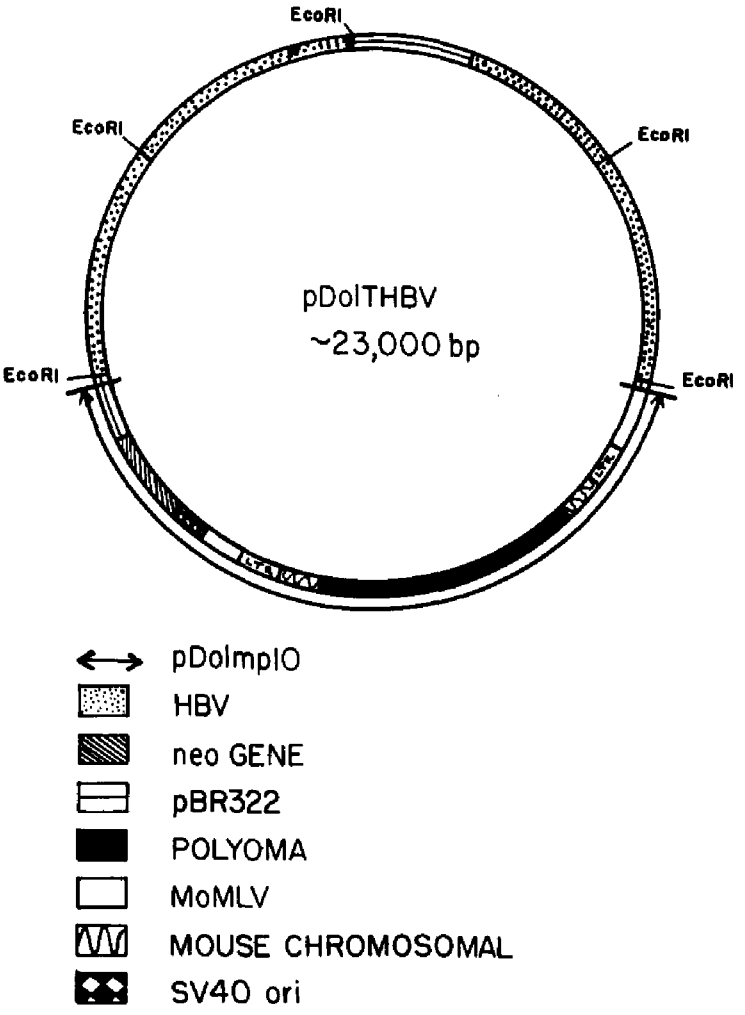

XhoI(129)
StuI(47)
HincII(221)
BamHI(2906)
XbaI(249)
BglII(2839)
BamHI(490)
RsaI(2508)
BglII(2425)
HincII(2169)
HincII(993)
EcoRI
BstNI(2110)
(1402)BamHI
HBV
1,3182
BglII(1986)
HBV
HpaII
TaqI(1910)
HincII(1686)
HincII
P
S
pSI
pSI
C
pC
X
HpaII(1672)
BamHI(1402)
TaqI
BglII
X
pC
BstNI
HincII
C
2
3
3
2
HincII(993)
BglII
RsaI
pTHBV-1
BamHI
BglII
XbaI
BamHI
HincII
EcoRI
4363,3182
EcoRI
1,1
XhoI
StuI
HindIII(29)
BamHI(375)
pBR322

Figure 3. The structure of plasmid pDolTHBV-1. The regions corresponding to various sequences are represented by different sets of lines which are shown and defined below the diagram.


EcoRI
EcoRI
EcoRI
EcoRI
EcoRI
pDolTHBV
~23,000 bp
pDolmpIO
HBV
neo GENE
pBR322
POLYOMA
MoMLV
MOUSE CHROMOSOMAL
SV40 ori

der Eb, 1973) - following the method of Wigler et al. (1980). ψAM22b cells and pDo1THBV-1 are described in the following paragraphs. After transfection, cells were selected for G418 (analogue of neomycin) resistance. The clone D1113 was derived from G418-resistant cells which produced significant levels of HBsAg. It has been maintained in the same medium and under the same atmospheric conditions as 4.10 cells, except - instead of methotrexate - 300 μg/ml of G418 is included in the medium in order to selectively maintain the transfected genotype.

The cell line ψAM22b (a generous gift from R. C. Mulligan), derived from NIH 3T3 cells, contains permanently integrated MoMLV provirus in which the ψ sequences were deleted, and the 3' end replaced with the same region from another MLV which possesses an amphotropic envelope gene (allowing recognition of receptors on broad spectrum of mammalian cells). The ψ sequences, which are adjacent to the 5' LTR, are necessary in cis for recognition and packaging of the viral nucleic acid. Due to their absence, this cell line constitutively produces all the viral proteins (including reverse transcriptase), but does not assemble viable virions.

Plasmid pDo1THBV-1 was derived from a retroviral vector pDo1mp10 (a generous gift from R. C. Mulligan) by insertion of BanI fragment from pTHBV-1, containing two tandem copies of the HBV genome and the majority of the ß-lactamase gene, into its unique XhoI site through blunt end ligation after Klenow treatment of both DNAs. After transfection of E. coli HB101 cells, ampicillin resistant colonies were isolated and plasmids harvested by standard methodologies (Maniatis et al., 1982). Restriction enzyme analysis of the plasmid showed that it contains two copies of head-to-tail HBV dimers, joined together through the pBR322

sequences and situated in opposing polarities with respect to each other (fig. 3). The AM22b cell line and plasmid pDolmp10 represent the retroviral vector system which was developed, and described in detail by Mann et al. (1983), and Cone and Mulligan (1984).

Cellular DNA Isolation.

The DNA was isolated essentially as described by Jeffrey and Flavell (1977). Cells in confluent monolayers were washed with 1 x phosphate buffered saline [1 x PBS (0.14 M NaCl, 2.7 mM KCl, 8.1 mM Na_2HPO_4, and 1.5 mM KH_2PO_4; pH 7.4)], treated with 0.05% trypsin and pelleted by centrifugation at 250g for 10 min at 4°C. After two additional washes with 1 x PBS, the cells were resuspended in 20 volumes (relative to the volume of the cell pellet) of 0.15 M NaCl/0.1 M ethylenediaminetetraacetic acid-disodium salt (EDTA), pH 8.0 solution. Subsequently sodium dodecyl sulfate (SDS) was added to a final concentration of 1% and the cells were lysed (on ice) with a lose-fitting Dounce Homogenizer, 20 to 30 times. After addition of proteinase K (final concentration 400 µg/ml) the mixture was incubated for 3 hrs at 37°C. At the end of this incubation the solution was made 1 M with respect to $NaClO_4$ and extracted two times with an equal volume of chloroform/isoamylalcohol [in 24/1 (v/v) ratio, respectively]. The two phases were separated by centrifugation at 4°C and 6,000g for 10 min, the aqueous phase was removed with a wide tip pipette, and DNA was precipitated after adding 0.25 volumes of 10 M NH_4OAc (final concentration 2 M) and 2 volumes of cold 95% ethanol. The precipitated DNA was sedimented by centrifugation at 5,000g for 30 min at 4°C, washed with 75% ethanol, vacuum dried and resuspended in TE buffer (10 mM Tris-HCl, pH 7.5; 0.05 mM EDTA). After addition of pancreatic

ribonuclease A (RNase A) and RNase T1 to a final concentration of 20 µg/ml and 20 U/ml, respectively, the sample was incubated for 1 hr at room temperature (rm temp). Subsequently, proteinase K was added to a final concentration of 100 µg/ml and the incubation was continued for additional 2 hrs. At the end of this incubation the sample was extracted with an equal volume of phenol/chloroform/isoamylalcohol mixture (v/v ratio 25/24/1, respectively), the two phases were separated by centrifugation and DNA was recovered from the aqueous phase by precipitation, as described above. The resuspended DNA was stored at 4°C in TE buffer until used in Southern analysis.

RNA Isolation.

When working with RNA, all solutions were either made up using diethylpyrocarbonate (DEPC) treated dH_2O or were treated overnight with 0.1% DEPC and then autoclaved. All materials - such as eppendorf centrifuge tubes, pipette tips, etc. - were autoclaved for at least one hr.

All RNAs were isolated by the method developed by Chirgwin et al. (1974). Cells were lysed by the addition of 2 ml of guanidinium thiocyanate solution (4 M guanidinium thiocyanate, 0.5% sarcosyl NL-30, 25 mM sodium citrate (pH 7.0), 0.1 M 2-mercaptoethanol and 0.1% antifoam A) per 175 cm^2 tissue culture flask with cells growing in a semi-confluent monolayer. The cell lysate was then dounced 10 times on ice, transferred to a 30 ml Corex centrifuge tube and mixed with 0.025 volume of 1 M HOAc and 0.75 volume of 95% ethanol. The tube was sealed, shaken thoroughly and nucleic acids were precipitated at -20°C overnight. The precipitated material was collected by centrifugation for 10 min at -10°C and 6,000 revolutions per minute (rpm) in a

Sorvall HB-4 rotor. The supernatant was discarded and the pellet resuspended, with vigorous shaking, in 0.5 volume (relative to the original volume of the cell lysate) of guanidine hydrochloride solution [neutralized to pH 7.0, buffered with 0.025 volume of 1 M NaOAc (pH 7.0), and made 5 mM in dithiothrietol (DTT)]. Following the addition of 0.025 and 0.5 volume (relative to the guanidine hydrochloride solution) of 1 M HOAc and 95% ethanol, respectively, the nucleic acids were precipitated and centrifuged as before. This last step was then repeated, again halving the total volume, and the precipitated material was sedimented by centrifugation for 5 min at $-10^{o}C$ and 6,000 rpm in HB-4 rotor. The RNA pellet was washed with 75% ethanol to remove excess guanidine hydrochloride, vacuum dried and dissolved in DEPC treated dH_2O. For storage, the aqueous solution of RNA was mixed with 0.1 volume of 5 M KOAc (pH 5.0) and 2 volumes of 95% ethanol and placed at $-20^{o}C$.

Isolation of Polyadenylated RNAs.

Polyadenylated [Poly(A)$^{+}$] RNAs were separated from ribosomal RNA species by two cycles of binding to oligo(dT)-cellulose (Type T-3) as described by Aviv and Leder (1972). The total RNA was taken up in ETS buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA, 0.5% SDS) at concentration not greater than 1 mg/ml, heated for 5 min at $85^{o}C$ and rapidly cooled to rm temp. After addition of 5 M NaCl, to a final concentration of 0.5 M, the sample was applied to 0.25 gm of oligo(dT)-cellulose column - equilibrated with the same buffer. The flow-through RNA was collected and designated poly(A)$^{-}$-(1). The column was washed with approximately 15 volumes (relative to the column volume) of the binding buffer (ETS/0.5 M NaCl) and the bound RNA was eluted with 4 volumes of 10 mM

Tris-HCl (pH 7.5), 1 mM EDTA and 0.05% SDS (ET/0.05% SDS) buffer. The eluted fraction was designated poly(A)$^+$-(1). After washing the column, consecutively, with 10 volumes of ETS, ETS/0.1 M NaOH and ETS/0.5 M NaCl, the poly(A)$^-$-(1) fraction was reapplied to it. The flow-through poly(A)$^-$ RNA was collected and precipitated, after adjusting the NaCl concentration to 1 M, by adding 2 volumes of 95% ethanol. The bound RNA was washed and eluted as described before, and pooled with fraction poly(A)$^+$-(1). This combined solution was made 0.5 M with respect to NaCl and reapplied to the column, previously washed with ETS, ETS/0.1 M NaOH and ETS/0.5M NaCl as described above. The column was washed again with 15 volumes of the binding buffer, the poly(A)$^+$ RNA was eluted with 4 volumes of ET/0.05% SDS buffer, made 1 M in NaCl, and precipitated by adding 2 volumes of 95% ethanol. After precipitating overnight at -20^{o}C, the poly(A)$^+$ and poly(A)$^-$ RNAs were collected by centrifugation at 4^{o}C and 10,000g for 30 min. The RNA pellets were washed with 75% ethanol, vacuum dried, resuspended in DEPC treated dH_2O, and stored in 70% ethanol-KOAc suspension, as described previously.

Agarose Gel Electrophoresis and Southern Blotting.

All restriction enzyme digestions referred to in this study were carried out under the conditions recommended by the supplier. Digested 4.10 or D1113 genomic DNAs were loaded, in the digestion buffers supplemented with 0.2 volume of 6 x loading buffer [0.25% brom phenol blue (BpB), 0.25% xylene cyanol FF (XCFF), and 30% glycerol], onto a 1% agarose gel and electrophoresed in 1 x TAE (40 mM Tris-acetate, pH 8.0; 2 mM EDTA) buffer. After the electrophoresis, part of the gel containing molecular size markers was stained in 1 µg/ml ethidium

bromide solution and DNA bands were visualized by short wave ultra violet (UV) light illumination. DNA in the remaining part of the gel was denatured by soaking it twice in 1.5 M NaCl/0.5 M NaOH solution, 15 min each time. After rinsing it with dH_2O, the gel was neutralized by two washes in 1M Tris-HCl (pH 8.0)/1.5 M NaCl solution, 15 min each time. Subsequently, DNA in the gel was transferred overnight to BA85 nitrocellulose membrane (pore size 0.45 μm) using 20 x SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0) as the blotting buffer. Following blotting, the nitrocellulose paper was rinsed in 2 x SSC for 10 min, air dried, and baked in a vacuum oven at 80°C for 2 to 3 hrs.

The above blotting procedure was developed by Southern (1975), and was outlined in greater detail by Maniatis et al. (1982).

Gel Electrophoresis of RNA and Northern Blotting.

RNAs were electrophoresed on vertical 1% agarose gels following the method of Lehrach et al. (1977). The gel running buffer was 40 mM in morpholinopropanesulfonic acid (MOPS), pH 7.0, 10 mM in sodium acetate, and 1 mM in EDTA; and the gel buffer was in addition 1.1 M in formaldehyde. The running buffer was recirculated throughout the electrophoresis in order to provide an even concentration of all the solutes in top and bottom chambers of the apparatus. The samples were loaded in solutions of the same composition as the running buffer, which in addition contained 45% formamide [deionized by passing through AG501-X8(D) resin], 2.5% ficoll 400, 0.04% BpB and 0.04% XCFF. Before loading, all the samples were incubated at 65°C for 3 min and quickly chilled on ice. After electrophoresis, part of the gel containing size markers was cut off and stained in 5 μg/ml ethidium bromide solution. Following the destaining in dH_2O for at least 1 hr, the RNA ladder was

visualized by short wave UV illumination. The remaining portion of the gel was soaked 2 times in 20 x SSC, 10 min each time, and RNA was transferred overnight to BA85 nitrocellulose membrane, using 20 x SSC as the blotting buffer. After the transfer, the blots were air dried and baked in a vacuum oven for 2.5 hrs. Following the hybridization and autoradiography, in order to visualize the 18S and 28S ribosomal RNAs, nitrocellulose membranes were first soaked in 5% HOAc for 10 min, then in 0.5 M sodium acetate (pH 5.2)/0.04% methylene blue solution for 10 min, and rinsed several times in dH_2O.

Preparation of Probes.

1. RNA probes (Riboprobes).

Strand specific riboprobes were synthesized using subgenomic HBV fragments (fig. 2) cloned into pSP64 and/or pSP65 in vitro transcription vectors [developed and described in detail by Melton et al. (1984)] as template and the SP6 RNA polymerase. The transcription reaction (40 mM Tris-HCl, pH 7.5; 6 mM $MgCl_2$; 10 mM NaCl; 2 mM spermidine; 10 mM DTT; 350μCi ^{32}P-α-GTP, specific activity 410 Ci/mmole; 0.5 mM CTP; 0.5 mM UTP; 0.5 mM ATP; 0.02 μg/μl of template DNA; 1 U/μl RNasin; and 10 U/μg of template DNA of SP6 RNA polymerase, in a total volume of 50 μl) was carried out in siliconized 1.5 ml eppendorf centrifuge tubes at 40°C for 1 hr. Following the incubation RQ1™ DNase was added to a final concentration of 2 U/μg of template DNA and the mixture was incubated at 37°C for 10 min. Subsequently, the reaction volume was adjusted to 100 μl and the solution was extracted with phenol/chloroform/isoamylalcohol mixture. The aqueous phase was transferred to a new eppendorf centrifuge tube and RNA was

precipitated by addition of 0.1 volume of 5 M KOAc (pH 5.0), 3 volumes of 95% ethanol and incubation at -70°C for 30 min. Precipitated nucleic acids were collected by 15 min centrifugation in an eppendorf table top centrifuge at 10,000g and 4°C, washed with 75% ethanol and vacuum dried. Samples were then taken up in dH_2O containing 2.5% ficoll 400, 0.04% of BpB and XCFF, and electrophoresed at 120 V in a horizontal 1.4% agarose gel with 1 x TAE as the running buffer. After electrophoresis the, gel was stained for 30 min in 1 µg/ml ethidium bromide solution and RNA bands were visualized by long wave UV illumination. The gel regions which contained the riboprobes were then cut out, and subsequently crushed by passage through a 1 ml Hamilton syringe. The RNA probes were eluted from the crushed gel at 37°C for 2 hrs in 10 mM Tris-HCl (pH 7.5)/1 mM EDTA/0.2 M NaCl buffer and precipitated with 50 µg of E. coli tRNA after addition of 0.1 volume of 5 M KOAc (pH5.0) and 2 volumes of 95% ethanol at -70°C for 30 min. The precipitate was collected by centrifugation at 4°C and 8,000g, washed once with 75% ethanol and vacuum dried. Subsequently, the probes were taken up in 0.5 ml of TE buffer (10 mM Tris-HCl, pH 8.0; 1 mM EDTA) and passed through the Sephadex G-50 (medium) column equilibrated with the same buffer. The radioactive material was recovered in three 1 ml fractions and first two of those were directly added to the hybridization solution. The third fraction was discarded.

