MOLECULAR AND CELLULAR BIOLOGY, Oct. 1983, p. 1766–1773
0270-7306/83/101766-08$02.00/0

Vol. 3, No. 10

Transcription of Hepatitis B Virus by RNA Polymerase II

LESLIE B. RALL,† DAVID N. STANDRING, ORGAD LAUB,‡ AND WILLIAM J. RUTTER*
Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143

Received 4 May 1983/Accepted 30 July 1983

We employed an in vitro cell-free transcription system to locate RNA polymerase II promoters on the hepatitis B virus genome. The strongest promoter precedes the surface antigen (HBsAg) gene, which is comprised of a long (500 base pairs) presurface region as well as the mature HBsAg coding sequence. The origin of this transcript was localized by using truncated templates and S1 endonuclease mapping. The activity of the promoter was confirmed in transfection experiments in which the complete HBsAg gene was introduced into monkey kidney cells via a simian virus 40 expression vector. A second RNA polymerase II promoter preceding the HBcAg gene was also active in the cell-free system. The presence of multiple promoters in the hepatitis B virus genome suggests that the relative levels of viral-specific proteins detected in liver and serum may reflect differential or regulated promoter efficiency.

Hepatitis B virus (HBV) is a partially single-stranded DNA virus whose host range is restricted to humans and chimpanzees (reviewed in reference 37). The complete nucleotide sequences of the cloned HBV (11, 27, 38, 40) and the related woodchuck hepatitis virus (10) define the major viral protein coding regions. The organization of genetic information is compact: four reading frames (A, B, C, and S) are found in the viral long strand and one small open reading frame (D) is found in the short strand. The viral surface (S) and core (C) protein coding sequences have been identified and incorporated into both procaryotic (9, 27) and eucaryotic (8, 19a, 25, 32, 35) expression vectors, generating products that react with antibodies directed against these antigens (HBsAg and HBcAg). The predicted sequences suggest that the surface antigen is a transmembrane protein, whereas on the basis of the protamine-like COOH terminal region (40), the core antigen is a DNA-binding protein. The S gene has a long 5′ leader sequence and the C gene has a shorter 5′ leader upstream from the mature coding regions, suggesting that both proteins could be synthesized as precursors. Of the remaining two long-strand open reading frames, the larger (A, ~80% of the viral genome) presumably encodes the viral DNA polymerase. This gene overlaps part of the C gene, all of the S gene, and part of the B gene but would be translated with a different register. No specific protein has yet been associated with the smaller reading frame (B).

This sequence information has not resulted in a detailed characterization of HBV promoters of RNA transcripts, mainly due to the lack of a reproducible tissue culture system for viral propagation. The definition of HBV promoters is important for (i) determining how the viral genetic information is expressed, (ii) assessing a possible oncogenic role of integrated HBV sequences, and (iii) studying replication, which has recently been proposed to involve reverse transcription of a full-length viral RNA transcript (36). The HBV DNA sequence provides some information on potential transcription signals such as TATA boxes and AATAAA-directed polyadenylation sites. However, this is not sufficient information to define the actual viral transcription units. We have, therefore, investigated the transcription of purified HBV DNA in a cell-free transcription system, which has been shown to accurately transcribe a variety of genes including globin (20, 21), ovalbumin (41), polyoma (17), simian virus 40 (SV40) early and late genes (16), and cauliflower mosaic virus (15). Cell-free transcription systems are capable of reproducing subtle transcriptional effects such as the multiple initiation sites found for the SV40 promoters (16). In addition, Green and Roeder (14) have shown that a previously unknown adenovirus promoter detected in in vitro experiments also functions in vivo, demonstrating the predictive power of this experimental system.

In the present work we define two major RNA polymerase II promoters in the HBV genome by

† Present address: Chiron Corporation, Emeryville, CA 94608.
‡ Present address: Department of Genetics, Weizmann Institute of Science, Rehovot, Israel.

using cell-free extracts and the runoff transcription assay. Moreover, one of these promoters inserted in an SV40-based vector is functional in permissive monkey kidney cells (12).

MATERIALS AND METHODS

DNA templates. Details of the templates used in this work have been presented by Standring et al. (34) and are summarized below. Nucleotide numbering of the HBV genome is given relative to the *Eco*RI site as the zero point (0/3,221) in accordance with the generally accepted system (11) and differs from that presented previously (35), in which the *Eco*RI site occurs at nucleotide 1,404. The plasmid containing the adenovirus major late promoter (pSmaF) was obtained from R. Roeder.

