

Synthesis of Hepatitis B Surface Antigen in Mammalian Cells: Expression of the Entire Gene and the Coding Region

ORGAD LAUB,^{1,2} LESLIE B. RALL,¹ MARTHA TRUETT,¹ YOSEF SHAUL,² DAVID N. STANDRING,² PABLO VALENZUELA,¹ AND WILLIAM J. RUTTER^{2*}

Chiron Corporation, Emeryville, California 94608,¹ and Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143²

Received 8 March 1983/Accepted 19 July 1983

We have constructed two simian virus 40 early replacement recombinants that have the coding sequences for hepatitis B virus surface antigen (HBsAg). One construction, LSV-HBsAg, has the coding region for HBsAg but not the portion encoding the putative pre-surface antigen leader. Transformed monkey kidney cells (COS) infected with this recombinant express large quantities of the characteristic partially glycosylated HBsAg molecule, which are assembled into 22-nm particles that appear similar to those produced by human liver cells infected with hepatitis B virus. This result indicates that the pre-surface antigen sequences are not required for the synthesis of HBsAg or its assembly into particulate structures. The second recombinant, LSV-HBpresAg, has the entire surface antigen gene, including the putative promoter and pre-surface antigen region. COS cells infected with this recombinant plasmid produce 40- to 50-fold less HBsAg than those infected with the LSV-HBsAg recombinant plasmid. RNA mapping studies suggest that the transcription of the HBsAg gene is initiated at more than one site, or alternatively, that RNA splicing of transcripts occurs in the pre-surface antigen region.

The classic marker for infection by hepatitis B virus (HBV) is the HBV surface antigen (HBsAg), which appears in two distinct particles of 42 and 22 nm. The larger particle consists of a core containing the viral genome, the core protein, a DNA polymerase, and a phospholipid envelope carrying the surface antigen (31). The smaller 22-nm particle is produced in substantial quantities in infected individuals and contains only the elements of the surface envelope. These HBsAg particles are about 10³-fold more immunogenic than the unassembled HBsAg protein (4). Four protein-coding regions are located in the long strand of the circular HBV DNA (8, 27, 30); the surface antigen gene occupies one of these. The sequence of this DNA has shown that there is a long open reading frame containing three ATG residues preceding the coding region of the mature HBsAg. Recently, using a cell-free extract (18) and truncated HBV DNA templates, we mapped an *in vitro* initiation site of a transcript coding for HBsAg. The results indicate that a strong promoter site precedes the putative leader sequence of HBsAg (L. B. Rall, D. N. Standing, O. Laub, and W. J. Rutter, *Mol. Cell. Biol.*, in press). Further, the successful production of HBsAg in heterologous systems appears to depend upon the presence of this promoter (25) or a substitute promoter (22). In

addition, studies on the mRNA in a hepatoma cell line that produces HBsAg (1) suggest that the major HBsAg-specific transcript(s) contains at least a portion of the pre-surface antigen region (6).

In this study, we have inserted defined portions of the HBV genome into the early region of the simian virus 40 (SV40) virus. These recombinants have been amplified and expressed in SV40-transformed CV-1 cells (COS cells) which synthesize T antigen in sufficient quantities to support the replication of SV40 (9). Two vectors, one with and one without the pre-surface antigen region, have been compared in terms of their replicative, transcriptional, and translational efficiencies. In addition, the sites of initiation of HBsAg-specific transcription have been mapped.

MATERIALS AND METHODS

Enzymes and radioisotopes. Restriction enzymes, polynucleotide kinase, T4 DNA ligase, and *Escherichia coli* polymerase I were purchased from Bethesda Research Laboratories or New England Biolabs. S1 nuclease was obtained from Miles Laboratories. [γ -³²P]ATP was from ICN Pharmaceuticals, Inc. L-[³⁵S]cysteine was purchased from New England Nuclear Corp. Reverse transcriptase was a generous gift from J. Beard.

Cells and virus recombinants. SV40-transformed

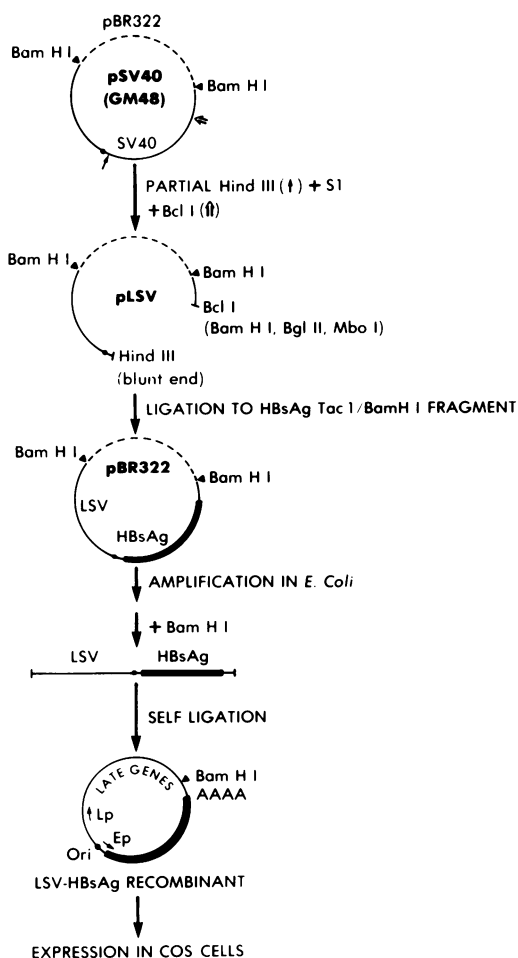


FIG. 1. Strategy used for the construction of the HBV-SV40 recombinants. HBV sequences are indicated by dark solid lines. The figure shows the procedure used for the LSV-HBsAg recombinant. The SV40-HBpresAg recombinant was constructed by an identical procedure with an HBV *Aval* blunt-ended-*Bam*HI fragment instead of the HBV *Tac*I-*Bam*HI fragment.

monkey kidney cells (COS cells) were maintained in Dulbecco modified Eagle medium containing penicillin, streptomycin, and 10% fetal calf serum. The construction of the two LSV-HBV recombinants is described below. Subconfluent COS cultures (10^6 cells per 10-cm plate) were transfected (24) with 5 to 10 μ g of circularized recombinant DNA, and virus stocks were prepared from cell lysates. This stock was used to reinfect COS cells, and the final virus titer (about 10^8 PFU/ml) was estimated by comparing the amounts of viral DNA with a known stock of wild-type SV40 (strain 777) analyzed under the same conditions.