2. Nick translation of viral DNA.

For the preparation of nick translated probes, HBV DNA was isolated from pTHBV-1 after cleaving the plasmid with EcoR1, separating it from pBR322 by electrophoresis in a 1% agarose gel and electroeluting the HBV genomic DNA into 0.5 x TAE for 30 min at 120 V. The 0.5 x TAE buffer

containing eluted HBV DNA was then applied to NACS PREPAC™ Convertible mini columns, previously equilibrated with the binding buffer (10 mM Tris-HCl, pH 7.5; 1 mM EDTA; and 0.5 M NaCl). Subsequently, the column was washed with 5 ml of this buffer and bound material was eluted with 0.4 ml of 10 mM Tris-HCl (pH 7.5)/1 mM EDTA/2 M NaCl buffer. The sample was extracted with phenol/chloroform/ isoamylalcohol and the aqueous phase was transferred to a new 1.5 ml eppendorf centrifuge tube. The DNA was precipitated by adding 2 volumes of 95% ethanol and incubating at -20°C overnight or at -70°C for 30 min. The precipitate was collected by centrifugation at 4°C and 10,000g, washed with 75% ethanol, vacuum dried, and resuspended in dH_2O at concentration of 50 µg/ml. Nick translation [described by Rigby et al. (1977)] was carried out at 14°C for 3 hrs in 25 µl of solution containing the following: 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 0.1 mM DTT, 50 µg/ml bovine serum albumin (BSA), 60 µM dATP and dGTP, 10 µM TTP, 6 µM dCTP, 250 µCi ^{32}P-α-dCTP (specific activity 3000 Ci/mmole), 250 ng of HBV DNA, 50 pg of DNase (from bovine pancreas), and 12 U of E. coli DNA polymerase I. After incubation, the reaction mixture was made 20 mM in EDTA, diluted to 0.3 ml with TE (pH 8.0) and passed through the Sephadex G-50 (medium) column, as described for riboprobes, in order to separate the probe from unincorporated radioactivity. This way prepared probe was ready to be used in the hybridization solution.

3. 5' and 3' labeled probes.

Probes labeled at either the 5' or 3' end were prepared by kinating dephosphorylated 5' extensions and Klenow filling of the recessed 3' ends, respectively (Maxam and Gilbert, 1980). In order to obtain a

probe labeled only in one strand, labeled restriction fragments were redigested with an additional restriction enzyme and run on 4, 6, or 8% polyacrylamide gels, depending on a size of the fragment sought. The desired band was cut out from a gel after staining in 1 µg/ml ethidium bromide and visualizing by long-wave UV illumination. The probes were isolated by electroelution and purified over NACS PREPAC^{TM} Convertible mini columns as described in the preceding section.

A. Dephosphorylation reaction.

The phosphatase reaction was carried out for 1 hr at 66^{o}C in 200 µl of 50 mM Tris-HCl (pH 8.0) solution, supplemented with not less than 1 U of bacterial alkaline phosphatase per 7 pmoles of 5' ends. Following this reaction, 2 µl of 0.5 M EDTA (pH 8.0) were added and the solution was heated at 50^{o}C for 10 min.

B. Kination of protruding 5' ends.

The kination reaction was carried out at 37^{o}C for 1 hr using 1 U of phage T4 polynucleotide kinase for 2.5 to 15 pmoles of 5' ends, and ^{32}P-γ-ATP (specific activity 5000 Ci/mmole) at concentration greater or equal to 1 pmole of radiolabeled ATP per 1 pmole of 5' ends. The Reaction buffer was 50 mM Tris-HCl (pH 7.5)/10 mM $MgCl_2$/5 mM DTT/1 mM spermidine/1 mM EDTA solution, and the reaction volume was usually 50 µl.

C. Kination of recessed or blunt 5' ends.

The DNA fragments (usually 100 pmoles of 5' ends or 50 µg of total DNA) were taken up in 35 µl of "A" buffer (20 mM Tris-HCl, pH 9.5; 1 mM spermidine; 0.1 mM EDTA), heated at 90^{o}C for 2 min, and quick chilled on ice. Immediately after chilling, 5 µl of "B" buffer (500 mM Tris-HCl, pH 9.5; 100 mM $MgCl_2$; 50 mM DTT; and 50% glycerol), and the

appropriate amounts of radiolabeled ATP and polynucleotide kinase were added. The final reaction volume was 50 µl. The reaction was carried out at 37°C for 1 hr and then the pH of the sample was neutralized by adding 1.6 µl of 2 M Tris-HCl.

D. Klenow reaction.

The recessed 3' ends were filled with E. coli DNA polymerase I "large fragment" (Klenow) enzyme. This reaction was carried out in total volume of 30 µl for 1 hr at rm temp. The reaction solution consisted of 50 mM NaCl, 10 mM $MgCl_2$, 10 mM Tris-HCl (pH 7.5), 25 µM of TTP/dGTP/dCTP, 5 µg of DNA fragments, 200 µCi of ^{32}P-α-dATP (specific activity 3000 Ci/mmole), and 10 U of Klenow enzyme.

At the end of each of the above procedures (A through D), every solution was extracted two times with phenol/chloroform/isoamylalcohol, once with diethyl ether (residual ether was blown away with nitrogen gas), and precipitated at -70°C after addition of 0.25 volume 10 M NH_4OAc and 3 volumes of 95% ethanol.

Polyacrylamide Gel Electrophoresis.

Next two sections (1 and 2) give examples for the preparation of 8% polyacrylamide gels. In order to prepare gels with lower polyacrylamide concentrations, the same procedures were followed, except that a lesser amount of acrylamide stock solution was used and the difference in volume was balanced by addition of dH_2O. The gel electrophoresis techniques, discussed below, were previously described by Maxam and Gilbert (1980).

1. Neutral gels.

The gels with 8% polyacrylamide content were prepared by mixing together 10 ml of 40% acrylamide stock solution [38 gm acrylamide and 2

gm of N,N'-diamine-bis-acrylamide (bisacrylamide) per 100 ml of dH_2O)], 5 ml of 10 x TBE (0.89 M Tris base, 0.89 M boric acid, 25 mM EDTA) (pH 8.0), 1.6 ml of 1.6% ammonium persulfate, and 33.4 ml of 37.5% glycerol solution. Polymerization was catalyzed by addition of 50 µl of N,N,N',N'-Tetramethyl-ethylenediamine (TEMED). Gels were formed in 1.5 mm thick, 14.5 cm wide, 10 cm long glass plates, and run at 200 V (constant voltage) with 1 x TBE as running buffer. The DNA samples were loaded with 10% glycerol, 1 x TBE, 0.042% BpB, and 0.042% XCFF.

2. Denaturing polyacrylamide/7 M urea gels.

These gels were prepared as follows: 30 gm of urea was dissolved in 12 ml of acrylamide stock solution (above) and 16.9 ml of dH_2O; the dissolved material was stirred with 0.5 gm of AG501-X8(D) resin for 10 min and filtered through a Nalgene 0.45 µm filter; to 43.4 ml of this filtrate, 5 ml of 10 x TBE, 1.6 ml of 1.6% ammonium persulfate, and 50 µl of TEMED, were then added; subsequently, the gels were poured and allowed to polymerize as described above. The gels were run at approximately 300 V (constant voltage), so that the surface temperature of the plates remained between 50 and 55°C. After electrophoresis the gels were rinsed in dH_2O, dried, and autoradiographed by exposure to X-Omat™ AR film.

3. SDS/polyacrylamide gels.

For the electrophoresis of proteins, gels were prepared in 0.3 mm thick, 14.5 cm wide, and 10 cm long glass plates essentially as described by Laemmli (1970). The lower gel consisted of following components: 12% polyacrylamide (37.5 acrylamide/1 bisacrylamide), 0.375 M Tris-HCl (pH 8.8), 0.1% SDS, 0.1% ammonium persulfate, 0.1% TEMED. Following the polymerization of the lower gel, the top of the plates was

sprayed with 0.1% SDS and upper gel was poured. This gel contained 3% polyacrylamide, 0.125 M Tris-HCl (pH 6.8), 0.1% SDS, 0.1% ammonium persulfate, and 0.1% TEMED. The gel was run at 13 mA (constant current) until the BpB dye went through the upper gel, and then the current was increased to 21 mA. The running buffer was 50 mM Tris-HCl (pH 8.3)/0.384 M glycine/0.1% SDS solution. All samples were loaded after heating at 100°C for 5 min in 62.5 mM Tris-HCl (pH 6.8), 10% glycerol, 2% SDS, 5% 2-mercaptoethanol, and 0.05% BpB solution.

S1 Nuclease and Primer Extension Analyses.

For S1 nuclease analysis (Berk and Sharp, 1977), radiolabeled probe [600-2000 counts per second (cps)] and the amount of RNA indicated in figure legends plus 100 µg of carrier *E. coli* tRNA were resuspended in 30 µl of hybridization buffer [80% deionized formamide; 0.4 M NaCl; 40 mM piperazine-N,N'-bis(2-ethane-sulfonic) acid (PIPES), pH 6.8; 0.5 mM EDTA), denatured for 10 minutes at 85°C and hybridized at 52°C for 3 hours. The samples were then diluted with 320 µl of ice cold, S1 buffer (50 mM NaOAc, pH 4.6; 0.28 M NaCl; 4.5 mM $ZnCl_2$, pH 5.0; 20 µg/ml denatured salmon sperm DNA). Digestion was carried out for 2 hours, at 20°C with 12,000 U of S1 nuclease. At the end of this reaction, S1 nuclease was deactivated by adding 6 µl of 2 M Tris-base and 8 µl of 0.5 M EDTA, in order to raise the pH and complex the metal ions, respectively. After adding 0.25 volume of 10 M NH_4OAc and 2 volumes of ethanol, the nucleic acids were precipitated at -20°C overnight. Following the centrifugation for 5 min at 10,000g samples were vacuum dried and taken up in 5 to 10 µl of 2 mM Tris-HCl (pH 7.5)/0.05 mM EDTA buffer. Before loading onto the 7 M urea/polyacrylamide gel, samples were mixed with 2 volumes of loading

buffer (88% deionized formamide, 10 mM EDTA, 0.2% BpB and XCFF) denatured at 90°C for 3 min and quick chilled on ice.

Hybridization for primer extension analysis (Broom and Gilbert, 1985) was carried out as described in the preceding paragraph, except the time of hybridization was decreased to 2 hrs, carrier tRNA was omitted, and only 500 to 1000 cps of radiolabled probe were used in each reaction. The total amount of RNA analyzed is also indicated in the figure legends. After hybridization, samples were diluted with 80 µl of cold 0.625 M KOAc solution and precipitated with 3 volumes of 95% ethanol at -70°C for 30 min. Following centrifugation at 10,000g (in eppendorf microcentrifuge) and 4°C for 15 min, two washes with 75% and one wash with 95% ethanol, all samples were dried and taken up in 2 x reverse transcription buffer (0.2 M Tris, pH 8.3; 20 mM $MgCl_2$; 2 mM DTT; 0.28 M KCl). The deoxyribonucleotides, RNasin, and avian myeloblastosis virus (AMV) reverse transcriptase were then added to final concentrations of 1 mM, 1 U/µl and 1 U/µl, respectively, and the reaction mixtures were incubated at 42°C for one hour. Samples were then precipitated with 3 volumes of 95% ethanol and 0.1 volume of 5 M KOAc (pH 5.0) at -70°C for 30 min; centrifuged for 15 min at 10,000g and 4°C in eppendorf microcentrifuge; washed with 75% ethanol; vacuum dried; and resuspended in 250 µl of TE buffer. This solution was then extracted with phenol/chloroform/isoamylalcohol and again precipitated with 5 M KOAc (pH 5.0) and 95% ethanol as described above. After centrifuging, washing, and vacuum drying, the samples were resuspended and loaded onto an appropriate 7 M urea/polyacrylamide gel as described above for the S1 nuclease analysis.

Characterization of Viral Proteins.

1. Assays for HBV antigens.

The HBsAg and HBeAg were detected by solid-phase radioimmuno assay (RIA) and HBcAg by inhibition of binding of ^{125}I-radiolabeled anti-HBcAg to HBcAg-coated beads using kits from Abbott Laboratories. Media conditioned by exposure to cells were either tested directly (100-200 µl/assay) or diluted to give an S/N ratio between 5 and 10. A sample was considered positive with an S/N ratio > 2.1. Units were calculated as [(cpm-background)/200 µl test sample] x 5 x dilution factor x total sample volume. The cytoplasmic "e" and core antigens were detected by using 50 µl of cell lysates (total lysate volume was usually 400 µl) in a given assay. In order to prepare the cell lysates, cells were first sedimented by centrifugation at 4°C and 500g, washed several times in 1 x PBS and then sonicated three times at 20 watts, 10 sec each time.

2. Purification of HBsAg.

For purification of HBsAg, five day old medium was collected from confluent 4.10 cells and separated from cellular debris by centrifugation at 10,000g for 10 minutes. The medium was fractionated by isopycnic CsCl centrifugation (1.38 gm CsCl/4.5 ml medium) in a Ti45 rotor at 36,000 rpm for 60 hrs at 4°C. Fractions were collected and HBsAg positive fractions were pooled and dialyzed against TNE buffer (10 mM Tris-HCl, pH 7.5; 0.1 M NaCl; 1 mM EDTA). The dialyzed material was layered on top of a discontinuous 20 to 60% sucrose gradient and centrifuged in a SW27 rotor at 25,000 rpm for 18 hrs at 4°C. HBsAg positive fractions were found at the interface between 60% and 20% sucrose layers. The purified HBsAg was diluted in a SDS/polyacrylamide

gel loading buffer denatured for 5 min at 100°C, and electrophoresed on SDS/polyacrylamide (12%) gel. One part of the gel was stained with Coomassie blue and proteins in the remaining part of the gel were electrophoretically transferred to nitrocellulose membranes (3 hrs at 60 V in 25 mM Tris-HCl, pH 8.3; 0.192 M glycine; 20% methanol).

3. Staining of the electrophoresed proteins.

Coomassie blue staining was done by soaking the gel for 30 min in a solution containing 0.1% dye, 50% ethanol, and 10% HOAc. The protein bands were visualized after destaining in 5% ethanol/10% HOAc solution. The destaining was continued, with several changes of the above solution, until the parts of the gel which did not contain proteins became clear.

The immune staining of proteins on the nitrocellulose paper (Towbin et al., 1979) was performed as follows: the nitrocellulose membrane was washed and incubated for 30 min at room temperature in TBST buffer (10 mM Tris-HCl, pH 8.0; 150 mM NaCl; 0.05% Tween) with 1% BSA; the blocking solution was removed and the paper incubated for 1 hr at room temperature in rabbit anti-pre-S2 (Neurath et al., 1985) or anti-S (raised against a synthetic peptide containing amino acids 135-155 of the S region, generously provided by Dr. R. Neurath) anti-sera, preadsorbed with the excess of cell culture medium from confluent 3T3 cells; the primary antibody was removed and nitrocellulose paper washed 3 times in TBST, 5 min each time; the membrane was then incubated in TBST for 30 min at room temperature with alkaline phosphatase conjugated goat anti-rabbit IgG (Promega Biotech, Madison, WI), washed 3 times in TBST and dried on a filter paper. The bound alkaline phosphatase was detected using a system from Promega Biotech, Madison, WI.

Hybridization Conditions.

The blots were prehybridized at 45°C with denatured salmon sperm DNA (0.25 mg/ml) in 5 x PSSC (50 mM sodium phosphate, pH 7.0; 3% sodium pyrophosphate; 0.75 M NaCl; 75 mM sodium citrate), 50% deionized formamide, 0.1% Denhardt's solution (0.1% ficoll, 0.1% polyvinylpyrrolidine, and 0.1% BSA) for 2 hours or overnight, depending whether riboprobe or nick translated probe was used, respectively. For hybridization with RNA probes, the blots were transferred to a solution containing probe, 0.1 mg/ml denatured salmon sperm DNA, 0.5 mg/ml yeast RNA, 50% deionized formamide, 5 x PSSC and 0.02% Denhardt's reagent, and incubated for 16 hours at 60°C. For nick translated probe and the Southern blots, the hybridization solution additionally contained 10% dextran sulfate, and yeast RNA was omitted; incubation proceeded overnight at 45°C. Non-specifically bound probe was removed by extensive washing with solutions maintained at 55°-60°C and 60°-65°C for DNA:DNA/DNA:RNA and RNA:RNA hybridizations, respectively. The highest stringency wash was for 1 hour with 0.1 x PSSC and 1% SDS. After washing, blots were rinsed with dH_2O, air dried and exposed to X-Omat™ AR film.

Commercial Sources of Reagents and Enzymes.

Below is a list of manufacturers or commercial suppliers of enzymes and important chemicals used in this study. All common chemicals, which are not indicated below, were purchased either from Fisher Scientific, or Sigma.

Acrylamide: Bio-Rad Laboratories.

Agarose, SeaKem ME: FMC BioProducts.

Ammonium persulfate: Bio-Rad Laboratories.