In vitro transcription. HeLa cell extracts were prepared according to the procedure of Manley et al. (22). HBV DNA templates were digested with the specified restriction enzyme and then extracted with phenol-chloroform, precipitated twice with ethanol, and dissolved in sterile water. Transcription reactions were performed as described by Manley et al. (22), using 0.5 to 1 μg of DNA per 25-μl reaction. The final reaction mixture contained 15 μl of HeLa cell extract; 12 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; pH 7.9); 60 mM KCl; 7.5 mM $MgCl_2$; 0.2 mM EDTA; 1.3 mM dithiothreitol; 10% glycerol; 600 μM each of ATP, CTP, and UTP; 25 μM GTP; and 10 μCi of [α^{32}P]GTP (410 Ci/mmol; Amersham Corp.). After 60 min at 37°C, the reaction was terminated by the addition of 200 μl of 0.25 M sodium acetate, pH 6.0, containing 0.5% sodium dodecyl sulfate and 100 μg of tRNA per ml. After extraction with phenol-chloroform, the RNA was precipitated twice with ethanol and analyzed on a 5% polyacrylamide–8 M urea gel (24). RNA for S1 mapping (42) was gel purified from a 10× reaction.

Transcription from an SV40/HBsAg recombinant. The construction and propagation of the LSV/HBsAg recombinant is described in detail elsewhere (19a). Cytoplasmic, polyadenylated RNA was isolated (19) from COS cells (12) infected 72 h earlier with the recombinant virus. HBV probes were prepared by digesting plasmid DNAs with either *Ball* (nucleotide 3,010) or *Eco*RI (nucleotide 0/3,221) and labeling with [γ-^{32}P]ATP and T4 polynucleotide kinase. After digestion with *Bgl*II (nucleotide 2,432), the subfragments were purified on gels and annealed to RNA under standard conditions (42). S1 analysis was performed as described in the legends to Fig. 3 and 4. Sequence reactions were performed according to the protocol of Maxam and Gilbert (24). The products of S1 and sequencing reactions were analyzed on 5% polyacrylamide–8 M urea gels and visualized by autoradiography.

RESULTS

Runoff transcription assays. As schematically summarized in Fig. 1, two different transcription templates were analyzed in the HeLa cell-free extracts (22): the gel-purified HBV 3.2-kilobase (kb) insert. which contains the entire viral genome cloned at the unique *Eco*RI site (40), and the plasmid pHBV1.85, which has a 1.85-kb

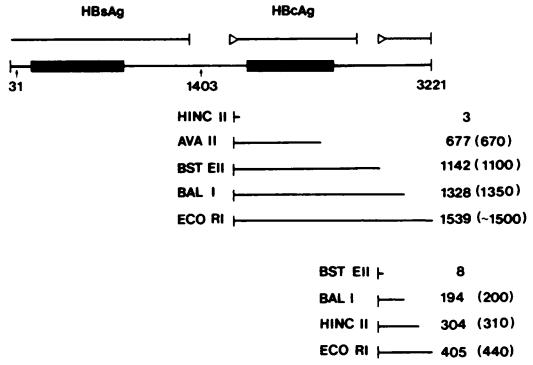

FIG. 1. Summary of the transcripts generated from HBV DNA. The HBV genome is depicted as the 3.2-kb *Eco*RI insert. The mature surface and core antigen proteins are shown as black boxes, and the transcript initiation sites are indicated by open triangles. The *Bam*HI sites are indicated by arrows and are located at nucleotides 31 and 1,403 on the HBV genome (40). The runoff transcripts obtained from the various digests are summarized below with the expected lengths of the RNAs. The estimated transcript sizes are given in parentheses.

*Bam*HI fragment of the HBV genome that encodes the entire C antigen and most of the preS region but not the mature S protein. This plasmid was cut with a variety of restriction enzymes to yield suitable linear templates. The estimated sizes of the runoff transcripts determined relative to ^{32}P-labeled DNA standards and their predicted values, based on the final assignments of the RNA initiation sites, are also presented in Fig. 1.