DNA preparation. SV40 strain 777 DNA was used for the construction of the expression vectors. The HBV DNA fragments were purified from pHBV3.2H (29) in which the previously cloned virus, pHBV3.2 (28, 30), was circularized and recloned via the unique *Hae*II site (nucleotide 1440) into the *Hind*III site of

pBR322. All DNA fragments were purified by agarose gel electrophoresis and were eluted by shaking the crushed gel slice in 0.01 M Tris, pH 7.5, containing 1 mM EDTA and 0.2M NaCl. The eluted DNA was filtered through GF/C filters and concentrated with *n*-butanol. Fragments were ligated as described by Maniatis et al. (17). SV40-HBV recombinants cloned in pBR322 (pLSV-HBV recombinants) were amplified in *E. coli* (5) and purified (26). Viral recombinant DNA was extracted from infected COS cells as described by Hirt (11).

RNA analysis. Cytoplasmic polyadenylated [poly (A)] RNA was isolated from infected COS cells as described previously (14, 15). RNA was analyzed by the Northern blot procedure described by Alwine et al. (2). Mapping of RNAs by the S1 nuclease method of Berk and Sharp (3) was done with 5'-end-labeled DNA probes (20). After hybridization at 50°C for 4 h, the DNA:RNA hybrids were digested with excess S1 nuclease and analyzed on a 5% acrylamide-8 M urea sequencing gel (20).

An *Aval*I-*Xba*I* (nucleotides 185 through 250) primer was prepared, and the labeled single strand was obtained by gel electrophoresis followed by elution and ethanol precipitation in the presence of polyribadenylic acid (10 μ g). Primer and LSV-HBsAg poly(A) RNA were hybridized for 8 h at 46°C in 80% formamide containing 0.4 M NaCl and 10mM PIPES [piperazine-*N,N'*-bis(2-ethanesulfonic acid)], pH 6.4. The samples were ethanol precipitated, dried, and suspended in 20 μ l of buffer containing 20 mM Tris-hydrochloride (pH 8.4), 10 mM NaCl, 6 mM MgCl₂, 1 mM dithiothreitol, 1 mM deoxynucleotide triphosphates, and 26 U of reverse transcriptase. Samples of 10 μ l were incubated for 50 min at 45 or 50°C. After phenol extraction, the reactions were analyzed on a 5% polyacrylamide-8 M urea gel.

Protein analysis. COS cells infected for a period of 48 h with the two LSV-HBV virus stocks at a multiplicity of about 20 PFU per cell were maintained for an additional 12 h in cysteine-depleted medium followed by a 6-h labeling with 50 μ Ci of L-[³⁵S]cysteine per ml. A 200- μ l sample of the labeled medium was immunoprecipitated with human anti-HBsAg serum by the SAC technique (19), and the products were analyzed on 12% acrylamide-sodium dodecyl sulfate gels (13). Quantitative radioimmunoassays were done with the AUSRIA II diagnostic kit (Abbott Laboratories). HBsAg particles were purified by affinity chromatography with goat anti-HBsAg covalently bound to Sepharose. The HBsAg was eluted from the column with 3 M potassium thiocyanate and dialyzed against phosphate-buffered saline for structural analysis.

RESULTS

Construction of the LSV-HBV recombinants.

Figure 1 summarizes the general strategy used for the construction of the SV40-HBsAg recombinants. The construction takes advantage of the fact that the early genes of the SV40 DNA are flanked by sites for the restriction endonucleases *Hind*III and *Bcl*I. The *Hind*III site is located at nucleotide number 5171 on the SV40 genome and is 8 nucleotides 5' to the translation initiation codon of the large T antigen gene. The *Bcl*I site is located at nucleotide number 2770,

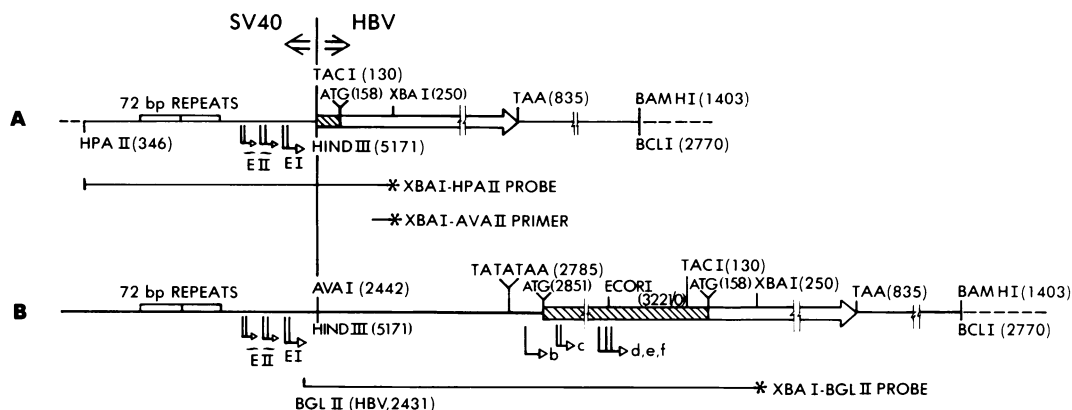


FIG. 2. Detailed map of the SV40 recombinants showing the SV40 early promoter region and the HBV inserts. (A) LSV-HBsAg. (B) LSV-HBpresAg. The vertical line marks the leftward boundary between the SV40 promoter-origin region and the HBV insert. The rightward boundary is at the *Bam*HI-*Bcl*I site. Relevant restriction sites are shown with their nucleotide locations in parentheses; HBV sites are above each figure. SV40 sites are below. The HBV surface antigen gene is represented by an arrow showing the presurface (hatched) and mature surface (open) antigen regions. The putative protein initiation (ATG) and termination (TAA) codons are shown together with the TATA box (TATATAA) preceding the presurface ATG. The SV40 enhancer element is also indicated (72-base pair repeats). Beneath each recombinant is shown the relevant end-labeled probe or primer used to map the 5' end of the RNA species. The asterisk (*) denotes the position of the 32 P label. The major sites of RNA initiation within the SV40 early region are designated by arrows marked EI (early early) and EII (late early). The sites mapped by S1 nuclease analysis within and around the HBV sAg gene are also marked by arrows (b, c, d, e, and f; see Fig. 7).