Bisacrylamide: Bio-Rad Laboratories.

Bacterial alkaline phosphatase: Bethesda Research Laboratories.

Bovine serum albumin: Sigma.

Brom phenol blue: Alied Chemical.

Cell culture reagents and media: Gibco Laboratories.

Cesium chloride: Bethesda Research Laboratories.

Coomassie blue: Bio-Rad Laboratories.

Deoxyribonuclease I, from bovine pancreas: Worthington, Inc.

Deoxyribonuclease I, RQI^{TM}: Promega Biotech.

Deoxyribonucleotides: Sigma.

Dextran sulfate: Pharmacia.

DNA, from salmon sperm: Sigma.

DNA Polymerase I, from E. coli: Boehringer and Mannheim.

DNA Pol I "Klenow", from E. coli: Boehringer and Mannheim.

Formaldehyde: Fluka AG.

Formamide: Fluka AG.

Glycerol: Fluka AG

Guanidine hydrochloride: Heico, Inc.

Guanidinium thiocyanate: Fluka AG.

Mixed bed resin AG501-X8(D): Bio-Rad Laboratories

Nitrocellulose membranes, BA 85: Schleicher and Schuell, Inc.

Nuclease S1: Boehringer and Mannheim.

Oligo(dt)-cellulose (type T-3): Collaborative Research, Inc.

Polynucleotide kinase, T4: Bethesda Research Laboratories.

Proteinase K: EM Sciences.

pSP64 and pSP65 vectors: Promega Biotech.

Radiolabeled nucleotides: Amersham, Inc.

Restriction enzymes: New Engaland BioLabs, Inc.

Reverse transcriptase: Seikagaku America, Inc.

Ribonuclease A: Sigma.

Ribonuclease T1: Sigma.

Ribonucleotides: Sigma.

RNA, from Yeast: Sigma.

RNA Polymerase, SP6: Promega Biotech.

RNasin: Promega Biotech.

Sephadex G-50, medium: Pharmacia.

Sodium dodecyl sulfate: Bio-Rad Laboratories

Sucrose: Schwarz/Mann Biotech.

tRNA, from E. coli: Sigma.

Tween: Mann Research Laboratories, Inc.

Urea: Schwarz/Mann Biotech.

X-Omat™ AR film: Estman Kodak Company.

Xylene cyanol FF: Fluka AG.

RESULTS

Restriction Analysis of HBV Sequences in DNA from Cells of Clone 4.10.

The total number of copies of HBV DNA in 4.10 cells is approximately 80 as estimated by dot-blot hybridization (data not shown). Hybridization of the 4.10 cell DNA on Southern blots (figure 4) with ^{32}P-labeled nick-translated HBV DNA indicated that 1) all of the HBV sequences were integrated into high molecular weight DNA, since the size of undigested HBV sequences was larger than 25,000 base pairs and no smaller episomal forms were detected (lane a); and 2) the majority of HBV sequences were present in HindIII fragments large enough to accommodate intact HBV dimer sequences (lane e). Since there is only one HindIII site (in the pBR322 sequences) in pTHBV-1, preservation of the entire plasmid during integration should give HindIII fragments >10,300 bp in length, while preservation of the smallest fragment of pTHBV-1 including both the HindIII site and the complete tandem genome would give fragments >6,400 bp in length. The major HBV DNA-containing restriction fragments observed after HindIII digestion of 4.10 DNA were >9,000 bp long. One minor band of approximately 3,200 bp was detected indicating the presence of at least few partial copies of the HBV genome. However, more detailed restriction analysis of the 4.10 cell DNA provided evidence that the integrity of the head-to-tail dimer was preserved during integration and amplification of the bulk of the transfected HBV sequences. In summary, it was found that: 1) the majority of HBV sequences were present in full genome length; after the EcoRI digestion, >95% (determined by densitometric scanning of the autoradiograph) of the HBV fragments generated from 4.10 DNA were apparently unit genome (3182 bp) length (lane b), the same length as the

Figure 4. Restriction analysis of HBV DNA sequences in 4.10 cells. Undigested DNA (lane a). DNA digested with EcoRI (lane b); BglII (lane c); BamHI (lane d) and HindIII (lane e). Arrows indicate the mobility of molecular weight markers in the same gel. The positions of above restriction enzymes with respect to the head-to-tail HBV dimer and pBR322 are indicated in figure 2.


a b c d e
23130
9416
6682
4361
2322
2027
564

EcoRI flanked HBV sequences in the plasmid used for transfection (figure 2); 2) the junction between the head-to-tail HBV dimers present in the transfected plasmid had been preserved, since both the BglII and BamHI digests of 4.10 cell DNA revealed the presence of fragments which could only have been derived from a head-to-tail dimer (figure 4) - these were the 2329 bp BglII fragment (lane c) and the 766 bp BamHI fragment (lane d). Digestion with BglII and BamHI also yielded the appropriate internal HBV fragments (439-lane c, 1504 and 912-lane d).

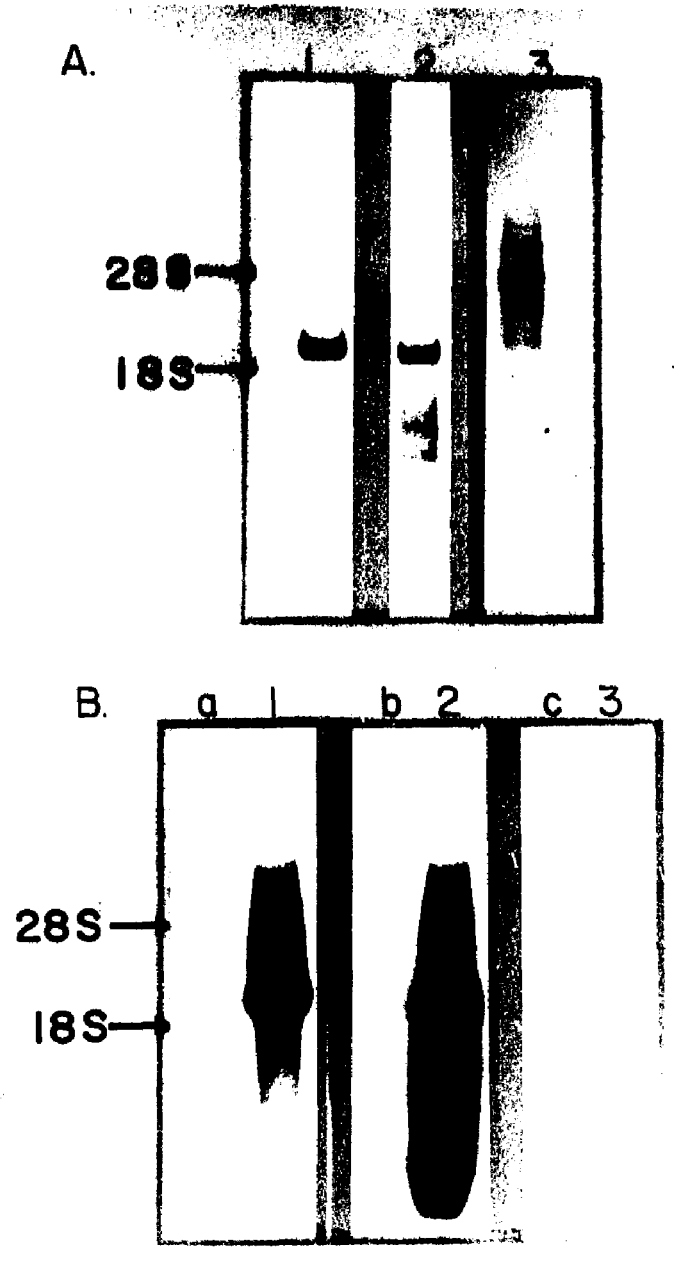
Characterization of Polyadenylated RNAs Transcribed from the Minus Strand of HBV DNA in 4.10 and D1113 cell lines.

Hybridization of poly(A)$^+$ RNAs isolated from 4.10 cells with ^{32}P-labeled nick-translated HBV DNA on Northern blots detected a variety of transcripts containing HBV sequences that ranged in size from approximately 900 to >5000 bases in length. The most prevalent transcript was approximately 2.1 kb long (data not shown). To determine which parts of the HBV genome were represented in these various transcripts, the restriction fragments indicated in fig. 2 were cloned into vectors that would direct the synthesis of radiolabeled sequence specific RNAs for use as probes (Methods). As can be seen in fig. 5A, probe 1 [bases 2839-(3182/1)-221], which should be specific for transcripts that include pre-S and amino-terminal S region codons, gave the strongest signal with an RNA of approximately 2.1 kb (lane 1), although several minor transcripts of longer-than-genome length were also detected. Probe 2 (bases 1686-1986) includes sequences which lie 5' to and include the only polyadenylation sequence in the viral (+)-strand (5'TATAAA3', bases 1918-1923). It also contains sequences from the unassigned reading frame, "X". Like the pre-S-S specific

probe 1, probe 2 annealed most strongly with the 2.1 kb poly(A)$^+$ RNA (lane 2). This result is compatible with the proposal that the 2.1 kb HBV transcript in 4.10 cells is analogous to that found in HBV-infected livers (Cattaneo et al., 1984; fig. 8, lane b) and in cells transfected with cloned HBV DNA (Laub et al., 1983; Gough, 1983; Pourcell et al., 1982), i.e., a transcript that starts in the pre-S region (see below for precise mapping of the 5' end of this transcript) and terminates shortly after the polyadenylation signal. Such a transcript would include both the S and "X" sequences. Probe 2 also detected three smaller RNA species (400-900 bp) that did not anneal with probe 1 and thus did not contain sequences from the pre-S and S regions. This suggests a possibility that the "X" specific mRNAs are synthesized in 4.10 cells. Probe 3 (bases 2169-2425) contains sequences from the coding region of the core gene. This fragment lies 3' to the polyadenylation sequence in the viral (+)-strand and should thus be specific for the core region transcripts that either pass through or start 3' to this signal. Probe 3 did not anneal with either the 2.1 kb transcript, or the small transcripts containing the "X" region sequences. Instead, it detected primarily the transcripts of longer-than-genome length (lane 3), with the most distinct species having lengths of approximately 3.5, 4.5 and 6.5 kb. Less discrete bands were also noted at molecular lengths ranging from 3 to 1.5 kb.

The lanes in fig. 5A are from autoradiographs obtained at different exposure times to optimize detection of the strongest signals from a particular probe. Figure 5B depicts the results obtained when all blots were exposed to autoradiography for the same period of time (24 hrs).

Figure 5. Northern analysis of poly(A)$^+$ RNAs isolated from 4.10 cells: A) Poly(A)$^+$ RNAs from 4.10 cells (2.5 µg per lane) annealed with probes specific for the pre-S-S (lane 1), "X" (lane 2), and core gene (lane 3) sequences. Exposure time for the autoradiography was 6 hrs for lanes 1 and 2, 48 hrs for lane 3. B) Poly(A)$^+$ RNAs from 4.10 (lanes 1,2,3) and 3T3 cells (lanes a,b,c) annealed with probes specific for pre-S-S (lanes a,1), "X" (lanes b,2), and core gene (lanes c,3) sequences. Exposure time for the autoradiography was 24 hrs for all lanes.


A.
1
2
3
28S
18S
B.
a 1
b 2
c 3
28S
18S

It is clear that the major transcript copied from the integrated HBV sequences is 2.1 kb in length and that the longer-than-genome length core region specific mRNAs represent only a minor fraction of total HBV RNA. It can also be seen from the autoradiographs that all of the probes utilized for these experiments are highly specific for HBV sequences since they did not anneal with poly(A)$^+$ RNA from 3T3 cells, the parental line for 4.10 (lanes a,b,c).

Northern blotting and hybridization of poly(A)$^+$ RNAs isolated from the D1113 cells with a nick translated HBV probe showed only two major RNA bands (fig. 6, lane a) which were about 2.1 and 3.5 kb in length. The same result was obtained when these RNAs were hybridized with the above described riboprobe 1 (fig. 7, lane b), which is specific for transcripts containing the S region of the HBV genome. When riboprobe 3 was used (specific only for C region sequences), the 3.5 kb RNA signal remained prominent, but the 2.1 kb band was not detected (fig. 6, lane c). These results demonstrate that, although the S gene sequences are present in both transcripts, the C gene sequences are only contained in the longer RNA. In addition, the relative abundance and molecular weight of these two RNAs are reminiscent of the 2.1 kb and 3.5 kb transcripts observed in analysis of the poly(A)$^+$ RNAs from HBV infected chimpanzee livers (fig. 8, lane b).

The S1 nuclease and primer extension methods were used to define the 5' ends of the above described transcripts. The 5' end-labeled XbaI-RsaI* fragment [nucleotides 2508-(3182/1)-249*, asterisk indicates the labeled nucleotide], spanning bases in the pre-S and early S regions, was hybridized to 4.10 and NIH 3T3 poly(A)$^+$ RNAs and the hybrids were subjected to S1 nuclease digestion (fig. 9A, lane 2 and 1,

Figure 6. Northern analysis of poly(A)$^+$ RNAs from D1113 cells hybridized with the nick translated HBV DNA, and riboprobe specific for transcripts containing the core region sequences only (see figure 2, probe 3). Lanes a and c correspond to 2.5 µg of D1113 poly(A)$^+$ RNA hybridized with the nick translated HBV DNA and the above described riboprobe, respectively. The negative control was 2.5 µg of poly(A)$^+$ RNA from NIH 3T3 cells hybridized with the nick translated probe, lane b; and riboprobe, lane d.

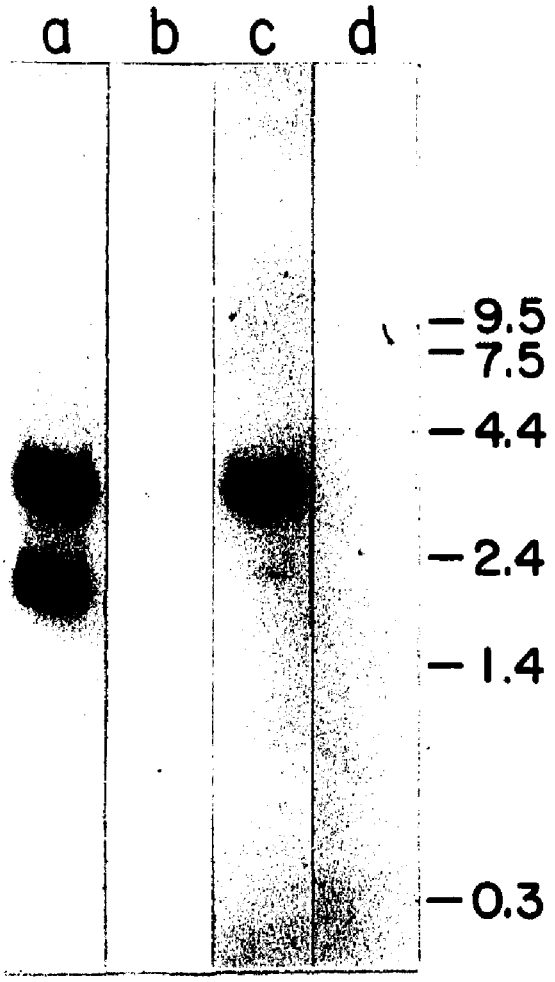

a
b
c
d
9.5
7.5
4.4
2.4
1.4
0.3

Figure 7. Northern analysis of poly(A)$^+$ RNAs isolated from D1113 cells with the riboprobe specific for pre-S-S sequences (see figure 2, probe 1). Lanes a and b show hybridization of this probe with 2.5 µg of poly(A)$^+$ RNA from NIH 3T3 and D1113 cells, respectively.

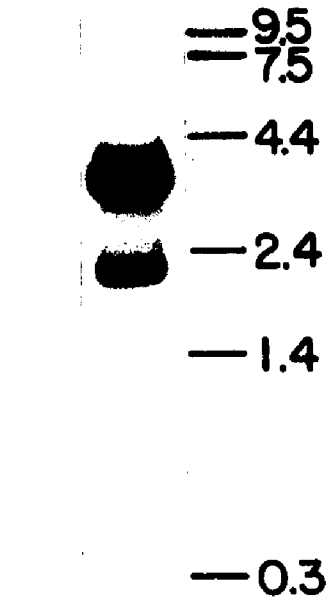
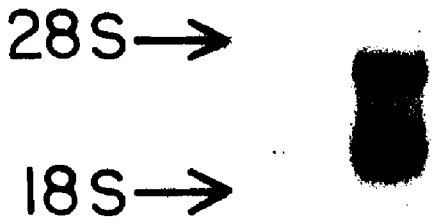

a
b
9.5
7.5
4.4
2.4
1.4
0.3

Figure 8. Northern analysis of poly(A)$^+$ RNAs from HBV infected chimpanzee livers. The HBV infected chimpanzee liver RNA was generously provided by Dr. J. Gerrin. The poly(A)$^+$ RNA was isolated as described in methods section. Lane b shows hybridization of 2.0 µg of poly(A)$^+$ RNA with the riboprobe specific for pre-S-S sequences (probe 1, see figure 2). Lane a represents a negative control showing hybridization of 20 µg of total human fetal liver RNA with the same riboprobe.