Surface antigen gene. Transcription of the HBV 3.2-kb insert gave rise to an approximately 440-base RNA species (Fig. 2A, lane 4) that was sensitive to low levels of α-amanitin (Fig. 2A, lane 5) and was therefore synthesized by RNA polymerase II. Since the length of this RNA species closely corresponded to that expected for a transcript originating from the predicted HBsAg promoter in the preS region (40), a series of templates was prepared to specifically test this possibility. A *Bam*HI digest of pHBV1.85 gave a slightly longer transcript (about 470 bases; Fig. 2A, lane 6), consistent with the fact that the *Bam*HI site is about 30 bases downstream from the *Eco*RI site (see Fig. 1). Similarly, *Hinc*II and *Ball* digests of pHBV1.85 gave rise to α-amanitin-sensitive transcripts of, respectively, 310 (Fig. 2B, lanes 2 and 3) and 200 bases (data not shown). A *Bst*EII-digested pHBV1.85 template did not generate the S antigen transcript, consistent with the fact that the enzyme cuts only about eight bases downstream from the anticipated transcription initiation site. From the size estimate of the smallest runoff transcript, an initiation site preceding the preS

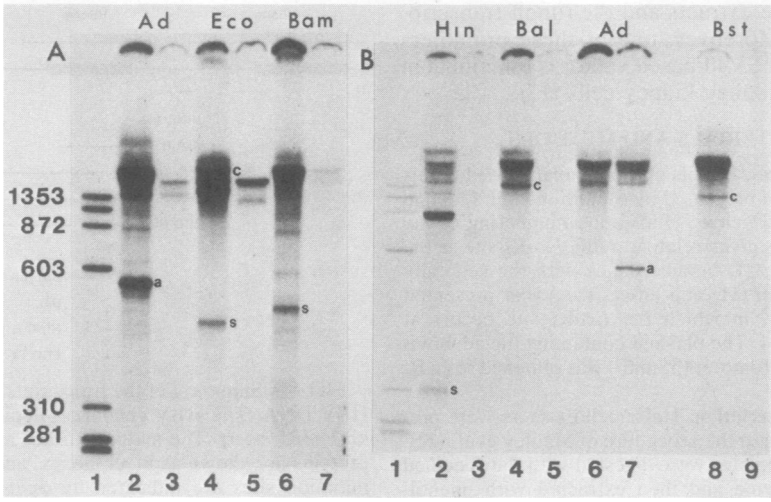

FIG. 2. Autoradiograms of in vitro transcripts directed by HBV DNA. Transcription reactions were performed in the presence of [α-^{32}P]GTP and analyzed as described in the text. The templates used were: (A) lane 4, HBV 3.2-kb *Eco*RI insert; lane 6, *Bam*HI-digested pHBV1.85; (B) pHBV1.85 cleaved with *Hin*cII (lane 2), *Bal*I (lane 4), and *Bst*EII (lane 8). The effect of 1 μg of α-amanitin per ml on the respective reactions is displayed in (A) lanes 5 and 7 and (B) lanes 3, 5, and 9. ^{32}P-labeled, *Hae*III-digested φX174 DNA size markers were included in (A) lanes 1 and 8 and (B) lane 1. Adenovirus pSmaF controls are included in (A) lane 2 and (B) lane 7. The corresponding reactions with α-amanitin are shown in (A) lane 3 and (B) lane 6. The variability in α-amanitin sensitivity among the upper bands was generally not observed when low concentrations (0.1 μg/ml) were used (data not shown). The upper bands may represent end-to-end transcripts or addition of label to the ends of DNA templates in some instances. Transcripts from the core or surface promoters are indicated with the letter c or s, respectively. In panel A, lane 4, the assignment of the core antigen transcript was made on the basis of transcription experiments involving additional truncated templates (data not shown). The origin of the third α-amanitin-sensitive species has not been clarified.

region was located at nucleotide 2,810 ± 10 on the HBV genome. Further analysis of this site by S1 mapping is presented below.

The transcription data indicated that the S antigen promoter is strong in vitro. A comparison between the adenovirus major late transcript generated from the pSmaF control (Fig. 2A, lane 2) and the HBV S transcript of similar size (e.g., Fig. 2A, lane 6) showed only a two- to threefold difference in the intensity of the transcripts.

Core antigen gene. The transcription data provided evidence for the existence of a second in vitro promoter. Transcription of *Bal*I- or *Bst*EII-digested pHBV1.85 yielded RNA species of about 1,350 and 1,100 bases, respectively (Fig. 2B, lanes 4 and 8). This is consistent with initiation occurring in the region preceding the start of the core antigen protein. Further experiments confirmed this result: the *Eco*RI insert of HBV DNA gave rise to an ~1,500-base transcript (Fig. 2A, lane 4), whereas an *Ava*II digest of pHBV1.85 gave a faint transcript of roughly 670 bases (data not shown). All of these transcripts were sensitive to low levels of α-amanitin (Fig. 2A, lane 5; Fig. 2B, lanes 5 and 9) and hence are synthesized by RNA polymerase II. Attempts to position the initiation site more precisely by using shorter truncated templates were not successful because the anticipated RNAs were not above background in the autoradiograms, reflecting the weakness of the core promoter. From the relative intensities of the bands visible for the core (670 bases) and surface (440 bases) antigen RNAs, we estimated that the core promoter is at least 10-fold less effective in the cell-free system. Runoff transcripts of decreasing size give a disproportionately large decrease in incorporation (unpublished data; compare the core antigen transcripts in Fig. 2B, lanes 4 and 8). Therefore, comparative data must be generated from transcripts of similar size. In addition, we and others (R. Roeder, personal communication) have found that inefficient promoters are variably transcribed in different preparations of cell-free extract. For example, the human insulin gene was transcribed in three of nine extracts (unpublished data), and similarly, the core promoter functioned in only two. In contrast, adenovirus major late and HBV S RNAs were synthesized by all extracts.