which is 77 nucleotides 5' to the termination codon of the large T antigen gene. The SV40 genome cloned in the *Bam*HI site of pBR322 (pSV40) was amplified in *E. coli* GM48. pSV40 DNA was linearized by partial digestion with *Hind*III followed by S1 treatment to produce blunt ends (21). The linear pSV40 DNA was digested with *Bcl*I, and the 7.2-kilobase fragment (pLSV) was purified by preparative agarose gel electrophoresis.

The coding region for HBsAg was purified as a *Tac*I-*Bam*HI fragment (nucleotides 130 through 1403 on the cloned HBV genome [30]). The gene which includes the pre-surface antigen region was isolated as an *Ava*I-*Bam*HI fragment (nucleotides 2442 through 1403), and the *Ava*I site in this fragment was filled in to form a blunt end by using T4 DNA polymerase. These fragments were inserted in the pLSV vector. Ampicillin-resistant colonies were screened by hybridization to 32 P-labeled SV40 and HBV probes, and the positive colonies were analyzed by restriction endonuclease mapping. The resulting plasmids contained an SV40 origin of replication, a functional set of SV40 late genes, and HBsAg (with or without the pre-HBsAg region) cloned in the sense of the SV40 early promoter sequence and polyadenylation site. Figure 2 presents a detailed map of the SV40 early promoter region plus the HBV inserts for the two recombinants, which are termed LSV-HBsAg (Fig. 2A) and LSV-HBpresAg (Fig. 2B).

The *Bam*HI fragments containing the LSV-HBV recombinants were isolated, self-ligated to form circular DNA, and transfected into SV40-transformed monkey kidney cells (COS cells) that produce T antigen and are hence permissive for SV40 early replacement recombinants (9). Viruses containing the LSV-HBV DNA were efficiently propagated in this host, and high-titer stocks (10^8 PFU/ml) were obtained after a second round of viral production (see above). Restriction analysis of the purified viral DNA showed that no rearrangements had occurred during replication (data not shown).

For subsequent studies, monolayers of 2×10^6 COS cells were infected with 20 PFU/cell. DNA dot blot analysis (23) showed that both viruses replicated at roughly the same rate and generated comparable copy numbers per cell (Fig. 3).

Transcription of HBsAg sequences from LSV-HBV recombinants. The relative levels of cytoplasmic poly(A) RNA produced 72 h postinfection were assayed by the dot blot technique (23). The results show that the LSV-HBsAg recombinant lacking the presurface sequence accumulated about fourfold more HBsAg-specific RNA than did the LSV-HBpresAg recombinant, which included the entire pre-HBsAg region (Fig. 3).

These RNAs were further analyzed by Northern blot analysis (2). As shown in Fig. 4, the HBsAg-specific RNA transcribed from the

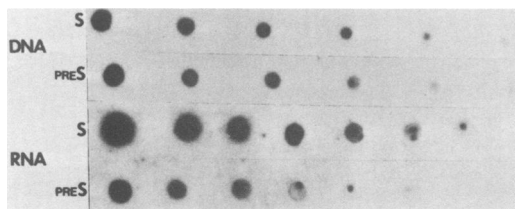


FIG. 3. Replication and transcription of the LSV-HBV recombinants. Monolayers of 2×10^6 COS cells were infected with 20 PFU of the LSV-HBV recombinant virus stocks per cell and were incubated for 72 h at 37°C. Hirt DNA (11) and cytoplasmic poly(A) RNA were prepared as described previously (14, 15). Two-fold serial dilutions of heat-denatured DNA and native poly(A) RNA were spotted on nitrocellulose paper. The initial spot contains RNA or DNA obtained from $\sim 5 \times 10^5$ cells (1/4 plate). The resulting dot blots were hybridized at 42°C with a ^{32}P -labeled HBV probe. The blots were washed in $0.1 \times \text{SSC}$ (SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate at 50°C and autoradiographed. (S) RNA or DNA extracted from COS cells infected with the LSV-HBsAg recombinant; (preS) RNA or DNA prepared from cells infected with the LSV-HBpresAg recombinant.

LSV-HBsAg recombinant migrated as a 1.6-kilobase band. This RNA corresponds in size to a transcript which initiates and terminates at the SV40 early promoter(s) and termination signal, respectively. Thus, the RNA contains about 1,270 nucleotides coding for HBsAg flanked by about 200 nucleotides of SV40 sequences as well as 100 to 200 poly(A) residues. In contrast, a similar analysis of the LSV-HBpresAg transcripts revealed a heterogeneous population of HBsAg-specific RNA ranging from 1,600 to 2,200 nucleotides in length.