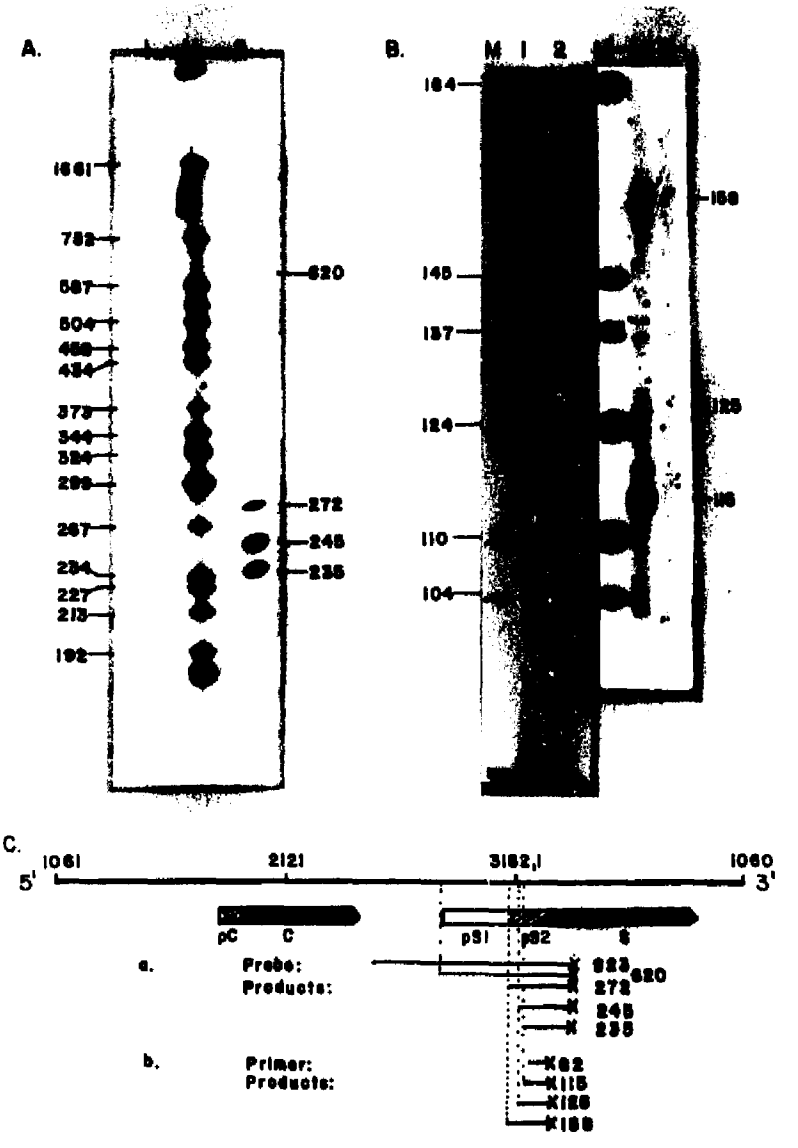
a b



respectively). Four protected digestion products were observed when 4.10 RNA was used in the reaction. The three most prominent fragments were 272±5 nucleotides (nt), 245±6 nt and 235±6 nt long, indicating transcript discontinuities (5' ends or splice junctions) at map positions 3159±5, 4±6 and 14±6, respectively. The fourth fragment, 620±15 nt, indicated the presence of a small percentage of transcripts with discontinuities mapping to base 2811±15. Hybridization of the same probe with poly(A)$^+$ RNA from D1113 cells and subsequent digestion with S1 nuclease gave rise to three protected fragments (fig. 11A, lane 4). The most prominent fragment was 272±4 nt long and indicated transcription discontinuity at nucleotide position 3159±4. Two additional transcription discontinuities were mapped to nucleotides 22±6 and 4±6, as indicated by the protected fragment lengths of 227±6 and 245±6 bases, respectively; however, the amount of the latter species was less than the former. In addition, some probe was completely protected, presumably by longer-than-genome length transcripts; and the discontinuity in proximity of base 2811 was not observed.

To determine if the above discontinuities were authentic transcript ends, or represented splice sites or junctions between HBV sequences and non HBV sequences in transcripts originating outside of the integrated HBV genomes, a 5' radiolabeled XhoI*-StuI fragment (nucleotides 47-129*) was annealed to RNA from 4.10 and 3T3 cells (fig. 9B). Extension of this primer on the 4.10 cell RNA (lane 1) yielded two sets of cDNA products with lengths of 115±3 and 158±3 nt, confirming the presence of primary initiation sites for S region transcripts at approximate positions 3153±3 (approximately 20 bases before the AUG codon marking

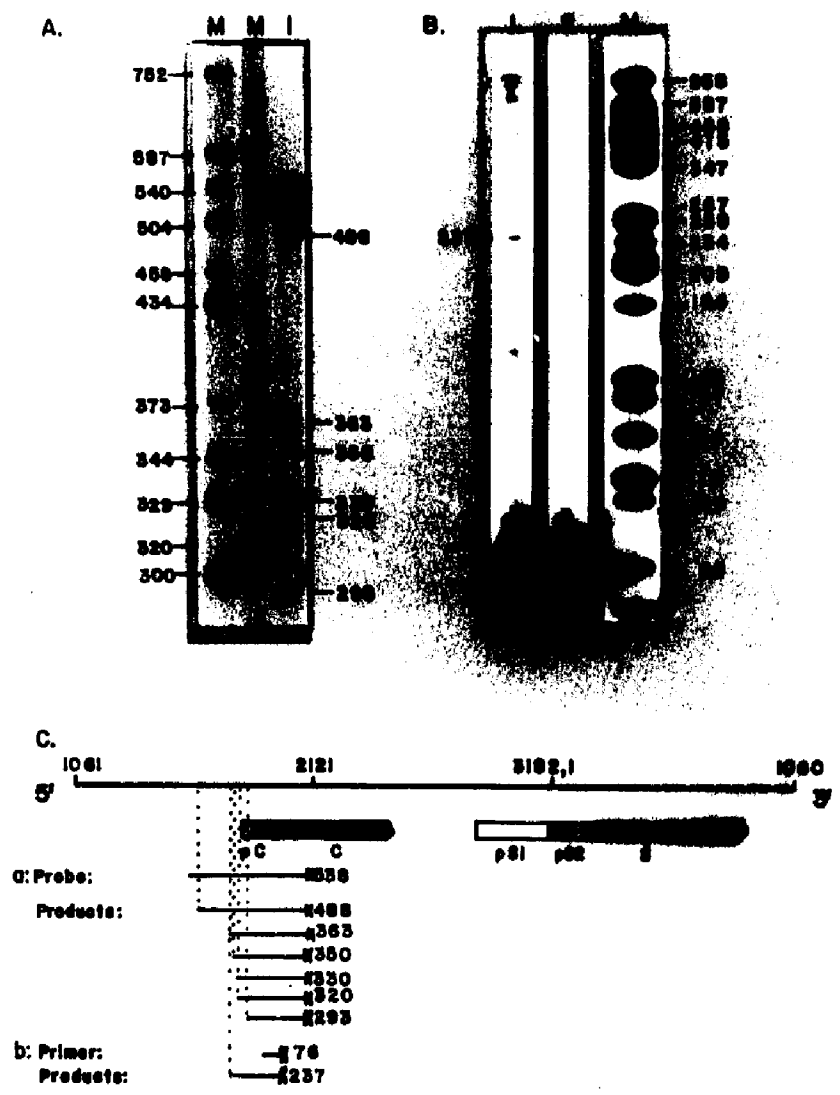
Figure 9. Characterization of the 5' ends of the 2.1 kb RNA isolated from 4.10 cells: A) S1 nuclease mapping of 5' end using a probe corresponding to bases 249-(1/3182)-2508, labeled on the 5' end at nucleotide 249, and 5 µg of poly(A)$^+$ RNA isolated from either 3T3 (lane 1) or 4.10 (lane 2) cells. Protected fragments were electrophoresed on a 6% urea/polyacrylamide gel. B) Primer extension analysis using a primer of bases 129-47, labeled on the 5' end at nucleotide 129 and 10 µg of poly(A)$^+$ RNA isolated from either 4.10 (lane 1) or 3T3 (lane 2) cells. Products were analyzed on an 8% urea/polyacrylamide gel. Size markers are shown in lanes designated M. Lanes M^1, 1^1, 2^1 show the effects of a 10 times increase in film exposure time to the same gel. C) Schematic localization of the start sites for the HBsAg mRNA as mapped by: a, S1 nuclease; b, primer extension. Sizes of the probes and protected or extended fragments are indicated and the labeled 5' end is marked with an asterisk.


A.
1661
783
587
504
458
434
373
344
324
298
267
234
227
213
192
620
272
248
235
B.
M 1 2
184
145
137
124
110
104
158
125
116
C.
5'
1061
2121
3182,1
1060
3'
pC
C
pS1
pS2
S
a.
Probe:
Products:
823
620
272
248
235
b.
Primer:
Products:
K82
K115
K126
K158

the start of the pre-S2 region) and 14±3 (within the pre-S2 coding region). Less abundant products with lengths in the range of 125±3, indicated that transcripts were also initiated around base 4±3. Termination of primer extension was not observed in the region of base 2811. Attempts to confirm the initiation site at base 2811 by extension of a 5' radiolabeled BstNI-BamHI fragment (nucleotides 2967*-2906) also failed. Thus, if the site at 2811 which lies approximately 25 bases 3' to a promoter-like TATATAA sequence is actually used to initiate transcription, such an initiation must occur with a very low frequency.

Similarly, in clone D1113, primer extension analysis also confirmed the three initiation sites for 2.1 kb RNA, which were mapped with S1 nuclease. The cDNA products were 162±5 nt, 123±4 nt and 115±4 nt long, indicating transcriptional start sites at positions 3149±5, 6±4 and 14±4 respectively (fig. 11B, lane 4). When the poly(A)$^+$ RNA from NIH 3T3 cells was subjected to the same analysis, above described extension products were not observed (fig. 11B, lane 3). However, a very strong artifactual band, corresponding to cDNA of about 137 bases, was detected. Since this band did not occur when reverse transcriptase was not included in the elongation reaction, this artifact was probably generated by extension of a small amount of probe which annealed upon itself. Examination of HBV DNA sequence showed that six bases at the 3' end of the probe (3'GGACAT) could base pair with a complementary sequence (3'AAGTCC), located in the same strand at nucleotide number 74, thus providing a 3' end for elongation by reverse transcriptase. The cDNA product expected from this synthesis would be 132 bases, approximately the size of the artifactual band.

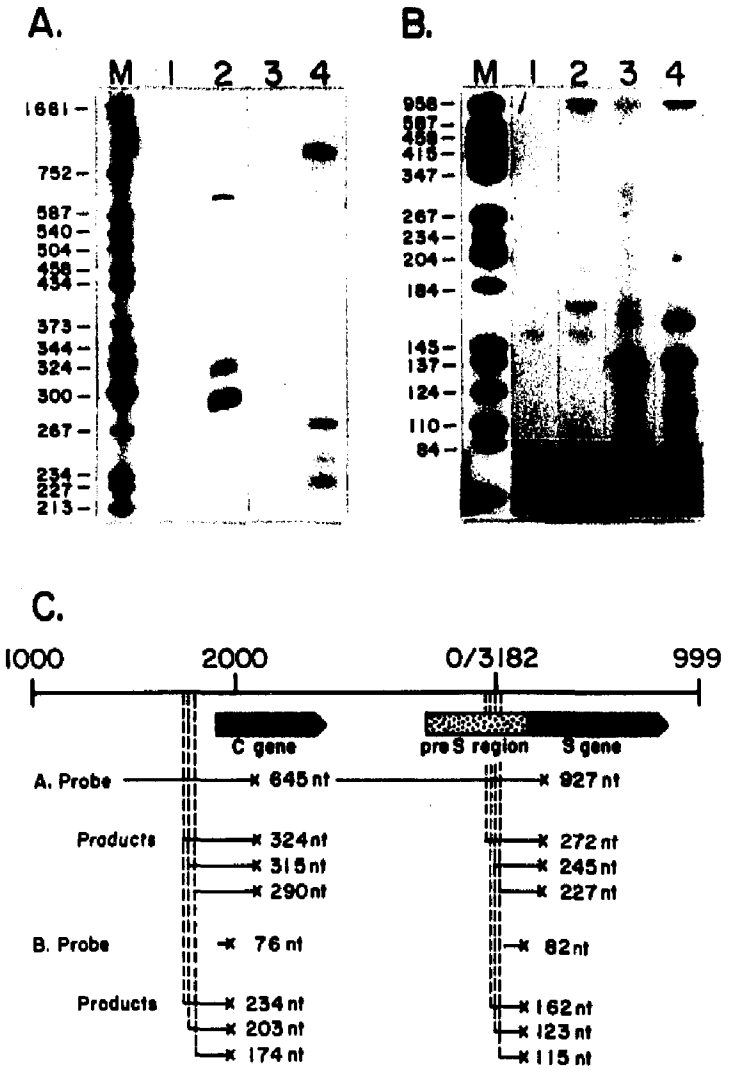
Figure 10. Characterization of the 3.5 kb (longer-than-genome length) HBV transcript isolated from 4.10 cells: A) S1 nuclease mapping of 5' end using probe corresponding to bases 2110 to 1572, labeled on the 5' end at nucleotide 2110, and 10 µg of poly(A)$^+$ RNA isolated from 4.10 cells (lane 1). The protected fragments were electrophoresed on a 6% urea/polyacrylamide gel. B) Primer extension analysis using a primer of bases 1986 to 1910, labeled on the 5' end at nucleotide 1986, and 20 µg of poly(A)$^+$ RNA isolated from 4.10 (lane 1) or NIH 3T3 (lane 2) cells. Extension products were separated on 8% urea/polyacrylamide gel. Size markers are shown in lanes designated M. C) Schematic localization of the start sites for the longer-than-genome length RNA as mapped by: a, S1 nuclease; and b, primer extension analysis. The probe sizes and protected or extended fragments are indicated.


A.
M M I
782
587
540
504
458
434
373
344
320
300
B.
C.
1061
2121
3182,1
1980
5'
3'
pC
C
p81
p82
a: Probe:
Products:
b: Primer:
Products:

Transcripts from which the core protein can be translated were identified by hybridization of a 5'-labeled HpaII-BstNI* fragment (nucleotides 1572-2110*, spanning bases from mid-core coding region through most of the "X" coding region) with 4.10 RNA, and digestion with S1 nuclease. Analysis of protected fragments revealed sites at 1747±6 and 1622±10 (fragment lengths 363±6 and 488± nt, respectively; fig. 10A, lane 1). Minor components with lengths of approximately 350±6, 330±6, 320±6 and 293±6 nt were also visualized, indicating discontinuities at nucleotides 1760±6, 1780±6, 1790±6 and 1817±6. Whether any or all of these discontinuities represent true initiation sites remains to be conclusively shown since the primer extension (fig. 10B, lane 1) only detected the site at nucleotide 1749±5. It may be presumed that the failure to confirm the other sites was the result either of the low level of these transcripts, or the possibility that some of these RNAs originated outside of the HBV genome and traverse the junction between the integrated virus and host sequences. However, the site at 1817±6, which lies in the pre-core region, would be analogous to one of the major initiation sites for transcription of the longer-than-genome length transcript of GSHV synthesized in livers of infected animals (Seeger et al., 1986), while the site at 1749±5 lies within 100 bp of a promoter-like sequence in HBV DNA which has been shown to function *in vitro* with RNA polymerase II (Rall et al., 1983).

The S1 nuclease analysis of the poly(A)$^+$ 3.5 kb RNA from D1113 cells, performed with a probe corresponding to the AvaI-BstNI fragment (1465-2110*), resulted in protected fragments of 324±7 nt, 315±7 nt and 290±7 nt, designating transcript discontinuities at nucleotide

Figure 11. S1 nuclease and primer extension analyses of the two major HBV RNAs (2.1 kb and 3.5 kb) isolated from D1113 cells. A) S1 nuclease analysis using either BstNI-AvaI HBV fragment (bases 2110 to 1465, labeled on the 5' end at nucleotide 2110) hybridized with 5 µg of poly(A)$^+$RNA from NIH 3T3 (lane 1) and D1113 (lane 2) cells, or XbaI-RsaI fragment [bases 249-(1/3182)-2508, labeled on the 5' end at nucleotide 249] hybridized with 5 µg of poly(A)$^+$RNA from NIH 3T3 (lane 3) and D1113 (lane 4) cells. The protected fragments were electrophoresed on 6% urea/polyacrylamide gel. B) Primer extension analysis using either BglII-TaqI primer (bases 1986 to 1910, labeled on the 5' end at nucleotide 1986) hybridized with 10 µg of poly(A)$^+$RNA from NIH 3T3 (lane 1) and D1113 (lane 2) cells, or XhoI-StuI primer (bases 129 to 47, labeled on the 5' end at nucleotide 129) hybridized with 10 µg of poly(A)$^+$RNA from NIH 3T3 (lane 3) and D1113 (lane 4) cells. Extended fragments were electrophoresed on 8% urea/polyacrylamide gel. Molecular size markers are in lanes designated M. C) Schematic localization of start sites for the two major HBV RNAs as mapped by: A, S1 nuclease; B, primer extension. Sizes of probes and protected or extended fragments are indicated and the labeled 5' end is indicated with an asterisk.