Based on our experience with the surface antigen transcript and on the observations of others (21), the size estimates for the RNA transcripts are accurate to within ±5% in instances where the DNA markers and RNA species are of similar size. Thus, the results of the

runoff assay with the *Ava*II digest maps the core promoter to a position 670 ± 35 bases from the *Ava*II site (nucleotide 2,359). This localizes the core promoter to nucleotide 1,689 ± 35 on the HBV genome. The only TATA-like sequence, TACATAA, in this region is at position 1,654–1,660, compatible with an initiation site around nucleotide 1,680.

S1 nuclease mapping of the in vitro S transcript. The site of initiation of the HBV S RNA in vitro has been determined by S1 nuclease analysis (42). The 440-base S RNA was synthesized in a 10× in vitro transcription reaction by using the *Eco*RI 3.2-kb insert and was purified by gel electrophoresis. pHBV3.2 DNA was cleaved with *Bal*I (which cuts at nucleotide 3,010), end labeled, and then subcut with *Bgl*II (nucleotide 2,432). Chemical degradation products generated from the 587-nucleotide probe provided a sequence ladder against which the protected fragments were sized. A number of fragments were protected by the in vitro transcript (Fig. 3). Some of the larger species were also visible as a smear in the control lane. However, a single strong band, corresponding to an initiation site at nucleotide 2,809 or 2,810, is visible in the experimental lane but not in the control. A fainter ladder of bands is present beneath this product, stretching as far as the A residue at nucleotide 2,816. Since S1 digestion is known to leave undigested overhangs (14), this latter residue may be the true initiation site.

Initiation of the HBV S RNA in vivo. We also studied the expression of the HBV S gene from SV40 recombinants that either do or do not contain the preS region, but replicate in COS cells (12). Details of these constructions and the general analysis of RNA and protein produced from them will be presented elsewhere (19a). Here, we report that one such recombinant, which contains the entire HBV S and preS regions as an *Ava*I (nucleotide 2,442)-*Bam*HI (nucleotide 1,403) fragment inserted downstream from the SV40 early promoter, appears to in part use the HBV preS promoter to transcribe the S gene region. Two end-labeled probes were used: the 578-base pair (bp) *Bal*I-*Bgl*II probe described above and a 789-bp *Eco*RI-*Bgl*II fragment (labeled at the *Eco*RI site) which contained an additional 211 nucleotides extending toward the start of the mature surface antigen protein coding sequence. The results of the S1 mapping experiment are shown in Fig. 4. A substantial fraction of both probes was fully protected, corresponding to RNAs initiated at the SV40 early promoter region (19a). In addition, a 430-nucleotide fragment of the *Eco*RI-*Bgl*II probe (Fig. 4B) and a 196-nucleotide fragment of the *Bal*I-*Bgl*II probe (Fig. 4A) were present in the experimental lanes but not in the control lanes. (The 500-base product present in the *Bal*I-*Bgl*II experiment, including the control lane, was due to a minor contaminant in the probe preparation, which remained undigested under the conditions of the experiment.) The 430- and 196-nucleotide products correspond to RNA species that map to the initiation site found for the in vitro transcript. Sizing the fragments against samples of the original probes cleaved with the enzyme *Bst*EII (which cuts at nucleotide 2,824) revealed that the protected fragments were approximately 10 nucleotides larger than the *Bst*EII-digested probes (data not shown), suggesting an initiation site at approximately 2,814 for these RNAs.