The 5' termini of the RNA molecules coded by these recombinants were determined by S1 nuclease mapping with the 5'-end-labeled DNA probes depicted in Fig. 2. A unique *Xba*I site within the HBsAg coding region was labeled with $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ and polynucleotide kinase. Two end-labeled DNA fragments were purified by gel electrophoresis, namely a 1,040-base pair *Bgl*III-*Xba*I* HBV DNA probe and a 537-base pair *Hpa*II-*Xba*I* LSV-HBsAg DNA fragment. The *Bgl*III-*Xba*I* DNA probe contained the entire HBV pre-surface antigen region, including the promoter, but was devoid of SV40 sequences. The *Hpa*II-*Xba*I* DNA probe contained 119 nucleotides of the HBsAg gene and an additional 418 bases of SV40 sequence, including the early promoter region (see Fig. 2). The ^{32}P -labeled probes were denatured and hybridized for 3 h at 41°C to poly(A) RNA isolated from infected COS cells under conditions favoring RNA-DNA duplex formation (80% formamide, 0.4 M NaCl, 10 mM PIPES [pH 6.4]). The

resulting hybrids were digested with S1 nuclease and analyzed on a 5% acrylamide–8 M urea gel. As shown in Fig. 5, the RNA coded by the LSV-HBsAg recombinant protected several bands when hybridized to the *Hpa*II-*Xba*I* probe. Two major bands about 176 and 181 (doublet marked c in Fig. 5) nucleotides long correspond to transcripts initiating in the SV40 E1 promoter region, position 5231 to 5237 on the SV40 genome (10). The somewhat diffuse band (Fig. 5, band marked b) of about 203 nucleotides most likely represents initiation in the SV40 E2 promoter region ascribed to position 14 to 17 on the SV40 genome (10). The larger band (Fig. 5, band marked a) of about 246 nucleotides corresponds to a minor upstream initiation site described by

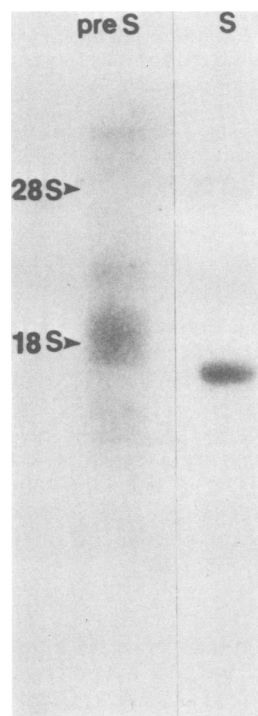


FIG. 4. Northern analysis of the HBsAg-coding mRNA made by COS cells infected with the LSV-HBV recombinants. Monolayers of 2×10^6 COS cells were infected with 20 PFU of either LSV-HBsAg or LSV-HBpresAg virus stocks per cell and were grown for 72 h at 37°C. Poly(A) RNA was prepared from the cytoplasmic fraction (14), denatured in 10 mM methyl mercury, electrophoresed through a 5 mM methyl mercury–1.5% agarose gel, and transferred to nitrocellulose paper. The resulting blots were hybridized to a ^{32}P -labeled HBV DNA, washed in $0.1 \times \text{SSC}$ –0.1% sodium dodecyl sulfate at 50°C, and autoradiographed. (pre-S) RNA extracted from COS cells infected with the LSV-HBpresAg recombinant; (S) RNA encoded by the LSV-HBsAg recombinant. The 18S and 28S ribosomal RNA markers were visualized by staining the gel before blotting.

Hansen et al. (10). An additional protected fragment of 420 nucleotides (mapping to around nucleotide 235 on the SV40 genome) was also seen, whereas the uppermost band comigrated with undigested probe DNA and was seen in some, but not all, experiments (data not shown). This material probably represents reannealed probe, but we cannot exclude the possibility that this protection arose from a very long RNA transcript. The sites of RNA initiation were also analyzed by primer extension with a 66-nucleotide *AvaII-XbaI** HBV DNA fragment (see Fig. 2). Figure 6 shows the results obtained when this primer (lane A) was hybridized with LSV-HBsAg poly(A) RNA and extended with reverse transcriptase at either 45°C (lane B) or 50°C (lane C). The extended products included a prominent band (S) at about 94 nucleotides followed by three groups of bands in the size range of 170 to 175 (c), 203 to 212 (b), and 238 to 245 (a)

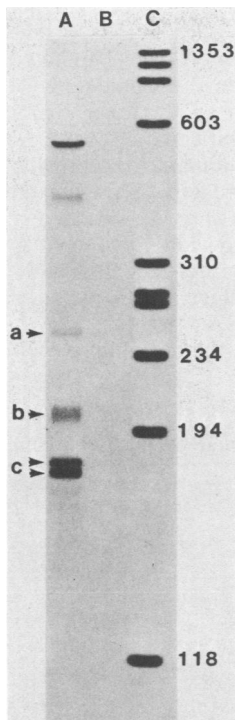


FIG. 5. Mapping the 5' end of RNA encoded by the LSV-HBsAg recombinant. COS cells were infected with about 20 PFU of the LSV-HBsAg virus stock per cell and were grown for 48 h at 37°C. Poly(A) cytoplasmic RNA was extracted (14). The RNA was hybridized to an end-labeled *HpaII-XbaI** LSV-HBsAg probe. The hybridization mixture, which contained 80% formamide, 0.4 M NaCl, and 0.01 M PIPES (pH 6.4), was digested for 1 h with 1,000 U of S1 nuclease and subjected to a 5% acrylamide-8 M urea sequencing gel. (A) LSV-HBsAg RNA; (B) no RNA; (C) DNA size markers (³²P-labeled *HaeIII* φX174).

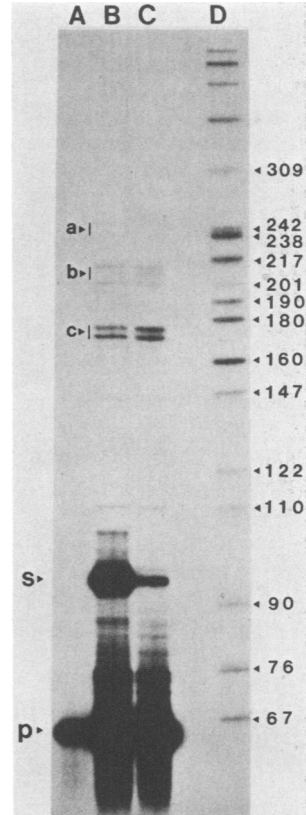
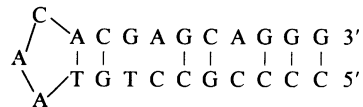


FIG. 6. Analysis of the 5' end of LSV-HBsAg by primer extension. Details of the primer extension method can be found in the text. The lanes in the autoradiogram are as follows: (A) sample of *AvaII-XbaI** primer; (B) primer extension reaction at 45°C; (C) reaction at 50°C; (D) ³²P-end-labeled *MspI* digest of pBR322 as size markers. The sizes of the markers are indicated on the right of the figure. The position of the primer (P) and the 94-base self-primed material (S) are indicated on the left of the figure. The bands marked a, b, and c represent the products extended into the SV40 early promoter region. Nucleotide sequence analysis of the self-primed product(s) reveals that the self-priming originates from a stem-loop structure (below) formed by the last 23 3' nucleotides of the *XbaI-AvaII* probe (data not shown).



nucleotides. The last groups of bands correspond precisely with the products described by Hansen et al. (10) from RNA isolated during the late lytic cycle of wild-type SV40. They reflect initiation events at nucleotides 5231 to 5237 (EI), 21 to 32 (EII), and 56 to 65 on the SV40 genome.