A.
M 1 2 3 4
1661
752
587
540
504
458
434
373
344
324
300
267
234
227
213
B.
M 1 2 3 4
958
587
458
415
347
267
234
204
184
145
137
124
110
84
C.
1000
2000
0/3182
999
C gene
pre S region
S gene
A. Probe
645 nt
927 nt
Products
324 nt
315 nt
290 nt
272 nt
245 nt
227 nt
B. Probe
76 nt
82 nt
Products
234 nt
203 nt
174 nt
162 nt
123 nt
115 nt

positions 1786±7, 1795±7 and 1820±7, respectively (fig. 11A, lane 2). The most intense signal indicated a possible initiation at nucleotide 1820±7, which lies within the pre-C region of the HBV genome. The other two signals were of minor, but equal intensities. Primer extension analysis, carried out with a primer corresponding to nucleotides 1910-1986* (TaqI-BglII* fragment) showed the above indicated transcription discontinuities to be the true 5' ends of the longer-than-genome length RNA (fig. 11B, lane 2). The most frequent initiation site was mapped to the nucleotide 1812±4 (cDNA size 174±4 nt), confirming the start site which was located by the S1 nuclease analysis to around base 1820. The other transcription initiation sites were mapped to regions 5' to the first pre-C AUG at nucleotides 1783±5 and 1752±5 (cDNA sizes 203±5 nt and 238±5 nt, respectively).

Although the 5' ends mapped with the primer extension analysis do not agree precisely with those mapped by S1 nuclease technique, these differences are within the limits of an error associated with estimation of molecular sizes by this type of analysis. Furthermore, since the S1 nuclease has a tendency to leave one to five nucleotides overhanging ends (Grosschedl and Birnsteil, 1980) and to end-nibble (Shenk et al., 1975), the primer extension analysis probably represents a more accurate mapping of the mRNAs' 5' ends. Also, the most accurate readings can be obtained in the lower molecular weight region of a given gel and this accuracy markedly decreases toward the top. In addition, the smiling effect of the gel usually adds in obscuring the results, especially in the lower gel regions. In order to facilitate an easier analysis of the results, all major HBV transcription start sites, which were mapped in 4.10 and D1113 cells, are summarized in Table I.

Characterization of Polyadenylated RNAs Transcribed from the Plus Strand of HBV DNA in 4.10 cells.

Northern transfer and hybridization of poly(A)$^+$ RNAs from 4.10 cells with radiolabeled strand specific RNA probes, capable of detecting only transcripts from the (+)-strand (S-strand) of HBV, revealed the presence of three RNA species: one major band which was approximately 2.4 kb long, and two minor bands of 2.0 and 2.8 kb in length (fig. 12, lane b). These bands were not detected in the poly(A)$^+$ RNA from untransfected mouse 3T3 cells, although three minor bands with different mobilities were visualized suggesting a low level of non-specific interaction with the probe (data not shown).

Since transcripts originating from the (+)-strand of HBV in cells containing integrated viral DNA could result from initiation of transcription in non-HBV sequences, the transcripts were further characterized by mapping their 5' and 3' ends. The S1 nuclease protection analysis with a probe corresponding to bases 1402*-2183 (BamHI*-HaeIII fragment) yielded DNA bands indicative of transcript discontinuities at bases 1917$\pm$10, 1892$\pm$10 and 1864$\pm$10 (fig. 13, lane a). The primer extension analysis carried out with a probe corresponding to bases 1726*-1804 (DraI*-FspI fragment) indicated an initiation site of RNA at base 1861$\pm$5 (fig. 13, lane b), confirming one of the sites found by S1 nuclease analysis. In addition, fifty bases upstream from this site (bases 1917 to 1921), is a 'TATA'-like sequence which could serve as a promoter.

Analysis of the 3' end of the (+)-strand transcript with probe corresponding to bases 2143-2906* (XbaI-BamHI* fragment, see fig. 14) resulted in a protected fragment corresponding to a termination site at

Figure 12. Northern analysis of poly(A)$^+$ RNA transcripts from the (+)-strand of HBV DNA isolated from 4.10 cells (2.5 µg of RNA per lane): a, 4.10 cells probed for (+)-polarity RNA with riboprobe containing bases 2839 to 221; b, 4.10 cells probed for (-)-polarity transcripts (transcripts from the plus-strand of HBV) with riboprobes containing bases 2839 to 221 and 1686 to 1986. See figure 2 or 14 for derivation of the fragments. Exposure time for this autoradiography was 24 hours.

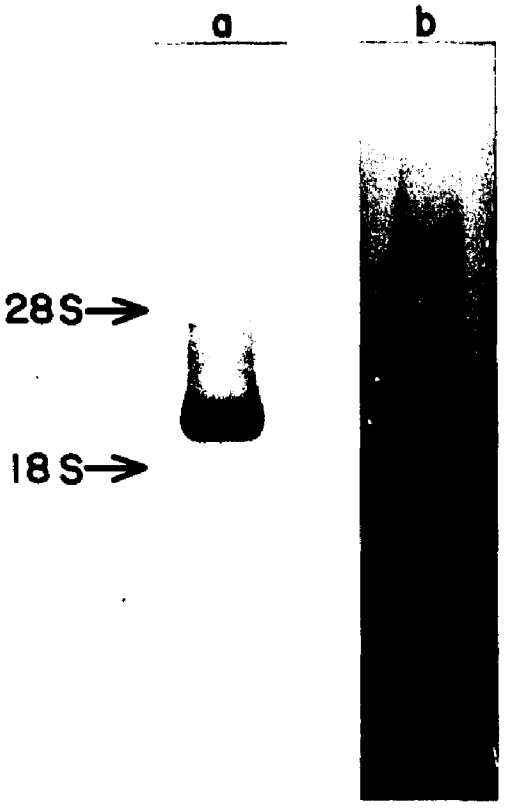

a
b
28S
18S

TABLE I. Summary of major initiation sites for HBV specific transcripts found in 4.10 and D1113 cells.

Cell line	RNA species	S1 mapping	Primer extension
4.10	2.1 kb	3159±5	3153±3
4.10	2.1 kb	4±5	4±3
4.10	2.1 kb	14±5	14±3
4.10	3.5 kb	1747±6	1749±5
D1113	2.1 kb	3159±4	3149±5
D1113	2.1 kb	4±6	6±4
D1113	2.1 kb	22±5	14±4
D1113	3.5 kb	1820±7	1812±4
D1113	3.5 kb	1795±7	1783±5
D1113	3.5 kb	1786±7	1752±5
4.10	antisense	1864±10	1861±5

Only the start sites detected by both the S1 nuclease and primer extension methods of analysis are indicated in the above table.

base 2381±10 (fig. 13, lane c). A poly(A)$^+$ HBV transcript starting at around base 1861 and ending in the proximity of base 2381 would be approximately 2.8 kb long, the size of the largest transcript detected by Northern blot analysis (fig. 12). No known polyadenylation signals are located close to the termination site at base 2381. However, about 100 bases upstream is the only "consensus" polyadenylation signal (AATAAA) in the HBV genome, a sequence conserved in the genomes of three other HBV subtypes (Galibert et al., 1979; Valenzuela et al., 1980; Ono et al., 1983) and in other hepadna viruses (Galibert et al., 1982;

Figure 13. Characterization of RNA Transcripts from the (+)-strand of HBV DNA isolated from 4.10 cells: A) S1 nuclease and primer extension analysis of antisense RNA: M, markers; a, S1 analysis of 5' end using a probe corresponding to bases 1402 to 2183, labeled at the 5' end of the (+)-strand at base 1402, and 5 µg of poly(A)$^+$ RNA; b, primer extension using a primer of bases 1726 to 1804, labeled at the 5' end of the (+)-strand at base 1726, and 15 µg poly(A)$^+$ RNA; c, S1 nuclease analysis of 3' end using probe of bases 2143 to 2906, labeled at the 3' end of the (+)-strand at base 2906, and 5 µg poly(A)$^+$ RNA. No extended or protected fragments were observed when poly(A)$^+$ RNA from NIH 3T3 cells were analyzed under identical conditions. Products of the S1 nuclease and primer extension analyses were electrophoresed on 4% and 8% urea/polyacrylamide gels, respectively. B) Schematic localization of the transcriptional signals for the antisense RNA on the HBV genome. The labeled 5' or 3' end is indicated with an asterisk.

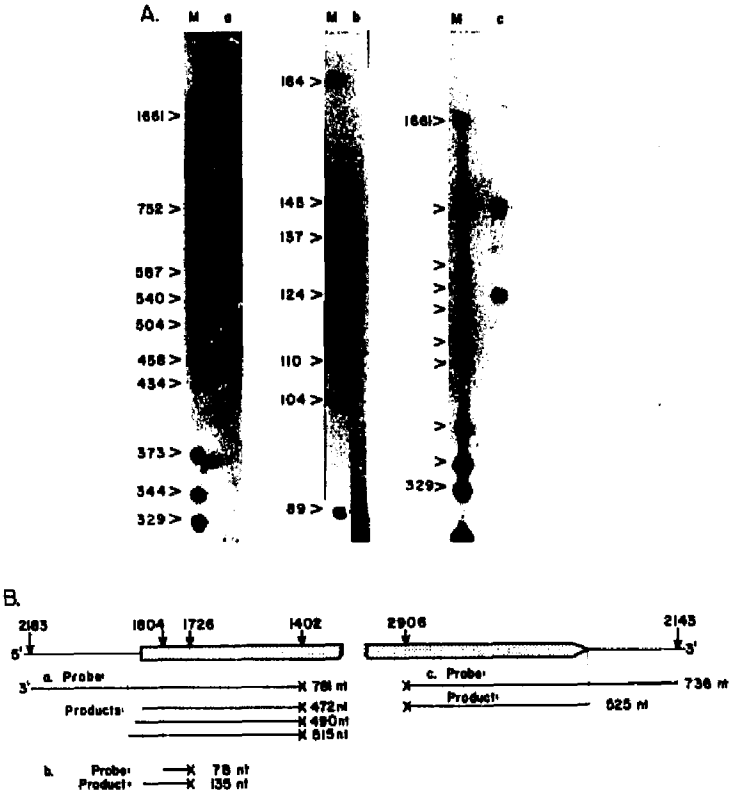
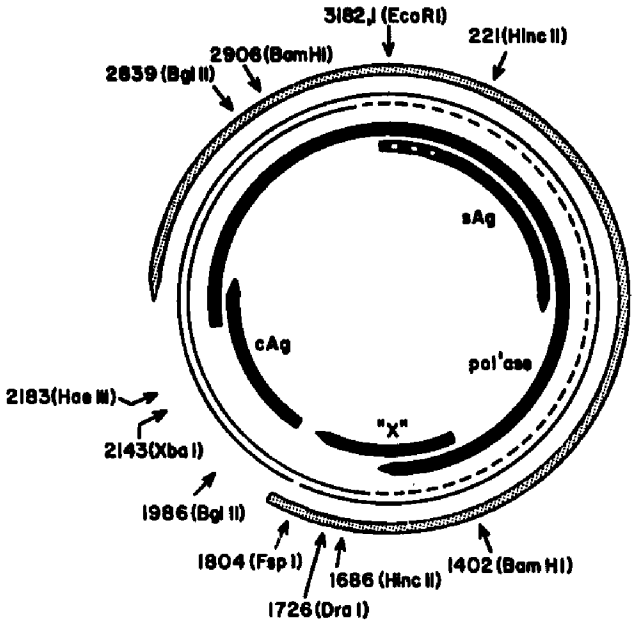

A.
M a
M b
M c
1661 >
752 >
587 >
540 >
504 >
458 >
434 >
373 >
344 >
329 >
184 >
148 >
137 >
124 >
110 >
104 >
89 >
1661>
329>
B.
2183
1804
1726
1402
2906
2145
5'
3'
a. Probe
3'
781 nt
Products
472 nt
490 nt
515 nt
c. Probe
736 nt
Product
525 nt
b.
Probe
78 nt
Product
135 nt

Figure 14. Position of the S-strand (antisense) transcript (shaded arrow) with respect to HBV genome, as mapped with S1 nuclease and primer extension analyses. Thin lines represent the two DNA strands, long and short; solid arrows are the potential coding regions of the L-strand transcript. Restriction enzyme sites are those which were used to prepare probes specific for the S-strand transcripts.


3182,1 (EcoRI)
221(Hinc II)
2906(BamHI)
2839(Bgl II)
sAg
cAg
pol'ase
"X"
2183(Hae III)
2143(Xba I)
1986(Bgl II)
1804(Fsp I)
1686(Hinc II)
1726(Dra I)
1402(BamHI)

Mandart et al., 1984; Seeger et al., 1984).

Characterization of Hepatitis B Virus Proteins Synthesized in 4.10 and D1113 Cells.

The clone 4.10 cells secrete relatively large amounts of HBsAg (Table II). By comparing dilutions of media exposed to 4.10 cells with dilutions from a standard preparation of HBsAg, it could be calculated that at confluence, 4.10 cells released 80-100 μg of HBsAg per 10^7 cells per day, or approximately 10 times the amount found in our previously described clones (Christman et al., 1982). However, although 22 nm spherical HBsAg particles were readily detected in all antigen positive samples, immunoelectronmicroscopic examination of both intracellular and released HBsAg failed to reveal the presence of 42 nm viral particles (Dane particles) or 22 nm tubular forms of HBsAg (data not shown).

In proportion with their higher level of HBsAg production, the 4.10 cells also released approximately 10 fold more HBeAg than those other clones. Average S/N ratios for medium conditioned for 5 days by 4.10 cells was ≥ 4 while intracellular HBeAg could only be reliably detected in 4.10 cell lysates if the concentration of cells was $> 10^7$/ml. Thus, the majority of HBeAg produced by 4.10 cells was released to the culture medium (Table II). In comparison, the D1113 cells also synthesized large amounts of HBsAg, and HBeAg [similarly present to a greater extent in the cell culture medium than in cellular cytoplasm (Table II)]. However, while the amount of HBsAg detected in the media from these cells was approximately 100 times lower than in 4.10 cell cultures, the levels of cytoplasmic and secreted HBeAg were significantly higher in D1113 cells. In addition, Table III shows

TABLE II. Detection of HBsAg and HBeAg in 4.10 and D1113 cell cultures.

Antigen	Culture time (d)	Units (cpm) per 10^6 cells
HBsAg	4.10 cell culture medium	7.5×10^6
HBeAg	4.10 cell culture medium	1.5×10^4
HBeAg	4.10 cell lysates	0.4×10^3
HBsAg	D1113 cell culture medium	12.6×10^4
HBeAg	D1113 cell culture medium	5.5×10^4
HBeAg	D1113 cell lysates	1.2×10^4

Aliquots of cell culture media and cell lysates (from $\geq 3 \times 10^7$ cells) were prepared so that an S/N ratios between 5 and 10 were obtained in the Abbott RIA kits for HBsAg or HBeAg. Units were calculated as [(cpm - negative control)/200 µl test sample] x 5 x dilution factor x total sample volume. The data represent the average of duplicate determinations on samples from 5 different culture flasks. Negative controls from the kits or equivalent dilutions of 3T3 cell medium were less than 400 cpm. All cultures were approximately 5 days old at the time of the assay and were fully confluent.

TABLE III. Detection of HBcAg in confluent cultures of D1113 cells.

Source	% Inhibition	Number of Cells
Medium	47.6	1.9×10^7
Cell lysate	70.4	2.7×10^7

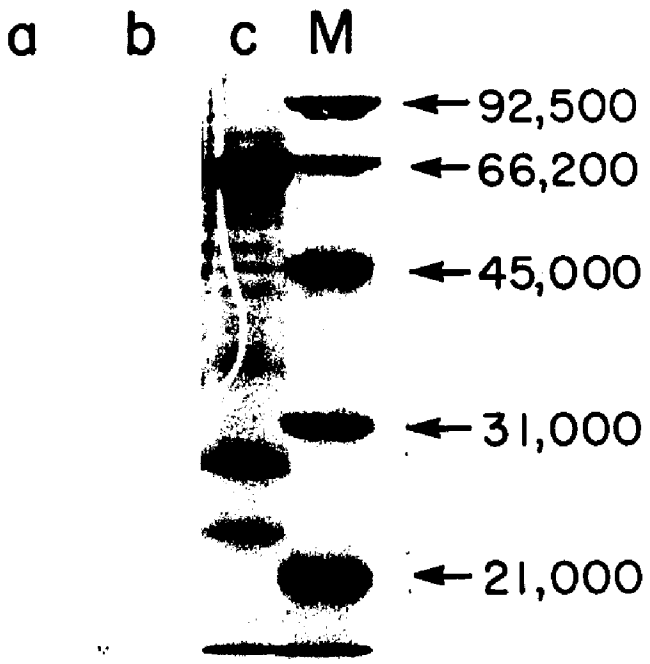
Percent inhibition of ^{125}I-labeled anti-HBcAg binding to immobilized HBcAg in the CORAB RIA assay. Either 50 µl of cell lysate (400 µl total volume), or 100 to 200 µl of cell culture medium were used in above determinations. Cell lysate or medium from control NIH 3T3 cells showed no significant inhibition.

that D1113 cells produced significant amounts of HBcAg which was found in their cytoplasms and culture media. In contrast, the HBcAg could not be reliably detected in 4.10 cell culture medium even after 25 fold concentration. The cell lysates from approximately 3 x 10^7 cells inhibited binding of ^{125}I-labeled anti-HBcAg to immobilized HBcAg [CORAB assay] by an average of only 20% and the immunoblotting of proteins from 4.10 cell lysates failed to reveal any anti-HBcAg reactive species (data not shown).