We conclude that the HBV S gene initiation

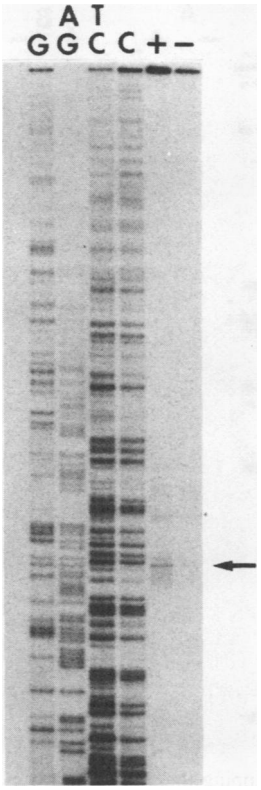

FIG. 3. S1 mapping of the in vitro HBV S transcript. A 578-nucleotide *Bal*I/*Bgl*II probe, ^{32}P end labeled at the *Bal*I sites, was annealed to gel-purified RNA from a 10× in vitro transcription reaction. Annealing was performed at 47°C overnight. Hybrids were quenched with 10 volumes (200 μl) of S1 buffer, and digestion with S1 nuclease (1,000 U; Miles Laboratories) was for 1 h at 37°C. In the control (−) lane, yeast RNA replaced the in vitro transcripts used in the (+) lane. The Maxam and Gilbert reactions (24) performed on the probe were G, A+G, T+C, and C reactions.

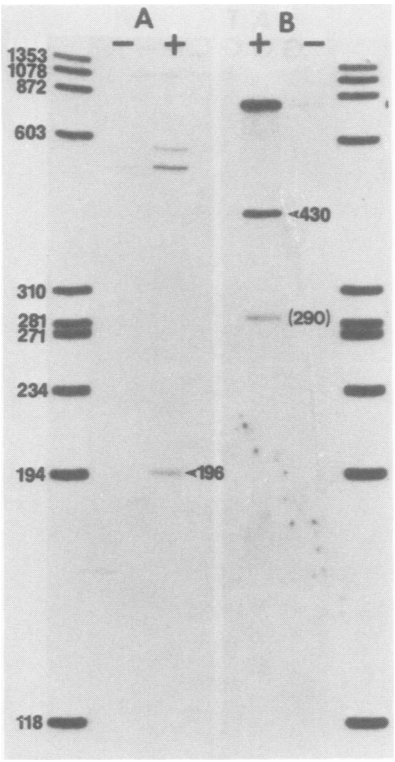

FIG. 4. Mapping the 5′ end of the in vivo HBV S transcript. Two probes were prepared from the cloned HBV DNA, digested either with *Eco*RI or *Bal*I and end labeled with [γ-^{32}P]ATP. The labeled fragments were digested with the restriction endonuclease *Bgl*II, and the resulting *Eco*RI*/*Bgl*II (789 bp) and *Bal*I*/*Bgl*II (578 bp) probes were purified by gel electrophoresis. The DNA probes were annealed with polyadenylated RNA encoded by the LSV-HBpreS recombinant at 50°C for 3 h. Hybrids were quenched with 10 volumes (200 μl) of S1 buffer and then digested with 1,000 U of S1 nuclease at 42°C for 1 h. Protected fragments were precipitated with ethanol in the presence of carrier and analyzed on a 5% polyacrylamide–8 M urea gel. The autoradiogram shows results from the *Bal*I*/*Bgl*II probe (A) and the *Eco*RI*/*Bgl*II probe (B). In each case, the control (−) experiments were performed exactly as for the polyadenylated RNA experiments (+) except that tRNA replaced polyadenylated RNA.

site detected in vitro is also used in vivo. Moreover, the HBV preS promoter is evidently a strong promoter in vivo as well as in vitro; comparison of the intensities of the full-length protected bands in Fig. 4 with the 430- and 196-nucleotide products indicates that the HBV promoter is used only two- to threefold less efficiently than the SV40 early promoter. The *Eco*RI-*Bgl*II probe also gives rise to a second protected species of approximately 290 bases (Fig. 4B). This may represent an RNA processing site or a minor initiation site not detected in the in vitro experiments.

DISCUSSION

In the present study, we defined the positions of two RNA polymerase II-dependent promoters in the HBV genome by using a cell-free transcription system. Both promoters give rise to RNAs transcribed from the viral long strand, which appears to contain all the information for the proteins encoded by the virus. In a separate study we have defined a 700-nucleotide RNA polymerase III transcript which is transcribed from the viral short strand (34).

The initiation site for the S gene transcript was delimited to nucleotide 2,810 ± 10 on the HBV genome by runoff transcription analysis. S1 mapping of the in vitro RNA transcript refined this estimate of the initiation site to within the region 2,809–2,816. On the basis of the consensus sequence (PyAPyPyPyPy) for mRNA initiation reported by Breathnach and Chambon (3), we propose that the A residue at position 2,816 is the most likely start site for the S gene transcript, with the A residue at position 2,813 being a second choice. Inspection of the HBV genome reveals that a TATA-like sequence, TATATAA, is present at position 2,785–2,791, which is approximately 25 nucleotides upstream from the putative initiation site at 2,816.