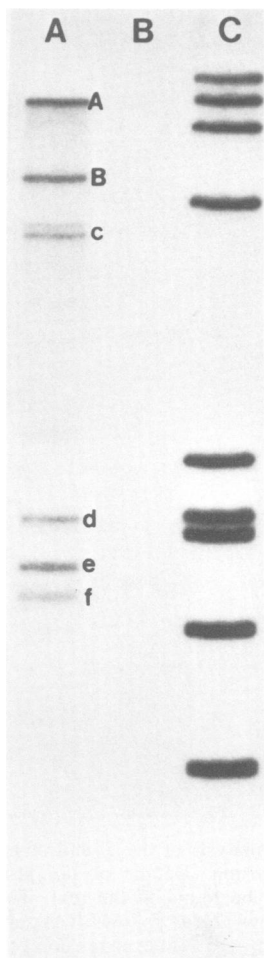


FIG. 7. Mapping of the 5' ends of RNA encoded by the LSV-HBpresAg recombinant. COS cells were infected with the LSV-HBpresAg virus stock (20 PFU/cell), and cytoplasmic poly(A) RNA was extracted 48 h later. The RNA was annealed to an end-labeled *BglII-XbaI** HBV probe. S1 nuclease mapping was done as described in the legend to Fig. 4. (A) LSV-HBpresAg RNA; (B) no RNA; (C) DNA size markers (³²P-*HaeIII* ϕ X174).

These initiation sites are summarized in Fig. 2. In contrast to these bands, the 94-nucleotide species was synthesized in control experiments involving an unrelated RNA (data not shown). Nucleotide sequence analysis of this product (data not shown) shows that it is generated by a self-priming reaction which is suppressed at higher temperatures (compare Fig. 6, lanes B and C). The stem-loop structure which is the basis for the self-priming reaction is shown in the legend to Fig. 6.

The *BglII-XbaI** HBV-specific probe which was annealed to RNA extracted from COS cells

infected by the LSV-HBpresAg recombinant also produced several nuclease S1-protected bands (Fig. 7). Band A, about 1,040 nucleotides long, extends from the *XbaI* site to the end of the DNA probe. This material presumably corresponds to RNA initiating at the SV40 early promoter region. Band B, about 650 nucleotides long, maps to the putative HBV presurface cap site about 25 nucleotides downstream from the TATAAA box at HBV nucleotide 2816. Accurate mapping of this fragment with shorter probes is reported elsewhere (Rall et al., in press). The bands designated c (a doublet) to f may reflect alternate initiation sites within the pre-surface antigen region or products of RNA processing. The locations of these sites are summarized in Fig. 2.

Hepatitis B surface antigen encoded by the LSV-HBV recombinants. The effect of the pre-surface antigen leader region on surface antigen gene expression was assessed by quantitative radioimmunoassays of cell extracts and culture medium (Table 1). The HBsAg levels were about twofold higher in cells infected with LSV-HBsAg than in those infected with LSV-preHBsAg. In contrast, the extracellular medium from COS cells infected with the LSV-HBsAg recombinant contained 40- to 50-fold more immunoreactive HBsAg than was secreted by the LSV-HBpresAg-infected cells. Quantitative assays showed that COS cells infected with the LSV-HBpresAg recombinant containing the entire pre-surface antigen gene secreted about 14 ng of HBsAg per 10⁶ cells during the 24-h period before the initiation of cell lysis (i.e., 48

TABLE 1. Surface antigen protein production^a

| SV40 construct | Immunoreactive HBsAg produced (ng/ml) | |
|----------------|---------------------------------------|-------------------|
| | Medium | Cellular fraction |
| LSV-HBsAg | 620 | 17.9 |
| LSV-HBpresAg | 14.2 | 6.7 |
| LSVinsC2 | <0.1 ^b | <0.1 ^b |

^a Monolayers of 2 × 10⁶ COS cells were infected with 20 PFU of either the LSV-HBsAg or the LSV-HBpresAg recombinant virus stock per ml and were incubated for 48 h at 37°C. The medium was then replaced with 10 ml of new medium and was incubated 24 h at 37°C. Medium and cell lysate (0.5% Nonidet P-40 and 0.5% deoxycholate in phosphate-buffered saline) were assayed for HBsAg by the AUSRIA II diagnostic kit (Abbott Laboratories). The control experiment was performed as above with a recombinant virus stock (LSVinsC2) which contains the human insulin cDNA in place of the HBsAg-coding sequences. This virus has been described elsewhere (15).

^b Same level of radioactivity as the AUSRIA II negative controls.

to 72 h after infection). Under the same conditions, COS cells infected with the LSV-HBsAg recombinant secreted up to 800 ng of HBsAg per 10^6 cells, an equivalent of over 10^7 molecules per cell. The purified HBsAg particles are stable and have been stored for several months without any loss in immunoreactivity. In addition, large quantities (5 to 10 $\mu\text{g/ml}$) of HBsAg accumulate if the viral infection is allowed to progress for 10 days.

The HBsAg molecules produced by the SV40-COS cell system were characterized by immunoprecipitation of ^{35}S -labeled proteins. Sodium dodecyl sulfate gel electrophoresis of the anti-HBsAg immunoreactive material (Fig. 8) disclosed three specific bands. The two predominant peptides, molecular weight 23,000 and 27,000, corresponded to the nonglycosylated and glycosylated forms of HBsAg, respectively. The third band, about 46,000 in molecular weight, may be a HBsAg dimer. The HBsAg protein produced by LSV-HBsAg-infected COS cells appeared similar to that produced in the Alexander cell line. The nature of the HBsAg produced by the LSV-HBpresAg construct has not been clarified further.