For further characterization, the secreted HBsAg was partially purified from contaminating serum proteins by isopycnic centrifugation in CsCl. The HBsAg particles, which banded at a density of 1.2 g/ml, were dissociated by boiling with SDS in the presence of 2-mercaptoethanol and the denatured proteins were separated by polyacrylamide gel electrophoresis. Coomassie blue staining of the gel (fig. 15, lane c) revealed multiple protein species with strongly stained bands at 24, 27 and 66 kilodaltons (kd), corresponding to the major HBsAg protein (P24), its glycosylated form (GP27) and the contaminating BSA, respectively.

Western blotting and immune staining of the proteins demonstrated that the 24 kd species reacted with antibody specific for an epitope lying within the major envelope protein domain [anti-S(135-155)]. This antibody also bound to epitopes in the 33 and 36 kd species as well as to a few minor species of > 45kd. GP27 reacted only weakly with this antibody (fig. 15, lane a). Since they bind anti-S and anti-pre-S2 antibodies, the 33 and 36 kd species appear to be HBV middle envelope proteins, GP33 and GP36, which are encoded by the pre-S2 and S genes. As would be expected, neither P24 nor GP27 reacted with the latter

Figure 15. Analysis of HBV envelope proteins from 4.10 cell culture medium: a, Western blotting using rabbit anti-S antibody; and, b, rabbit anti-pre-S2 antibody reacted with alkaline phosphatase conjugated goat anti-rabbit IgG; c, same sample as in a and b stained with Coomassie blue; M, molecular weight markers.


a
b
c
M
92,500
66,200
45,000
31,000
21,000

antibody (fig. 15, lane b)

These results demonstrate that the 22 nm particles released by 4.10 cells are composed primarily of protein that migrates with the mobility of the major envelope protein of HBV (24 and 27 kd; Peterson, 1981). However, the finding of additional species with molecular sizes >28,000 that react with antibody specific for the pre-S2 region epitopes indicates that these particles also contain significant amounts of envelope proteins with amino acid sequences that can only be translated from mRNAs containing at least the pre-S2 coding region (Heermann et al., 1984).

DISCUSSION

The hepatitis B virus and other hepadna viruses are highly species specific and highly similar with respect to virion and genome structure, nucleic acid sequence and structure of replicative intermediates (Robinson et al., 1982). A common feature of infection by these viruses is production of two major virally encoded polyadenylated RNAs, which are approximately 2.1 kb and 3.5 kb (longer-than-genome length or "pre-genome"), and which initiate in the pre-S and pre-core coding regions of the viral genomes, respectively (Cattaneo et al., 1984; Moroy et al., 1985; Buscher et al., 1985; Enders et al., 1985).

The HBsAg Encoding Transcripts.

The primary start sites for envelope gene transcripts in the livers of hepadna virus infected ducks (Buscher et al., 1985), ground squirrels (Enders et al., 1985), woodchucks (Moroy et al., 1985) and chimpanzees (Cattaneo et al., 1984) have all been mapped to sites within the pre-S region. Although transcripts starting at these sites can lead to production of envelope proteins containing part of the amino acid sequence encoded by the pre-S region (those between the second in phase AUG and the AUG coding for the first methionine in the major envelope protein), they can not be translated to give proteins containing all of the amino acids encoded between the first AUG of the pre-S region and the termination codon at the 3' end of the S ORF. Nevertheless, envelope proteins of the proper sizes and carrying immunogenic determinants, which can only be translated from mRNAs that include the codons of the entire pre-S region, are clearly synthesized and incorporated into viral envelopes and HBsAg particles found in the livers and sera of infected humans (Heermann et al., 1984) and

susceptible animals (Feitelson et al., 1983). An initiation site in HBV which would be appropriate for such transcripts has been detected as a secondary start site for S region transcription in COS cells (African green monkey kidney cells, which express the SV-40 large T antigen), transfected with a plasmid in which the SV-40 promoter was placed 409 bp upstream from the first ATG of the pre-S region (Laub et al., 1983). Siddiqui et al. (1986) detected the same initiation site for the HBsAg transcription in rat fibroblast cells, which were stably transfected with a plasmid carrying linearized copy of the HBV genome - situated upstream of the SV40 origin of replication - and a gene coding for neomycin resistance. This site, which was also detected by in vitro transcription of HBV sequences by RNA polymerase II (Rall et al., 1983), was localized to a position approximately 25 nucleotides downstream from the TATA-like promoter between nucleotides 2785-2791. Low level of activity from this promoter was also documented by Ou and Rutter (1985) in PLC/PRF/5 cells - human liver tumor line derived from patient with HBV infection (Alexander et al., 1976) that contain about 7 copies of HBV genome (Shaul et al., 1984), none of which are intact. Although usage of this initiation site in infected livers is below the level of detection (Cattaneo et al., 1984), low levels of transcript initiation around base 2811 on the integrated HBV genomes of 4.10 cells was visualized in S1 nuclease protection experiments. Since I was unable to confirm this start site by primer extension analysis it is possible that the result obtained from S1 nuclease mapping represents a detection of a 3' splice junction in the longer-than-genome length RNA. Utilization of splicing mechanism in order to facilitate an efficient expression of an internal coding region, through positioning of its functioning

initiation codon in the proximity of the mRNA 5' end, has been shown to occur in retroviruses (Weiss et al., 1982). Hepatitis B virus, which in many ways is similar to retroviruses (Miller and Robinson, 1986), could resort to splicing in order to express the pre-S1 region, however, no evidence exists at present to support this speculation; and in fact, some of the above described studies, in which this initiation site was precisely mapped with primer extension analysis (Laub et al., 1983) and in vitro RNA polymerase II (Rall et al., 1983) transcription assay, argue strongly against it.

Neither S1 nuclease, nor primer extension analysis showed the presence of pre-S1 region containing mRNA, which would initiate downstream from the above described promoter, in D1113 cell line. It is not known why this transcript initiation site is not consistently detected. In a variety of the above described heterologous systems exists the possibility that the SV40 origin of replication, placed either upstream or downstream from the HBV S coding region and its putative promoters, could significantly increase normally occurring low levels of transcription from this TATA-like promoter. The SV40 origin of replication contains a pair of 72 bp direct repeats, which are necessary for efficient transcription of SV40 early genes from a TATA-like promoter (Gruss et al., 1981; Fromm and Berg, 1982; Ghosh et al., 1981). It has also been shown that these elements can enhance transcription of SV40 early as well as heterologous genes in orientation independent manner, - even when placed thousands of base pairs upstream from their promoter element, or downstream from a given transcription unit (Banerji et al., 1981; Moreau et al., 1981). In systems such as the in vitro polymerase II assay and PLC/PRF/5 or 4.10 cells, where

the amount of the longer-than-genome length RNA is either very low or nonexistent, the possibilities that this RNA could interfere with S1 nuclease mapping (through hybridization either to unprotected region of the probe, or over its entire length) and primer extension analysis (through competition for the primer) are eliminated. This conclusion is consistent with the results obtained from D1113 cell line, which - although derived from the same parental line as the 4.10 cells - contains high levels of 3.5 kb RNA and similarly to the situation in infected livers shows no detectable levels of HBsAg mRNA initiated upstream of the pre-S1 region. However, the possibility that the MoMLV sequences, which are present in D1113 cells, could also affect the synthesis of this mRNA cannot be eliminated.

The major start sites for the envelope region transcripts in 4.10 and D1113 cells occur in proximity of nucleotides 3153 and 14, positions quite analogous to those which were mapped in a number of above described infected and transfected cell systems. Thus, my results indicate that the transcriptional apparatus of 4.10 and D1113 cells recognizes with the greatest efficiency the HBsAg mRNA initiation sites located in proximity to the SV-40 late promoter-like sequences, which lie between nucleotides 3117-3135; and although low level of transcription initiation at nucleotide 2811 - 3' to TATA-like sequence - was found in the former cell line, the latter cell line showed no prsence of this transcript. The difference in activities of these two promoters was eliminated, however, when they were taken out of the context of HBV genome and used to initiate transcription of heterologous genes (Siddiqui et al., 1986; Malpiece et al., 1983). In addition, a minor mRNA start site was also shown to occur in both cell lines at

around nucleotide 4, although the result of S1 nuclease analysis of mRNA from 4.10 cells indicated that the frequency of initiation at this position, and at around nucleotide 14 is equivalent. This discrepancy in the relative band intensities (indicating initiation around base 4), observed in S1 nuclease and primer extension analyses, is probably an artifact of S1 nuclease mapping. The reasons for this low level of transcripts initiated at around base 4 are not clear, since there are no AUGs between this and the downstream start site in proximity of base 14; and the next AUG (the start for P24 and GP27 translation) is located at nucleotide 156. Nevertheless, the heterogeneity of mRNA initiation in this region is analogous to the multiple 5' ends of the SV-40 late mRNAs (Canaani et al., 1979; Ghosh et al., 1978), the synthesis of which is controlled by DNA sequences similar to those found 25 to 65 bp upstream from the above described mRNA start sites in the HBV genome (Cattaneo et al., 1983).

The mRNA originating at nucleotide 2811 is the only one capable of encoding the full pre-S region, i. e. the only one that could direct synthesis of the large viral envelope protein; and the low level of this mRNA may account for the fact that I did not detect P39 and GP42 in the immuno-blots from 4.10 cells. The presence of three separate mRNAs for translation of HBV large, middle and major HBV envelope proteins would be an ideal condition for their differential expression, without invoking the use of polycystronic mRNAs. The possibility that the large envelope protein is translated from the longer-than-genome length RNA cannot be excluded, but is very unlikely. A large body of evidence (Kozak, 1981a, 1983, 1986) points to the fact, that in contrast to the mechanism of prokaryotic mRNA translation, eukaryotic ribosomes are not

capable of initiating protein synthesis by direct binding to mRNA at internal AUGs. In support of this model, polycystronic mRNAs are not found in eukaryotes and are very rare in eukaryotic viruses. Although reinitiation of translation from an upstream ORF - following termination of protein synthesis on a downstream cistron - was shown to occur during the cauliflower mosaic virus (a virus with a life cycle very similar to HBV) mRNA translation (Dixon and Hohn, 1984), the termination and initiation codons for the two different proteins synthesized in this way are separated by only one nucleotide. The region between termination of HBcAg ORF and the first AUG of the pre-S region is 417 bases long and contains three additional AUGs, one of which exists in frame within the polymerase coding region. Therefore, translation of the pre-S1 region from the longer-than-genome length RNA would probably be extremely inefficient and hence, existence of separate mRNAs for expression of a protein containing this region represents a more convincing possibility. Nevertheless, the "relaxed" scanning mechanism, proposed by Kozak (1981b), may play an important role in translation of pre-S2 and S sequences. This mechanism takes into account the fact that nucleotides surrounding functional initiation codons are not random, and that certain sequences, particularly purines at positions -3 and +4 (adenine nucleotide in sequence AUG is assigned position 1), increase the frequency of recognition of a given initiation codon by the 40S ribosome. Any deviation from this sequence would render a given AUG codon a weaker initiator, and would introduce certain degree of flexibility and/or "leakiness" to the overall translation initiation process. In support of this hypothesis, it has been found that the synthesis of majority of eukaryotic (cellular and viral) proteins is

initiated at AUG codons which, according to the above discussion, are surrounded by favorable sequences. In addition, all of the non-functional AUGs, which are located either upstream (this situation is very rare in eukaryotes) or downstream from the functional initiation codon, have pyrimidines in -3 and/or +4 position (Kozak, 1981b). The pre-S2 coding domain, in contrast to pre-S1 and S, begins with a less optimal initiation codon. In addition this pre-S2 AUG is located only about 20 bases from mRNA 5' end, whereas the majority of eukaryotic mRNAs have initiation codon normally 40 to 80 bases from their 5' ends (Kozak, 1981a). Although it has been shown that 5' untranslated leader sequences can be as short as 3 bases (Kelley et al., 1982), initiation of translation in such mRNAs is not restricted to the 5' proximal AUG. Taking into account the relative amounts of HBsAg mRNAs, as well as the characteristics of their primary structures discussed above, it is understandable why in a variety of previously described systems, the most abundant product of their translation is the HBV major envelope protein, and the middle and large polypeptides are present in relatively lesser amounts. This conclusion is thoroughly supported by experimental data, which shows that during HBV infection, the major protein is present in largest amounts and the other two polypeptides constitute a much smaller fraction of the total HBV envelope proteins (Heermann et al., 1984), - a situation, which cannot be fully accounted for by the stoichiometry of their corresponding mRNAs (Cattaneo et al., 1983, 1984). The results obtained from the 4.10 cells are also in agreement with the above discussion, since the relative amount of the major envelope protein as compared to the middle one did not correspond to the relative amounts which could be predicted solely on the basis of their

corresponding mRNA concentrations (fig. 9B, lane 1 and fig. 15, lane a). Finally, it is worth noting that in an analogous way herpes virus expresses its thymidine kinase gene (Marsden et al., 1983). The mRNA for this gene contains three in phase initiation codons, which direct synthesis of three different amino terminal proteins.

The Longer-Than-Genome Length HBV Transcripts.

Although my experimental results clearly show that essentially all polyadenylated core antigen specific transcripts in 4.10 cells are of longer-than-genome length, the small amount of these transcripts hinders their thorough characterization. Heterogeneity of 5' ends of these RNAs, which were mapped by S1 nuclease, is unlikely to indicate any characteristic of viral transcription and probably represents detection of junctions between cotranscribed host and integrated viral sequences. This possibility becomes even more convincing when we notice that the primer extension analysis showed only one site for initiation of longer-than-genome length RNA synthesis, this being at around nucleotide 1749 (fig. 10B, lane 1). This site lies approximately 100 nucleotides upstream from a promoter-like sequence (TACATAA, nucleotides 1653 through 1659), which was shown to function in an *in vitro* RNA polymerase II directed transcription system (Rall et al., 1983). A DNA fragment lacking the pre-core region, but including this proposed promoter, has also been found capable of driving expression of a bacterial gene in a hybrid plasmid construct (Shaul et al., 1985). Identification of the core specific transcripts with a length of approximately 3.5 kb suggests that at least some core transcripts, starting at the mapped initiation sites in the 5' member of a tandem genome and ending at the polyadenylation site, are identical to the 3.5

kb HBV transcripts found in liver during active infection. However, the larger 4.5 and 6.5 kb transcripts cannot be the same as the minor fraction of very large HBV transcripts formed during infection. Unless the 4.5 and 6.5 kb transcripts have initiated at envelope gene promoters and passed through at least one HBV polyadenylation signal, these large transcripts would have to include non-HBV sequences at either their 3' or 5' ends (initiation within the HBV genome and termination at a downstream polyadenylation signal outside of the tandem or initiation upstream from the tandem and termination at the polyadenylation signal in the 3' member). No transcript starting 5' to a tandem and ending at a polyadenylation signal in the 5' member of the tandem would be detected by our core specific probe.

Contrasting results were obtained from analysis of D1113 cell RNAs, where the longer-than-genome length RNA constitutes more than 50% of the total viral RNAs. Until now, these high amounts of the 3.5 kb transcript were observed only in infected livers (Cattaneo et al., 1984; and fig 8), and typically low levels of this RNA have been seen in transfected cell systems (Gough, 1983; Pourcell et al., 1982; Zelent et al., submitted). This could be one of the main reasons for the absence of HBV replication in tissue culture. The 5' ends of the longer-than-genome length RNA isolated from D1113 cells were easily mapped by S1 nuclease and primer extension analyses. A total of three start sites were detected. The major initiation site lays in a close proximity ($\pm$4 bases) to the pre-C AUG, and the 5' ends for the two minor transcripts were found to lay within 60 nucleotides upstream from it. These results are consistent with reports analysing the initiation of transcription of the longer-than-genome length RNA in livers of

animals infected with their corresponding species specific hepadna viruses. The longer-than-genome length RNA of all the mammalian viruses displays similar 5' end heterogeneity as I have detected for the 3.5 kb RNA in D1113 cells. For example, one major and two minor 3.5 kb transcripts are present in GSHV infected livers (Enders et al., 1985). The initiation site for the major transcript was mapped to about 6 nt downstream from the pre-C AUG, and the start sites for the two minor RNAs were found to lay within 25 bases of downstream sequences. In analogous way three start sites for this RNA were also found on WHV genome (Moroy et al., 1985). However, the situation here seems to be complicated by the fact that the pre-C region contains an extra in frame AUG, located 21 bases upstream from the 5' proximal initiation codon. Interestingly, the two major initiation sites for the longer-than-genome length transcripts of WHV were mapped to this pre-C region - one start site between the first and second AUG, and the other slightly downstream of the latter initiation codon. Additionally, a minor initiation site was also detected about 100 bases upstream from the first pre-C AUG. At present it is not known whether this WHV pre-C region structure, as well as the distribution of the initiation sites for the mRNAs encoding it, play any role in the expression of this region at a protein level. The DHBV, on the other hand, synthesizes only one species of longer-than-genome length RNA initiated at a unique site about 12 bases upstream from the first pre-C AUG. This discrepancy is not surprising, however, since analysis of the DHBV genomic organization and nucleotide sequence points to its earlier evolutionary origin than that of the three mammalian viruses. Since the 5' proximal AUG is only 12 bases from the mRNA 5' end, translation of this region is probably not

restricted to the first AUG and - in analogous fashion to the mammalian viruses - core and pre-C-core proteins are synthesized.