Gough and Murray (13) and Pourcel et al. (29) demonstrated that the preS region of the HBV genome is necessary for the production of HBsAg in transfected cells. We have directly shown that the preS promoter functions in vivo by using an SV40 recombinant in which the promoter plus the preS and S sequences of HBV were inserted downstream from the SV40 early promoter region. Cytoplasmic RNA isolated from permissive cells infected with the recombinant virus contains a heterogeneous population of HBV-containing sequences (19a). However, analysis of the RNA by S1 mapping experiments revealed that a major RNA species mapped to the in vitro transcription start site. Both the in vitro and SV40 results show that the HBV preS promoter is relatively strong, only two- to threefold weaker than the powerful adenovirus major late promoter or the SV40 early promoter.

The second promoter defined in this study is the core antigen promoter, which gives rise to a runoff transcript initiating at nucleotide 1,689 ± 35 on the HBV genome. The weakness of this promoter has prevented more accurate initiation site estimates from being obtained, either by runoff transcription assays of shorter truncated templates or by S1 nuclease analysis of the purified RNA. There is only one reasonable TATA box in this region of the HBV genome; the sequence TACATAA appears at position

1,654–1,660. We propose that this is part of the promoter sequence, with the A residue at 1,682 or 1,683 as the most probable initiation site. Chakraborty et al. (6) have identified a cell-free transcript unique to this region of the HBV genome, but the data do not distinguish between a promoter for the HBcAg or the putative viral DNA polymerase genes.

An unusual feature of the HBV genome is that the viral DNA found in the infectious particle has a partially single-stranded structure (33, 37). In addition, the nick in the viral long strand is situated just downstream from the core antigen promoter, thus separating the promoter from the coding sequences (33, 37). Furthermore, much of the S gene region may be single-stranded in many viral molecules (33, 37). Thus, we cannot be certain that the promoters defined here with double-stranded templates are functional in the virus. However, there is increasing evidence that covalently closed supercoiled forms of HBV DNA are found in the nuclei of infected cells and that these molecules may be the actual templates for viral transcription in vivo (30, 43). Furthermore, integrated forms of HBV DNA characteristic of hepatoma cells are double-stranded, and at least some of these are transcriptionally active (reviewed in reference 37).

Our studies have not yet detected independent promoters for either the putative viral B or DNA polymerase genes. The former has no obvious TATA sequence but the latter has a TATA sequence, TATAA, at nucleotides 1,919 to 1,924. These promoters may be under temporal or positive regulation by viral gene products, conditions which were not reproduced in our experiments. However, the in vitro experiments may detect only relatively strong promoters. The lack of detection of appropriate transcripts may simply reflect the low biological levels of these gene products and the correspondingly weak promoters for these genes. It is equally possible that two promoters are sufficient to express all the viral genes via polyfunctional messages or alternative splicing pathways or both.

A promoter preceding the preS region creates a paradox concerning surface antigen biosynthesis. Studies on the production of HBsAg in alternate hosts have clearly shown that the preS region is not required for S antigen synthesis, assembly into characteristic 22-nm particles, or subsequent secretion (8, 39). Yet the position of the promoter suggests that preS sequences are present in the initial S gene transcript and must therefore either be removed at the RNA level or there must be processing of a preS protein. However, we have not excluded the possibility that additional promoter sites, not detected in these experiments, can be used.

The initiation of transcription of the core gene (around nucleotide 1,680) occurs only 50 bases from the initiation site for the RNA polymerase III transcript. This suggests that the two promoters overlap and that coordinate regulation of these two genes may occur (see reference 34). In addition, both promoters are within the region of overlap between the viral long and short strands. Opening up of the HBV circular genome followed by filling in of the gaps would generate a linear molecule with the region between the nick in the long strand and the start of the short strand duplicated as a direct repeat at each end, analogous to the structure of a retrovirus long terminal repeat (23). The core promoter in this region is therefore of considerable interest since it might function, as does the retroviral promoter, to generate a full-length RNA transcript that would serve as the template in a reverse transcriptase-mediated replication of HBV (36).