The nature of the secreted HBsAg particles was further characterized by electron microscopy. Affinity-purified HBsAg coded by LSV-HBsAg was absorbed onto carbon film grids and negatively stained with uranyl acetate. When examined under the electron microscope (Fig. 9), the LSV-HBsAg preparation contained a somewhat heterogeneous population of spherical particles with a mean size of 22 nm, similar to the spherical particles seen in the serum of infected individuals (12) or in the medium of NIH 3T3 cells into which HBV DNA has been introduced by transfection (32).

DISCUSSION

We have used an SV40-based expression vector to study the role of the long open reading frame contiguous with the HBsAg coding sequences in the biogenesis of the HBV surface antigen. LSV recombinants have been constructed that retain the SV40 late genes, the early promoter(s), and the polyadenylation site, but heterologous sequences were substituted for the coding region of T antigen. The latter function is constitutively produced by the host COS cells, and therefore efficient replication and subsequent packaging of these molecules occurred in this system. This vector-host system was useful for the expression of exogenous DNA. All early SV40 splice junctions were deleted from the SV40 vector, so aberrant splicing caused by these sequences was eliminated. Further, the relatively strong SV40 early promoters and the high number of gene copies per cell achieved in

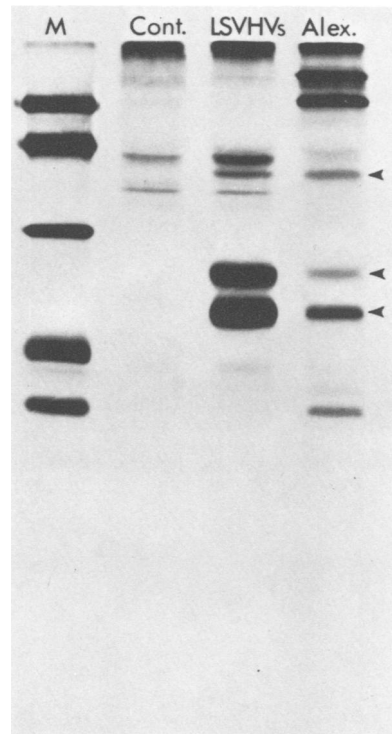


FIG. 8. Synthesis of HBsAg in infected COS cells. Monolayers of Alexander cells or infected COS cells were grown 12 h in cysteine-depleted medium followed by 6 h of labeling with L- ^{35}S]cysteine (50 $\mu\text{Ci/ml}$). The labeled medium (200 μl) was immunoprecipitated with human anti-HBsAg serum and analyzed on 12% Laemmli protein gels (13). (M) protein size markers; (Cont.) mock infected COS cells; (LSV-HVs) COS cells infected with the LSV-HBsAg recombinant; (Alex.) Alexander cells. The size markers were bovine serum albumin (69K), ovalbumin (46K), carbonic anhydrase (30K), β -lactoglobulin (18.4K), and cytochrome *c* (12.3K).

this system resulted in the formation of high cellular levels of recombinant RNA. Two LSV-HBV recombinants were prepared for this study; the recombinant designated LSV-HBsAg had the coding region for mature HBsAg, and the LSV-HBpresAg construct included the pre-surface antigen sequences and the HBsAg promoter (see Fig. 2). The results of hybridization studies and restriction enzyme analyses demonstrate that both viruses replicated to the same extent (about 10^8 PFU/ml) without deletions or rearrangements.

The size of the seemingly homogeneous population of HBsAg-specific transcripts produced from the LSV-HBsAg recombinant is consistent with initiation and termination at the SV40 early gene signals. The results of S1 nuclease and primer extension experiments confirm that the 5' end of the HBsAg mRNA mapped to the SV40