The initiation sites for the longer-than-genome length RNAs in HBV infected livers have not been described and data concerning their initiation is limited. From initial studies with RNAs from infected chimpanzee livers H. Will proposed that the major initiation site for this RNA is located few bases downstream from the pre-C AUG (personal communication). Recently, Roosinick et al. (1986) attempted to map the initiation of the core region containing mRNAs - initiated from the HBV genome using its endogenous promoters - in a rat fibroblast cells, stably transfected with a plasmid containing one copy of HBV genome. Their S1 nuclease results indicate initiation sites (of equal intensities) for HBV core region transcripts to occur about 10 nt upstream and 30 nt downstream from the pre-C AUG. However, they did not confirm these sites with primer extension analysis and the possibility exists that these transcription discontinuities may not represent true initiation sites. The estimation of their location from just the S1 nuclease analysis is also imprecise, since - as discussed in the methods section - this enzyme has a tendency to end-nibble and to leave undigested overhangs. Neither my data nor the results described above unambiguously show that the major initiation site for the longer-than-genome length RNA synthesis is located within the pre-C region of the HBV genome. The conservation of the pre-C region in all the hepadna viruses characterized, as well as recent experiments showing that translation of the pre-C sequences targets the HBcAg to cellular membranes and facilitates HBeAg secretion (Ou et al., 1986), argue

strongly in favor of a necessary and functional significance of this region. Although the expression of the pre-C sequences does not interfere with the self-assembly of this protein into particles, such particles are much less stable and are easily dissociated by salt (Miyanohara et al., 1986). The observation that the core particles produced in infected livers are not sensitive to salt treatment (Takahashi et al., 1979), suggests that a core protein without the pre-C sequences is the predominant product of this ORF. The most efficient translation of such protein would be achieved from the mRNA which does not contain the pre-C AUG, i.e. initiates downstream from it. Therefore, it is likely that the larger amounts of core protein, which is the major constituent of viral nucleocapsid, are synthesized from the mRNA initiated within the pre-C sequences, and smaller amounts of pre-C-core polypeptide, which could function in attachment of the viral core to membranes, are translated from the minor mRNA species which initiate upstream from the pre-C AUG.

In addition to being a template for the HBcAg translation and an intermediate in viral replication, the longer-than-genome length transcript is the only HBV mRNA which could serve as a template for the synthesis of viral polymerase. Recently, Will et al. (1986) documented presence of core-pol fusion proteins in HCC tissues that contain HBV DNA. In view of this finding, a mechanism of ribosomal frame shifting - shown to occur during the synthesis of some retroviral gag-pol fusion proteins (Jacks and Varmus, 1985) - seems likely to function in expression of C and P regions of the HBV genome, although possibility of a frame shift mutation can not be completely ruled out.

Thus, my results from analysis of D1113 cells demonstrate: 1) that the two major HBV transcripts are initiated at all of the proper sites in the integrated tandem HBV genomes; 2) that both envelope and core protein specific mRNAs of appropriate size are synthesized; 3) that the C and S region specific transcripts direct synthesis of substantial levels of core and envelope proteins, respectively. The presence of substantial amounts of secreted "e" protein in 4.10 cell cultures indicates that very low levels of longer-than-genome length RNAs are also translated in these cells. The absence of detectable levels of HBcAg could be due to either instability of the core protein in these cells and its rapid degradation to HBeAg, or change in the antigenicity of the core protein which is synthesized in 4.10 cells. The former possibility seems to be more probable and is supported by the fact that the major initiation site for the longer-than-genome length RNAs was mapped upstream of the pre-C AUG; therefore, the predominant translation product of such an RNA would be the less stable (the one containing the pre-C region) form of HBcAg. In contrast to the D1113 cell line, 4.10 cells contained high amounts of transcripts specific for the "X" protein coding region. Evidence exists that these RNAs may even be translated in the 4.10 cells to give a polypeptide reactive with polyclonal antibodies raised against ß-galactosidase-X gene fusion protein synthesized in bacteria (Vitvitski et al., 1986). Finally, much greater levels of HBsAg found in 4.10 cells (Table II) than in D1113 cell line most likely reflect the fact, that in contrast to D1113 cells, the 4.10 cell line was derived by selection for amplification of the HBsAg production and contains a high copy number of integrated HBV DNA.

The low level of 3.5 kb transcript found in 4.10 cells, in comparison to what has been found in infected livers and D1113 cells, may reflect the absence of tissue specific factors required for activity of the HBV enhancer which is thought to control the levels of the 3.5 kb RNA transcription. Since the D1113 cells are not of liver origin, it is not likely that the high amount of longer-than-genome length RNA synthesized in these cells is result of high activity of HBV enhancer. Most likely, the enhancer-like sequences located in the MoMLV LTRs, which flank HBV sequences in the plasmid used to transfect the ψAM22b cells (parental line for D1113 cell line), activate in cis the core region promoter. Interestingly, the levels of 3.5 kb RNA, as compared to 2.1 kb transcript, are higher in D1113 cells than in HBV infected chimpanzee livers. This situation could perhabs indicate greater activity of this putative MoMLV enhancer than the enhancer identified in the HBV genome. However, in view of lacking experimental evidence, this conclusions must remain purely hypothetical.

Transcripts from the Short Strand of HBV.

Although, at this time, I can only speculate on the role that the transcripts from the (+)-strand could have in life cycle of HBV, I have clearly shown the existence of at least one such transcript which initiates and terminates within the integrated viral genomes in HBV transfected 3T3 cells (4.10 cell line). The length of this transcript, as well as the fact that it is polyadenylated, indicates that it is an RNA polymerase II transcript. Its structure also suggests a number of possible functions.

1) Regulation of viral gene expression. In several systems (Izant and Weintraub, 1985; Marx, 1984), so called antisense RNA suppresses the

expression of exogenous and endogenous genes - presumably through hybridization with sense mRNAs. The HBV (-)-polarity RNA, which I have described, could function as such antisense RNA since it is complementary to all of the known HBV sense transcripts; therefore, it could potentially interfere with the synthesis of all known HBV proteins. Since the amount of antisense HBV RNA in 4.10 cells is much less than that of the 2.1 kb envelope mRNA, I could not expect to detect any effect which it might have on synthesis of HBsAg. However, annealing with the low level of longer-than-genome length transcript may either influence the extent to which these RNAs are used for core or "e" antigen production, or their ability to interact with core proteins for assembly into immature virions.

2) Message for a viral protein. The poly(A)$^+$ (-)-polarity RNA is transcribed from a region including the only major ORF encoded by the (+)-strand. It could also be processed to combine three smaller ORFs and thus code for as yet unidentified viral protein(s).

3) Regulation of viral replication. The initiation site of antisense HBV RNA is in the region of the HBV genome presumed to be the origin of its replication (Summers and Mason, 1982); therefore, this RNA is capable of affecting the synthesis of replicative intermediates.

Although the results presented here do not prove that minus polarity HBV transcripts are anything other than a curiosity of integrated genomes, they give an interesting and entirely new example of naturally occurring synthesis of an antisense RNA in eukaryotic cells. Further studies, with RNAs isolated from HBV infected chimpanzee or human livers, will be necessary in order to unequivocably implicate or rule

out the involvement of transcription from the (+)-strand of HBV in its life cycle.

Conclusions and Prospects.

Despite the fact that within the last several years a great deal of information on HBV has been revealed, many questions still remain unanswered. Perhaps the least understood is also the most important characteristic of this virus, i.e. the mechanisms by which it causes liver damage and/or HCC. Some of the other aspects, which need to be elucidated, concern the viral assembly and maturation, as well as the mechanism of entry into the hepatocytes.

The most important results of this study were obtained from experiments with D1113 cells and concern the synthesis of the "pre-genomic" 3.5 kb HBV mRNA. The D1113 is the first cell line which produces such great amounts of longer-than-genome length RNAs from the endogenous viral promoters. The high levels of these RNAs in D1113 cells allowed me to unambiguously map their 5' ends, which were found in agreement with initiation sites proposed for these RNAs in view of the results obtained from studies of other hepadna viruses. Since the longer-than-genome length RNA is perhaps the most important intermediate in HBV life cycle, other experiments were conducted to see if these cells showed additional markers of viral replication. Subsequently, evidence was obtained indicating the presence of episomally replicating HBV DNA, as well as immature viral particles containing HBV DNA and migrating with approximate density of 1.34 gm/ml in CsCl equilibrium gradients (Sells et al., 1986). Therefore, this cell line - neither tissue nor species specific - supports at least partial replication of this virus and may prove useful for future studies concerning

elements in the HBV genome which control viral replication and gene expression.

From several studies, including my own, we can visualize the following way in which HBV most likely expresses its genes: 1) the major, middle and large envelope proteins are synthesized from three different mRNAs, which are initiated at short distances upstream of their corresponding translation initiation codons (additionally, the relaxed scanning mechanism may play a role in the expression of the major and minor proteins from a polycystronic mRNA); 2) the longer-than-genome length RNAs serve as templates for synthesis of two forms of the HBcAg (with or without the pre-C region) and the viral polymerase, which may be synthesized in form of core-pol polyprotein precursor; the "X" protein is synthesized either from a spliced messenger, proposed by Simonsen and Levinson (1983), or from small mRNAs which may initiated slightly upstream of the X ORF (Saito et al., 1986). The HBV promoters which could direct synthesis of the above mentioned RNAs, although deduced from the DNA sequence by comparison with the sequences of promoters previously defined in other viruses, have not been studied. Further experiments (such as an <u>in vitro</u> mutagenesis) are necessary in order to precisely show which sequences in HBV genome - and to what extend - are responsible for control of transcription and/or translation of the viral genes. Perhaps the rapid progress in culturing the virus in our (Sells et al., 1986) and other laboratories (Sureau et al., 1986) will help to clarify these and other questions, which address the molecular biology of HBV and pathogenesis of the course of infection.

LITERATURE CITED

Alexander, J.J., E.M. Bey, E.W. Geddes, and G. Lecatsas. 1976. Establishment of a continuously growing cell line from primary carcinoma of the liver. S.Afr.Med.J. 50:2124-2128.

Aviv, H., and P. Leder. 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc.Natl.Acad.Sci. U.S.A. 69:1408-1412.

Banerji, J., S. Rusconi, and W. Schaffner. 1981. Expression of a ß-globin gene is enhanced by remote SV40 DNA sequences. Cell 27:299-308.

Beasley, R.P., L.Y. Hwang, C.C. Lin, and C.S. Chein. 1981. Hepatocellular carcinomas and hepatitis B virus. A prospective study of 22,707 men in Taiwan. Lancet 2:1129-1133.

Berk, A.J., and P.A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell 12:721-732.

Blum, H.E., A.T. Hasse, J.D. Harris, D. Walker, and G.N. Vyas. 1984. Asymetric replication of hepatitis B virus DNA in human liver: demonstration of cytoplasmic minus-strand DNA by blot analyses and *in situ* hybridization. Virology 139:87-96.

Broom, S., and W. Gilbert. 1985. Rous sarcoma virus encodes a transcriptional activator. Cell 40:537-546.

Budkowska, A., B. Kalinowska, and A. Nowoslawski. 1978. Identification of two HBeAg subspecificities revealed by chemical treatment and enzymatic digestion of liver-derived HBcAg. J.Immunol. 123:1415-1416.

Burrell, C.J., P. Mackay, P.J. Greenaway, P.H. Hofschneider, and K. Murray. 1979. Expression in *Escherichia coli* of hepatitis B virus DNA sequences cloned in plasmid pBR322. Nature 279:43-47.

Buscher, M., W. Reiser, H. Will, and H. Schaller. 1985. Transcripts and the putative RNA pregenome of duck hepatitis B virus: implication for reverse transcription. Cell 40:717-724.

Canaani, D., C. Kahana, A. Mukamel, and Y. Grouer. 1979. Sequence heterogeneity at the 5' termini of late simian virus 40 19S and 16S mRNAs. Proc.Natl.Acad.Sci., USA 76:3078-3082.

Cattaneo, R., H. Will, N. Hernandez, and H. Schaller. 1983. Signals regulating hepatitis B surface antigen transcription. Nature 305:336-338.

Cattaneo, A., H. Will, and H. Schaller. 1984. Hepatitis B virus transcription in the infected liver. EMBO J. 3:2191-2196.

Charnay, P., C. Pourcel, A. Louise, A. Fritsch, and P. Tiollais. 1979. Cloning in Escherichia coli and physical structure of hepatitis B virion DNA. Proc.Natl.Acad.Sci. USA 76:2222-2226.

Chirgwin, J.M., A.E. Przybyla, R.J. MacDonald, and W.J. Rutter. 1974. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294-5299.

Christman, J.K., M. Gerber, P.M. Price, C. Flordellis, J. Edelman, and G. Acs. 1982. Amplification of expression of hepatitis B surface antigen in 3T3 cells cotransfected with dominant-acting gene and cloned viral DNA. Proc.Natl.Acad.Sci. U.S.A. 79:1815-1819.

Cone, R.D., and R.C. Mulligan. 1984. High-efficiency gene transfer into mammalian cells: generation of helper free recombinant retrovirus with broad mammalian host range. Proc.Natl.Acad.Sci., USA 81:6349-6353.

DeGroot, J., J. Fevery, and L. Lepontre. 1978. Long term follow-up of chronic active hepatitis of moderate severity. Gut 19:510-513.

Dejean, A., L. Bougneleret, K.-H. Grzeschik, and P. Tiollais. 1986. Hepatitis B virus DNA integration in a sequence homologous to v-erb-A and steroid receptor genes in a hepatocellular carcinoma. Nature 322:70-72.

Dixon, L.K., and T. Hohn. 1984. Initiation of translation of the cauliflower mosaic virus genome from a polycystronic mRNA: evidence from deletion mutagenesis. EMBO J. 3:2731-2736.

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.

Feitelson, M.A., P.L. Marion, and W.S. Robinson. 1983. The nature of polypeptides larger in size than the major surface antigen components of hepatitis and like viruses in ground squirrel, woodchuck and ducks. Virology 130:79-90.

Flintoff, W.F., S.V. Davidson, and L. Siminovitch. 1976. Isolation and partial characterization of three methotrexate-resistant phenotypes from chinese hamster ovary cells. Somat.Cell.Genet. 2:245-261.

Fromm, M., and P. Berg. 1982. Deletion mapping of DNA regions required for SV40 early region promoter function in vivo. J.Mol.Appl.Genet. 1:457-481.

Galibert, F., T.N. Chen, and E. Mandart. 1982. Nucleotide sequence of a cloned woodchuck hepatitis virus genome: comparison with the hepatitis B virus sequence. J.Virol. 41:51-65.

Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, and P. Charnay. 1979. Nucleotide sequence of the hepatitis B virus genome (subtype awy) cloned in E. coli. Nature 281:646-650.

Ganem, D. 1982. Persistent infection of humans with hepatitis B virus: mechanisms and consequences. Rev.Inf.Dis. 4:1026-1042.

Gerlich, W.H., U. Goldmann, R. Muller, W. Stibbe, and W. Wolff. 1982. Specificity and localization of the hepatitis B virus associated protein kinase. J.Virol. 42:761-766.

Gerlich, W.H., and W.S. Robinson. 1980. Hepatitis B virus contains protein attached to the 5' terminus of its complete DNA strand. Cell 21:801-809.

Ghosh, P.K., P. Lebowitz, F.J. Frisque, and Y. Gluzman. 1981. Identification of a promoter component involved in positioning of the 5' termini of simian virus 40 early mRNAs. Proc.Nat.Acad.Sci., USA 78:100-104.

Ghosh, P.K., V.B. Reddy, J. Swinsloe, P. Lebowitz, and S.M. Weissman. 1978. Heterogeneity and 5'-terminal structures of the late RNAs of simian virus 40. J.Mol.Biol. 126:813-846.

Gough, N.M. 1983. Core and e antigen synthesis in rodent cells transformed with hepatitis B virus DNA is associated with greater than genome length viral messenger RNAs. J.Mol.Biol. 165:683-699.

Gough, N.M., and K. Murray. 1982. Expression of the hepatitis B virus surface, core and e antigen genes by stable rat and mouse cell lines. J.Mol.Biol. 162:43-67.

Graham, F.L., and A.J. van der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology 52:456-467.

Grosschedl, R., and M.L. Birnsteil. 1980. Identification of regulatory sequences in the prelude sequences of an H2A histone gene by the study of specific deletion mutants in vivo. Proc.Natl.Acad.Sci., USA 77:1432-1436.

Gruss, P., R. Dhar, and G. Khoury. 1981. The SV40 tandem repeats as an element of the early promoter. Proc.Natl.Acad.Sci., USA 78:943-947.

Heermann, K.H., U. Goldman, W. Schwartz, T. Seyffarth, H. Baumgarten, and W. Gerlich. 1984. Large surface proteins of hepatitis B virus containing the pre-S sequence. J.Virol. 52:396-402.

Izant, J.G., and H. Weintraub. 1985. Constitutive and conditional suppression of exogenous and endogenous genes by anti-sense RNA. Science 229:345-352.

Jacks, T., and H.E. Varmus. 1985. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science 230:1237-1242.

Jameel, S., and A. Siddiqui. 1986. The human hepatitis B virus enhancer requires trans-acting cellular factor(s) for activity. Mol.Cell.Biol. 6:710-715.

Jeffreys, A.J., and R.A. Flavell. 1977. A physical map of the DNA regions flanking the rabbit ß-globin gene. Cell 12:429-439.

Kelley, D., C. Coleclough, and R.P. Perry. 1982. Functional significance and evolutionary development of the 5'-terminal regions of immunoglobulin variable-region genes. Cell 29:681-689.

Koch, S., A.F. von Loringhoven, P.H. Hoffschneider and R. Koshy. 1984. Amplification and rearrangement in hepatoma cell DNA associated with integrated hepatitis B virus DNA. EMBO J. 3:2185-2189.