With these findings we can begin to consider whether HBV promoter sequences could play a role in the formation of primary hepatocellular carcinoma. The results of earlier studies indicate that HBV sequences are integrated into the cellular DNA of many hepatomas (4, 5, 7, 31) and in the livers of many carrier patients (18) thought to have a high risk of developing primary hepatocellular carcinoma (2). Recent studies have shown that bursal tumors in chickens may arise from a process initiated by avian leukosis virus integration close to the host oncogene, c-*myc*, resulting in an elevated level of expression (28). In some, though not all, instances, this increased level of expression arises because an avian leukosis virus long terminal repeat promoter reads through into the adjacent c-*myc* sequences (26). Thus, some forms of cancer may result from a promoter insertion event which overides the normal cellular control of the expression of a host oncogene. We have recently isolated integrated HBV sequences from the Alexander human hepatoma cell line (1), which contains at least seven integrated copies of HBV DNA, most of which are incomplete and some of which are rearranged (Y. Shaul and W. J. Rutter, personal communication). Analysis of the cloned integrated HBV fragments derived from this cell line indicates that in at least one clone the strong HBV preS promoter is adjacent to human cellular sequences (Shaul et al., manuscript in preparation), perhaps activating a cellular oncogene which may initiate the onset of HBV-associated hepatocellular carcinoma.

ACKNOWLEDGMENTS

We thank Leslie Spector for preparation of the manuscript.
This research was supported by Public Health Service grant AI 19744 from the National Institutes of Health.

LITERATURE CITED

1. **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.
2. **Beasley, R. P., C. C. Lin, L. G. Hwang, and C. S. Chien.** 1981. Hepatocellular carcinoma and hepatitis B virus: a prospective study of 22,707 men in Taiwan. Lancet **ii**:1129–1133.
3. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
4. **Brechot, C., M. Hadchouel, J. Scotto, M. Fonck, F. Potet, G. N. Vyas, and P. Tiollais.** 1981. State of hepatitis B virus DNA in hepatocytes of patients with HBsAg positive and HBsAg negative liver disease. Proc. Natl. Acad. Sci. U.S.A. **78**:3906–3910.
5. **Brechot, C., C. Pourcel, M. Hodchouel, A. Dejean, A. Louise, J. Scotto, and P. Tiollais.** 1982. State of hepatitis B virus DNA in liver disease. Hepatology **2**:27S–34S.
6. **Chakraborty, P. R., N. Ruiz-Opazo, and D. A. Shafritz.** 1981. Transcription of human hepatitis B virus core antigen gene sequences in an in vitro HeLa cellular extract. Virology **111**:647–652.
7. **Chen, D. S., B. H. Hoyer, J. Nelson, R. H. Purcell, and J. Gerin.** 1981. Detection and properties of hepatitis B viral DNA in liver tissues from patients with hepatocellular carcinoma. Hepatology **2**:42S–46S.
8. **Crowley, C. W., C. C. Liu, and A. D. Levinson.** 1983. Plasmid-directed synthesis of hepatitis B surface antigen in monkey cells. Mol. Cell. Biol. **3**:44–55.
9. **Edman, J. C., R. Hallewell, P. Valenzuela, H. M. Goodman, and W. J. Rutter.** 1981. Synthesis of hepatitis B surface and core antigens in E. coli. Nature (London) **291**:503–506.
10. **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.
11. **Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, and P. Charnay.** 1979. Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in E. coli. Nature (London) **281**:646–650.
12. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
13. **Gough, N. H., 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.
14. **Green, M. R., and R. G. Roeder.** 1980. Definition of a novel promoter for the major adenovirus-associated virus mRNA. Cell **22**:231–242.
15. **Guilley, H., R. K. Dudley, G. Jonard, E. Balazs, and K. E. Richards.** 1982. Transcription of cauliflower mosaic virus DNA: detection of promoter sequences and characterization of transcripts. Cell **30**:763–773.
16. **Hansen, U., D. G. Temen, D. M. Livingston, and P. A. Sharp.** 1981. T antigen repression of SV40 early transcription from two promoters. Cell **27**:603–612.
17. **Jat, P., U. Novak, A. Cowie, C. Tyndall, and R. Kamen.** 1982. DNA sequences required for specific and efficient initiation of transcription at the polyoma virus early promoter. Mol. Cell. Biol. **2**:737–751.
18. **Kam, W., L. B. Rall, E. A. Smuckler, R. Schmid, and W. J. Rutter.** 1982. Hepatitis B viral DNA in liver and serum of asymptomatic carriers. Proc. Natl. Acad. Sci. U.S.A. **79**:7522–7526.
19. **Laub, O., and Y. Aloni.** 1975. Transcription of simian virus 40. V. Regulation of simian virus 40 gene expression. J. Virol. **16**:1171–1183.