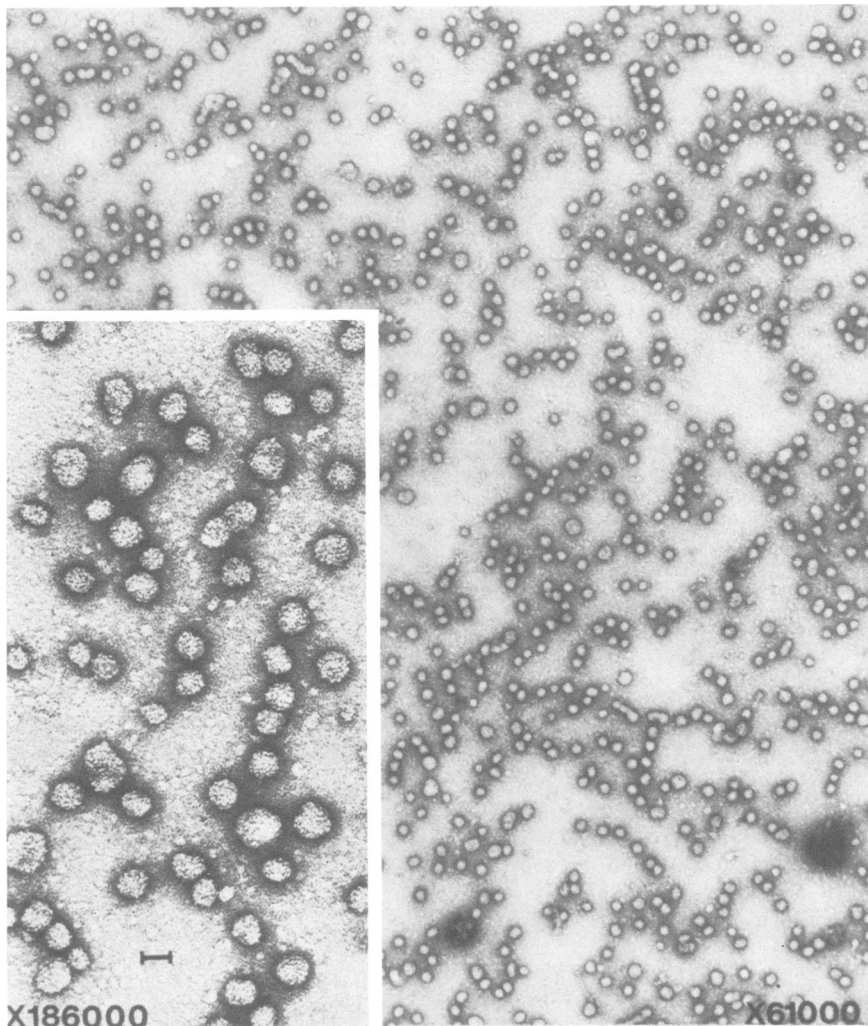


FIG. 9. Electron micrographs of HBsAg particles produced by COS cells infected with the LSV-HBsAg recombinant. HBsAg particles were absorbed to a column of goat anti-HBsAg-specific antibody covalently bound to Sepharose, eluted with 3 M potassium thiocyanate, and visualized by negative staining with 2% phosphotungstic acid. Bar, 22 nm.

early promoter(s) in a complex pattern as described by Hansen et al. (10). The major sites of initiation occurred in the region 5231 to 5237 on the SV40 genome designated EI. A minor set of transcripts mapped to the second promoter region, EII, at position 14 to 17 on the SV40 genome or nucleotides 30 to 32 with S1 nuclease mapping or primer extension, respectively. This anomaly in the two procedures is thought to reflect an effect of the A-T-rich sequence located at nucleotides 15 through 30. Identification of the actual 5' capped nucleotides will help explain this discrepancy. In both cases, other minor sites of initiation occurred in the region of nucleotides 56 to 65. In addition, a previously undetected initiation event at approximately nu-

cleotide 235 (corresponding to the end of the second 72-base pair repeat) was observed in the S1 analysis. This initiation event occurred in the region that is required for enhancer function but is not generally thought to exhibit promoter function. However, the assignment of promoter function in this region is complicated by the fact that the EII promoter is a complex unit which is active during viral replication but may not act continuously as a promoter.

We observe the simultaneous use of EI and EII in this system. The use of EII reflects the presence of functional T antigen in COS cells. This product is required for the use of EII (10) during viral infection. However, high levels of T antigen cause a substantial shutdown of the

early promoter region (autoregulation) late in viral infection. That we do not observe strong autoregulation effects on HBsAg mRNA levels may reflect the fact that COS cells produce relatively low levels of T antigen (9). This is an important feature in the use of SV40 early promoter-dependent constructs as expression vectors.

In contrast to the simple pattern obtained with LSV-HBsAg, Northern blotting and S1 nuclease mapping showed that the HBsAg-specific RNA transcribed from the LSV-HBpresAg is complex (see Fig. 2). A significant fraction of the RNA appeared to initiate upstream from the HBV sequences, presumably at the SV40 early promoter. A second major RNA species generated by the LSV-HBpresAg recombinant initiated at the HBpresAg promoter which has been detected in an *in vitro* transcription assay (Rall et al., in press). The other transcripts generated may reflect alternate initiation sites in the HBsAg presequence. However, an analysis of the nucleotide sequence in this region shows that there are putative splice acceptor sites in the HBsAg presequence (data not shown), so that RNA processing cannot presently be ruled out. The LSV-HBsAg recombinant produces about four-fold more stable HBsAg-coding mRNA than the LSV-HBpresAg recombinant which has both an efficient SV40 and an HBV promoter. Even though the long pre-surface antigen leader could attenuate both promoters, resulting in comparatively lower levels of transcription, it is also possible that the long pre-surface antigen region includes RNA processing signals which increase the rate of turnover of these transcripts.

COS cells infected with the LSV-HBsAg recombinant synthesize and secrete the characteristic mixture of glycosylated and nonglycosylated forms of HBsAg. Further, electron microscopic analysis showed that the secreted HBsAg particles coded by the LSV-HBsAg recombinant are similar in structure to the 22-nm particles detected in the serum of infected individuals. The amount of HBsAg secreted by COS cells infected with the LSV-HBsAg recombinant is substantial, about 10^7 HBsAg molecules per infected cell during the 24-h period before the initiation of cell lysis. This is about 10-fold more than that produced by a human hepatoma cell line (1) or in several HBV-SV40 late replacement recombinants studied elsewhere (16, 22; O. Laub, unpublished results).

The efficient production of HBsAg by the LSV-HBsAg recombinant indicates that the putative pre-surface antigen signal peptide preceding the coding region of mature HBsAg is not required for the assembly, secretion, or stability of the HBsAg particles. However, the conservation of the long open pre-surface antigen region

in three previously described HBV DNA sequences and the existence of an analogous region in the Woodchuck hepatitis virus genome (7, 8, 30) are strong arguments in favor of the translation and functionality of this region. We have recently mapped the HBsAg promoter *in vitro* and detected only one major initiation site at HBV nucleotide 2816 (Rall et al., in press) that precedes the pre-surface antigen region. In the present work, we have shown that this initiation site is used in the LSV-HBpresAg recombinant (additional data on this point are presented in Rall et al., in press). These results, together with the data from DNA transfection experiments (25), define a promoter before the start of the pre-surface antigen region. Thus, the initial HBsAg transcript should contain the long pre-surface antigen region, and the S1 studies shown in Fig. 7 demonstrate for the first time that this sequence is also present in cytoplasmic mRNA. In addition to this full-length mRNA, there are at least three minor mRNA species which map to the region between the pre-surface and surface AUG translational start codon. These species are not artifacts of the SV40 system since they are also seen in RNA from a cell line which expresses HBsAg after transfection with HBV DNA (unpublished results). As discussed above, these species may reflect discrete initiation events or RNA splicing. The pre-surface antigen region contains three in-frame ATG residues before the mature HBsAg initiation codon. At least two of the minor RNA species would contain one of these extra AUG codons (unpublished results) in addition to any that might be brought in by a splicing event. This raises the intriguing possibility that the pre-surface/surface gene might encode a number of protein products showing variation at the amino terminus. Such products could show altered cellular compartmentalization or could play a slightly different structural role compared with HBsAg within the HBV particles. The SV40 system described here should facilitate the search for and function of precursors or variants of the HBsAg protein.