Koretz, R., K.J. Levin, D.J. Rebhun, and G.L. Gifnick. 1978. Hepatitis B surface antigen carriers - to biopsy or not to biopsy. Gastroenterology 75: 860-863.

Koshy, R., A.F. von Loringhoven, S. Koch, O. Marquardt, and P.H. Hofschneider. 1984. Structure and function of integrated HBV genes in the human hepatoma cell line PLC/PRF/5, p. 265-273. In N. Vyas, J.L. Dienstag, and J.H. Hoofnagle (eds), Viral Hepatitis and Liver Disease. Grune and Stratton, Orlando, Florida.

Kozak, M. 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc.Natl.Acad.Sci., USA 83:2850-2854.

Kozak, M. 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol.Rev. 47:1-45.

Kozak, M. 1981a. Mechanism of mRNA recognition by eukaryotic ribosomes during initiation of protein synthesis. Curr.Top.Microbiol.Immunol. 93:81-123.

Kozak, M. 1981b. Possible role of flanking nucleotides in recognition of AUG iinitiation codon by eukaryotic ribosomes. Nucleic Acids Res. 9:5233-5252.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.

Landers, T.A., H.B. Greenberg, and W.S. Robinson. 1977. Structure of hepatitis B Dane particle DNA and nature of the endogenous DNA polymerase reaction. J.Virol. 23:368-376.

Laub, O., L.B. Rall, M. Truett, Y. Shaul, D.N. Standring, P. Valenzuela, and W.J. Rutter. 1983. Synthesis of hepatitis B surface antigen in mammalian cells: expression of the entire gene and the coding region. J.Virol. 48:271-280.

Lehrach, H., D. Diamond, J.M. Wozney, and H. Boedtker. 1977. RNA molecular weight determination by gel electrophoresis under denaturing conditions. A critical reexamination. Biochemistry 16:4743-4751.

Lien, J.-M., C.E. Aldrich, and W.S. Mason. 1986. Evidence that a capped oligoribonucleotide is the primer for duck hepatitis B virus plus-strand DNA synthesis. J.Virol. 57:229-236.

Liu, C.-C., D. Yansura, and A.D. Levinson. 1982. Direct expression of hepatitis B surface antigen in monkey cells from an SV40 vector. DNA 1:213-221.

Lozeron, H.A., and W. Szybalski. 1966. Suppression of RNA precipitation during Cs_2SO_4 density gradient centrifugation. Biochem.Biophys.Res.Comm. 23:612-618.

Machida, A., S. Kishimoto, H. Ohnuma, K. Baba, T. Ito, H. Miyamoto, G. Funatsu, K. Oda, S. Usuda, S. Togami, T. Nakamura, Y. Miyakawa, and M. Mayumi. 1984. A polypeptide containing 55 amino acid residues coded by the pre-S region of hepatitis B virus deoxyribonucleic acid bears the receptor for polymerized human as well as chimpanzee albumins. Gastroenterology 86:910-918.

Malpiece, Y., M.L. Michel, G. Carloni, M. Revel, P. Tiollais, and J. Weissenbach. 1983. The gene S promoter of hepatitis B virus confers constitutive gene expression. Nucleic Acids Res. 11:4645-4654.

Mandart, E., A. Kay, and F. Galibert. 1984. Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison with woodchuck and human hepatitis virus sequences. J.Virol. 49:782-792.

Maniatis, G., E.F. Fritsch, and J. Sambrook. 1982. Molecular Cloning Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Mann, R., R.C. Mulligan, and D. Baltimore. 1983. Construction of retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 33:153-159.

Marion, P.L., L. Oshiro, D.C. Regnery, G.H. Scullard, and W.S. Robinson. 1980. A virus in beechey ground squirrels which is related to hepatitis B virus in humans. Proc.Natl.Acad.Sci., USA 77:2941-2945.

Marsden, H.S., L. Haarr, and C.M. Preston. 1983. Processing of herpes simplex virus proteins and evidence that ranslation of thyamidine kinase mRNA is initiated at three separate AUG codons. J.Virol. 46:434-445.

Marx, J.L. 1984. New ways to "mutate" genes. Science 225:819.

Mason, W.C., C. Aldrich, J. Summers, and J.M. Taylor. 1982. Asymmetric replication of duck hepatitis B virus DNA in liver cells. Free minus strand DNA. Proc.Natl.Acad.Sci. U.S.A. 79:3997-4001.

Mason, W.S., G. Seal, and J. Summers. 1980. Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus. J.Virol. 36:829-836.

Maxam, A.M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base specific chemical cleavages, p. 499-560. In L. Grossman and K. Moldave (eds), Methods in Enzymology Vol. 65. Academic Press, N.Y.

Melton, D.A., P.A. Krieg, M.R. Regagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmid containing bacteriophage SP6 promoter. Nucleic Acids.Res. 12:7035-7056.

Meyers, M.L., L. Vitvitski-Trepo, N. Nath, and J.J. Sninsky. 1986. Hepatitis B virus polypeptide X: expression in Escherichia coli and identification of specific antibodies in sera from hepatitis B virus-infected humans. J.Virol. 57:101-109.

Michel, M.L., P. Pontiso, E. Sobczak, Y. Malpiece, R.E. Streeck, and P. Tiollais. 1984. Synthesis in animal cells of hepatitis B surface antigen particles carrying a receptor for polymerized human serum albumin. Proc.Natl.Acad.Sci. U.S.A. 81:7708-7712.

Miller, R.H., and W.S. Robinson. 1986. Common evolutionary origin of hepatitis B virus and retroviruses. Proc.Natl.Acad.Sci., USA 83:2531-2535.

Miller, R.H., C.-T. Tran, and W.S. Robinson. 1984. Hepatitis B virus particles of plasma and liver contain viral DNA-RNA hybrid molecules. Virology 139:53-63.

Molnar-Kimber, K.L., J.W. Summers, and W.S. Mason. 1984. Mapping of cohesive overlap of duck hepatitis B virus DNA and of the site of initiation of reverse transcription. J.Virol. 51:181-191.

Moreau, P., R. Hen, B. Wasylyk, R. Everett, M.P. Gaub, and P. Chambon. 1981. The SV40 72 base pair repeat has a striking affect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res. 9:6047-6067.

Moriarty, A.M., H. Alexander, R.A. Lerner, and G.B. Thornton. 1985. Antibodies to peptides detect new hepatitis B antigen: serological correlation with hepatocellular carcinoma. Science 227:429-433.

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.

Myanohara, A., T. Imamura, M. Araki, K. Sugawara, N. Ohtomo, and K. Matsubara. 1986. Expression of hepatitis B virus core antigen gene in Saccharomyces cerevisiae: synthesis of two polypeptides translated from different initiation codons. J.Virol. 59:176-180.

Neurath, A.R., S.B.H. Kent, N. Strick, and K. Parker. 1986. Identification and chemical synthesis of a host cell receptor binding site on hepatitis B virus. Cell 46:429-436.

Neurath, R.A., S.B.H. Kent, N. Strick, P. Taylor, and C.E. Stevens. 1985. Hepatitis B virus contains pre-S gene-encoded domains. Nature 315:154-156.

Ohori, H., M. Yamaki, S. Onodera, E. Yamada, and N. Ishida. 1980. Antigenic conversion from HBcAg to HBeAg by degradation of hepatitis B core particles. Intervirology 13:74-82.

Ono, Y., H. Onda, R. Sasada, K. Igarashi, Y. Sugino, and K. Nishioka. 1983. The complete nucleotide sequence of the cloned hepatitis B virus DNA: subtype adr and adw. Nucleic Acids Res. 11:1747-1757.

Ou, J.-H., O. Laub, and W. Rutter. 1986. Hepatitis B virus gene function: the precore region targets the core antigen to cellular membranes and causes the secretion of the e antigen. Proc.Natl.Acad.Sci., USA 83:1578-1582.

Ou, J.-H., and W.J. Rutter. 1985. Hybrid hepatitis B virus-host transcripts in human hepatoma cell. Proc.Natl.Acad.Sci., USA 82:83-87.

Pasek, M., T. Goto, W. Gilbert, B. Zink, H. Schaller, P. Mackay, G. Leadbetter, and K. Murray. 1979. Hepatitis B virus genes and their expression in E. coli. Nature 282:575-579.

Peterson, D.L. 1981. Isolation and characterization of the major protein and glycoprotein of hepatitis B surface antigen. J.Biol.Chem. 256:6975-6983.

Peterson, D.L., J.M. Roberts, and G.N. Vyas. 1977. Partial amino acid sequence of two major component polypeptides of hepatitis B surface antigen. Proc.Natl.Acad.Sci., USA 74:1530-1534.

Pourcel, C., A. Louise, M. Gervais, N. Chenciner, M.-F. Dubois, and P. Tiollais. 1982. Transcription of the hepatitis B surface antigen gene in mouse cells transformed with cloned viral DNA. J.Virol. 42:100-105.

Price, P.M., A.Z. Zelent, M.A. Sells, and J.K. Christman. 1986. Antisense transcripts synthesized by hepatitis B virus transfected cells. In R.C. Gallo, W. Haseltine, G. Klein, and H. Zur Hausen (eds), Viruses and human cancer. Alan R. Liss, Inc., New York. In press.

Price, P.M., S. Ostrove, C. Flordellis, M.A. Sells, S. Thung, M. Gerber, J. Christman, and G. Acs. 1983. Characterization of RNA transcripts and virally coded proteins synthesized in mouse fibroblasts transfected with hepatitis B DNA. Biosci.Rep. 3:1017-1026.

Price, P.M., S.Z. Hirschman, and E. Garfinkel. 1980. DNA cloned from the ayw subtype of hepatitis B virus. J.Med.Virol. 6:139-145.

Rall, L.B., D.N. Standring, O. Laub, and W.J. Rutter. 1983. Transcription of hepatitis B virus by RNA polymerase II. Mol.Cell.Biol. 3:1766-1773.

Redeker, A.G. 1975. Viral hepatitis: clinical aspects. Am.J.Med.Sci. 270:9-16.

Rigby, P.W.J., M. Dieckmann, C. Rhodes, and P. Berg. 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J.Mol.Biol. 113:237-251.

Robinson, W.S., P. Marion, and M. Feitelson. 1982. The hepadna virus group: hepatitis B and related viruses. In W. Szmuness, J. Alter, and J. Maynard (eds), Viral Hepatitis: 1981 International Symposium. Franklin Institute Press, Philadelphia.

Robinson, W.S., and L.T. Lutwick. 1976. The virus of hepatitis, type B. New Engl.J.Med. 295:1168-1175 and 1232-1236.

Robinson, W.S. 1975. DNA and DNA polymerase in the core of Dane particle of hepatitis B. Am.J.Med.Sci. 270:151-159.

Roossinck, M.J., S. Jameel, S.H. Loukin, and A. Siddiqui. 1986. Expression of hepatitis B core region in mammalian cells. Mol.Cell.Biol. 6:1393-1400.

Saito, I., Y. Oya, and H. Shimojo. 1986. Novel RNA family structure of hepatitis B virus expressed in human cells, using a helper-free adenovirus vector. J.Virol. 58:554-560.

Sattler, F., and W.S. Robinson. 1979. Hepatitis B viral DNA has cohesive ends. J.Virol. 32:226-233.

Seeger, C., D. Ganem, and H.E. Varmus. 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science 232:477-484.

Seeger, C., D. Ganem, and H.E. Varmus. 1984. Nucleotide sequence of an infectious molecularly cloned genome of ground squirrel hepatitis virus. J.Virol. 51:367-375.

Sells, M.A., M. Shvartsman, A.Z. Zelent, P.M. Price, and G. Acs. 1986. Expression of HBV core protein, pregenomic RNA and episomal DNA by transfected murine cells. Cold Spring Harbor Hepatitis B Viruses Meeting. Cold Spring Harbor, New York.

Shaul, Y. W.J. Rutter, and O. Laub. 1985. Human hepatitis B viral enhancer element. EMBO J. 4:427-430.

Shaul, Y., M. Ziemer, P.D. Garcia, R. Crawford, H. Hsu, P. Valenzuela, and W.J. Rutter. 1984. Cloning and analysis of integrated hepatitis virus sequences from a human hepatoma cell line. J.Virol. 51:776-787.

Shenk, T.E., C. Rhodes, P.W.J. Rigby, and P. Berg. 1975. Biochemical method for mapping mutational alterations in DNA with S1 nuclease: the location of deletions and temperature-sensitive mutations in simian virus 40. Proc.Natl.Acad.Sci., USA 72:989-993.

Siddiqui, A., S. Jameel, and J. Mapoles. 1986. Transcriptional control elements of hepatitis B surface antigen gene. Proc.Natl.Acad.Sci., USA 83:566-570.

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.

Smith, G.L., M. Mackett, and B. Moss. 1983. Infectious vaccinia virus recombinants that express hepatitis B virus surface antigen. Nature 302:490-495.

Sninsky, J.J., A. Siddiqui, W.S. Robinson, and S.N. Cohen. 1979. Cloning and endonuclease mapping of the hepatitis B viral genome. Nature 279:346-348.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J.Mol.Biol. 98:503-517.

Standring, D.N., W.J. Rutter, H.E. Varamus, and D. Ganem. 1984. Transcription of the hepatitis B surface antigen gene in cultured murine cells initiates within pre-surface region. J.Virol. 50:563-571.

Standring, D.N., L.B. Rall, O. Laub, and W.J. Rutter. 1983. Hepatitis B virus encodes an RNA polymerase III transcript. Mol.Cell.Biol. 3:1774-1782.

Summers, J. 1984. Replication of hepatitis B viruses, p. 87-96. In G.N. Vyas, J.L. Dienstag, and J.H. Hoffnagle (eds), Viral hepatitis and liver disease. Grune and Stratton, Orlando, Florida.

Summers, J., and W.S. Mason. 1982. Replication of the genome of a hepatitis B-like virus by reverse transcription of an RNA intermediate. Cell 29:403-415.

Summers, J., J.M. Smolec, and R. Snyder. 1978. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. Proc.Natl.Acad.Sci., USA 75:4533-4537.

Sureau, C., J.-L. Romet-Lemonne, and M. Essex. 1986. Production of hepatitis B virus-like particles in Hep G2 cells after transfection with cloned circular HBV DNA. Cold Spring Harbor Hepatitis B viruses Meeting. Cold Spring Harbor, New York.

Stibbe, W., and W. Gerlich. 1983. Structural relationships between minor and major proteins of hepatitis B surface antigen. J.Virol. 46:626-628.

Takahashi, K., Y. Akahane, T. Gotanda, T. Mishiro, M. Imai, Y. Miyakawa, and M. Mayumi. 1979. Demonstration of hepatitis B e antigen in the core of Dane particles. J.Immunol. 122:275-279.

Toh, H., H. Hayashida, and T. Miyata. 1983. Sequence homology between retroviral reverse transcriptase and putative polymerase of hepatitis B virus and cauliflower mosaic virus. Nature 305:827-829.

Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitocellulose sheets: procedure and some applications. Proc.Natl.Acad.Sci, USA 76:4350-4354.

Tur-Kaspa, R., R.D. Burk, Y. Shaul, and D.A. Shafritz. 1986. Hepatitis B virus contains a glucocorticoid-responsive element. Proc.Natl.Acad.Sci., USA 83:1627-1631.

Valenzuela, P., M. Quiroga, J. Zaldivar, P. Gray, and Rutter, W.J. 1980. The nucleotide sequence of the hepatitis B viral genome and the identification of major viral genes, p. 57-70. In B. Fields, R. Jaenisch, and C.F. Fox (eds), Animal Virus Genetics. Academic Press, New York.

Varmus, H.E. 1984. Do hepatitis B viruses make a genetic contribution to primary hepatocellular carcinoma? p. 411-414. In: N. Vyas, J.L. Dienstag, and J.H. Hoofnagle (eds), Viral Hepatitis and Liver Disease. Grune and Stratton, Orlando, Florida.

Vitvitski, L., M.L. Meyers, J.J. Sninski, P. Chevallier, and G. Acs. 1986. Immunofluorescent studies of hepa DNA X gene products expression in liver of man, woodchuck and duck: evidence for cross-reactivity and correlation with replication. Cold Spring Harbor Hepatitis B Viruses Meeting. Cold Spring Harbor, New York.

Weiss, R., N. Teich, H. Varmus, and J. Coffin. 1982. RNA tumor viruses. Cold Spring Harbor Laboratory Press, New York.

Wigler, M., M. Perucho, D. Kurtz, S. Dana, A. Pellicer, R. Axel, and S. Silverstein. 1980. Transformation of mammalian cells with an amplifiable dominant-acting gene. Proc.Natl.Acad.Sci., USA 77:3567-3570.

Will, H., J. Salfeld, E. Pfaff, C. Mauso, L. Theilmann, and H. Schaller. 1986. Putative reverse transcriptase intermediates of human hepatitis B virus in primary liver carcinomas. Science 231:594-596.

Will, H., R. Cattaneo, E. Pfaff, C. Kuhn, M. Roggendorf, and H. Schaller. 1984. Expression of hepatitis B antigens with simian virus 40 vector. J.Virol. 50:335-342.

Zelent, A.Z., M.A. Sells, P.M. Price, A. Mohamad, G. Acs, and J. Christman. 1986. Murine cells carrying integrated tandem genomes of hepatitis B virus DNA transcribe RNAs from endogenous promoters on both viral strands and express middle and major viral envelope proteins. J.Virol., submitted.