19a. **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.

20. **Luse, D. S., J. R. Haynes, D. Van Leeuwen, E. A. Schon, M. L. Cleary, S. G. Shapiro, J. B. Lingren, and R. G. Roeder.** 1981. Transcription of the beta-like globin genes and pseudogenes of the goat in a cell-free system. Nucleic Acids Res. **9**:4339–4354.
21. **Luse, D. S., and R. G. Roeder.** 1980. Accurate transcription initiation on a purified mouse beta-globin DNA fragment in a cell-free system. Cell **20**:691–699.
22. **Manley, J. L., A. Fire, A. Cano, P. A. Sharp, and M. L. Gefter.** 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. Proc. Natl. Acad. Sci. U.S.A. **77**:3855–3859.
23. **Marx, J. L.** 1982. Is hepatitis B virus a retrovirus in disguise? Science **217**:1021–1022.
24. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. **74**:560–564.
25. **Moriarty, A. M., B. H. Hoyer, J. W. Shih, J. L. Gerin, and D. H. Hamer.** 1981. Expression of hepatitis B virus surface antigen gene in cell cultures by using a simian virus 40 vector. Proc. Natl. Acad. Sci. U.S.A. **78**:2606–2610.
26. **Neel, B. G., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin.** 1981. Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. Cell **23**:323–334.
27. **Pasek, M., T. Goto, W. Gilbert, B. Zink, H. Schaller, P. McKay, G. Leadbetter, and K. Murray.** 1979. Hepatitis B virus genes and their expression in E. coli. Nature (London) **282**:575–579.
28. **Payne, G. S., J. M. Bishop, and H. E. Varmus.** 1982. Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphoma. Nature (London) **295**:209–214.
29. **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.
30. **Ruiz-Opazo, N., P. R. Chakraborty, and D. A. Shafritz.** 1982. Evidence for supercoiled hepatitis B virus DNA in chimpanzee liver and serum Dane particles: possible implications in persistent HBV infection. Cell **29**:129–138.
31. **Shafritz, D. A., and M. C. Kew.** 1981. Identification of integrated hepatitis B virus DNA sequences in human hepatocellular carcinomas. Hepatology **1**:1–8.
32. **Siddiqui, A.** 1983. Expression of hepatitis B virus surface antigen gene in cultured cells by using recombinant plasmid vectors. Mol. Cell. Biol. **3**:143–146.
33. **Siddiqui, A., F. Sattler, and W. S. Robinson.** 1979. Restriction endonuclease cleavage map and location of unique features of the DNA of hepatitis B virus, subtype adw_2. Proc. Natl. Acad. Sci. USA **76**:4664–4668.
34. **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.
35. **Stratowa, C., J. Doehmer, Y. Wang, and P. H. Hofschneider.** 1982. Recombinant retroviral DNA yielding high expression of hepatitis B surface antigen. Eur. Mol. Biol. Org. (EMBO) J. **1**:1573–1578.
36. **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.
37. **Tiollais, P., P. Charnay, and G. N. Vyas.** 1981. Biology of hepatitis B virus. Science **213**:406–411.
38. **Valenzuela, P., P. Gray, M. Quiroga, J. Zaldivar, H. M. Goodman, and W. J. Rutter.** 1979. Nucleotide sequence of the gene coding for the major protein of hepatitis B virus surface antigen. Nature (London) **280**:815–819.
39. **Valenzuela, P., A. Medina, W. J. Rutter, G. Ammerer, and B. Hall.** 1982. Synthesis and assembly of hepatitis B surface antigen particles in yeast. Nature (London) **298**:347–350.
40. **Valenzuela, P., M. Quiroga, J. Zaldivar, P. Gray, and W. J. Rutter.** 1980. The nucleotide sequence of the hepati-

tis B viral genome and the identification of the major viral genes, p. 57–70. *In* B. Fields, R. Jaenisch, and C. F. Fox (ed.), Animal virus genetics. Academic Press, Inc., New York.

41. **Wasylyk, B., C. Kedinger, J. Corden, O. Brison, and P. Chambon.** 1980. Specific in vitro initiation of transcription on conalbumin and ovalbumin genes and comparison with adenovirus-2 early and late genes. Nature (London) **285**:367–373.
42. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S beta-globin mRNA precursor and mature 10S beta-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.
43. **Will, H., R. Cattaneo, H. G. Koch, G. Darai, H. Schaller, H. Schellekens, P. M. C. A. van Eerd, and F. Deinhardt.** 1982. Cloned HBV DNA causes hepatitis in chimpanzees. Nature (London) **299**:740–742.