LITERATURE CITED

1. Alexander, J. J., E. M. Bey, E. W. Geddes, and G. S. Letcasas. 1976. Establishment of a continuously growing cell line from primary carcinoma of the liver. *Afr. Med. J.* 50:2124-2129.
2. Alwine, J. C., D. J. Kemp, and G. R. Stark. 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc. Natl. Acad. Sci. U.S.A.* 74:5350-5354.
3. 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.
4. Cabral, G. A., F. Marciano-Cabral, G. A. Funk, Y. Sanchez, F. B. Hollinger, J. L. Melnick, and G. R.

- Dreesman. 1978. Cellular and humoral immunity in guinea pigs to two major polypeptides derived from hepatitis B surface antigen. *J. Gen. Virol.* **38**:339-350.
5. Clewell, D. B., and D. R. Helinski. 1969. Supercoiled circular DNA-protein complex in *E. coli*: purification and induced conversion to an open circular DNA form. *Proc. Natl. Acad. Sci. U.S.A.* **62**:1159-1166.
 6. Edman, J. C., P. Gray, P. Valenzuela, L. B. Rall, and W. J. Rutter. 1980. Integration of hepatitis B virus sequences and their expression in a human hepatoma cell. *Nature (London)* **286**:535-538.
 7. Galibert, F., T. N. Chen, and E. Mandart. 1982. Nucleotide sequence of a cloned woodchuck hepatitis virus: comparison with the hepatitis B virus sequence. *J. Virol.* **41**:51-65.
 8. 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.
 9. Gluzman, Y. 1981. SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**:175-182.
 10. Hansen, U., D. G. Tenen, D. M. Livingston, and P. A. Sharp. 1981. T antigen repression of SV40 early transcription from two promoters. *Cell* **27**:603-612.
 11. Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell culture. *J. Mol. Biol.* **26**:365-369.
 12. Hitzeman, R. A., C. Y. Chen, F. E. Hagie, E. J. Patzer, C. C. Liu, D. A. Estell, J. V. Miller, A. Yaffe, D. G. Kleid, A. D. Levinson, and H. Oppermann. 1983. Expression of hepatitis B virus surface antigen in yeast. *Nucleic Acids Res.* **11**:2745-2763.
 13. Laemmli, U. K. 1977. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature (London)* **227**:680-685.
 14. Laub, O. and Y. Aloni. 1975. Transcription of simian virus 40. V. Regulation of simian virus 40 gene expression. *J. Virol.* **16**:1171-1183.
 15. Laub, O. and W. J. Rutter. 1983. Expression of the human insulin gene and cDNA in a heterologous mammalian system. *J. Biol. Chem.* **258**:6043-6050.
 16. 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.
 17. Maniatis, T., R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quon, D. K. Sim and A. Efstratiadis. 1978. The isolation of structural genes from libraries of eucaryotic DNA. *Cell* **15**:687-701.
 18. Manley, J. L., A. Fire, A. Cano, P. A. Sharp, and M. L. Gelfer. 1980. DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. *Proc. Natl. Acad. Sci. U.S.A.* **77**:3855-3859.
 19. Martial, J. A., R. A. Hallowell, J. D. Baxter, and H. M. Goodman. 1979. Human growth factor: complementary DNA cloning and expression in bacteria. *Science* **205**:602-607.
 20. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base specific chemical cleavages. *Methods Enzymol.* **65**:499-560.
 21. McKnight, S. L., and E. R. Davis. 1980. Expression of the herpes thymidine kinase gene in *Xenopus laevis* oocytes: an assay for the study of deletion mutants constructed in vitro. *Nucleic Acids Res.* **8**:5931-5948.
 22. Moriarty, A. M., B. H. Hoyer, J. W. Shih, J. L. Gerin, and D. H. Hamer. 1981. Expression of the hepatitis B virus surface antigen gene in cell culture by using a simian virus 40 vector. *Proc. Natl. Acad. Sci. U.S.A.* **78**:2606-2610.
 23. Palmiter, R. D., H. Y. Chen, and R. L. Brinster. 1982. Differential regulation of metallothionein-thymidine kinase fusion genes in transgenic mice and their offspring. *Cell* **29**:701-710.
 24. Parker, B. A. and G. R. Stark. 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. *J. Virol.* **31**:360-369.
 25. 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.
 26. Radloff, R., W. Bauer, and J. Vinograd. 1967. A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: the closed circular DNA in HeLa cells. *Proc. Natl. Acad. Sci. U.S.A.* **57**:1514-1521.
 27. Summers, J., A. O'Connell, and I. Millman. 1975. Genome of hepatitis B virus: restriction enzyme cleavage and structure of DNA extracted from Dane particles. *Proc. Natl. Acad. Sci. U.S.A.* **72**:4597-4601.
 28. Valenzuela, P., P. Gray, M. Quiroga, J. Zaldivar, J. 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.
 29. Valenzuela, P., A. Medina, W. J. Rutter, G. Ammerer, and B. D. Hall. 1982. Synthesis and assembly of hepatitis B virus surface antigen in yeast. *Nature (London)* **298**:347-350.
 30. Valenzuela, P., M. Quiroga, J. Zaldivar, P. Gray, and W. J. Rutter. 1980. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes. p. 55-77. *In* B. N. Fields, R. Jaenisch, and C. F. Fox (ed.), *Animal virus genetics*. Academic Press, Inc., New York.
 31. Vyas, G. M., S. N. Cohen, and R. Schmid. (ed.) 1978. *Viral hepatitis: a contemporary assessment of etiology, epidemiology, pathogenesis and prevention*. p. 32. Franklin Institute Press, Philadelphia.
 32. Wang, Y., M. Schafer-Ridder, C. Stratowa, T. K. Wong, and P. H. Hofschneider. 1982. Expression of hepatitis B surface antigen in unselected cell culture transfected with recircularized HBV DNA. *EMBO J.* **1**:1213-1216.