JOURNAL OF VIROLOGY, Jan. 1987, p. 35–41
0022-538X/87/010035-07$02.00/0

Vol. 61, No. 1

5′-Terminal Sequences Influence the Segregation of Ground Squirrel Hepatitis Virus RNAs into Polyribosomes and Viral Core Particles

GREG H. ENDERS,[1] DON GANEM,[2,3] AND HAROLD E. VARMUS[1,2]*

Division of Genetics, Departments of Biochemistry and Biophysics,[1] Microbiology and Immunology,[2] and Medicine,[3] University of California, San Francisco, California 94143

Received 15 July 1986/Accepted 26 September 1986

To determine which of the major ground squirrel hepatitis virus RNAs serve as mRNAs and which serve as templates for reverse transcription of the genome, we analyzed the subcellular distribution of these RNAs in livers of infected ground squirrels. Both major classes of viral RNA, the 2.3- and 3.5-kilobase (kb) classes, are unspliced, are polyadenylated at a common position, and display heterogeneous 5′ ends that can encode proteins with different amino termini (G. H. Enders, D. Ganem, and H. Varmus, Cell 42:297–308, 1985). Both of the 2.3-kb RNAs, which encode surface antigens, appear to be predominantly associated with polyribosomes. Of the three 3.5-kb RNAs, the two longer, which can encode a protein initiated from the first methionine codon in the core antigen gene, appear to be predominantly associated with polyribosomes, and a minority of the shortest 3.5-kb RNAs, which can encode a protein initiated from the second methionine in the core antigen gene, appears to be associated with polyribosomes. This last RNA is instead found predominantly within viral core particles, consistent with evidence that indirectly implicates it in two steps of viral DNA synthesis (C. Seeger, D. Ganem, and H. E. Varmus, Science 232:477–484, 1986). None of the other viral RNAs is detectably packaged into cores. These findings provide independent evidence that the shortest 3.5-kb RNA is the template for synthesis of the viral genome and reveal a novel selectivity in viral RNA packaging.

The hepatitis B viruses (hepadnaviruses) are the smallest DNA viruses known to infect vertebrates. Viral replication occurs primarily in the liver and is associated with a range of pathology that encompasses the late development of hepatocellular carcinoma (2, 6). The hepadnaviruses synthesize their genomes from RNA templates, using a strategy that differs from the canonical reverse transcription scheme employed by retroviruses and several eucaryotic transposons (21, 23, 25). In this paper we present evidence distinguishing hepadnaviral RNAs that play a role in genome synthesis from those that function solely as mRNAs.

We recently determined the primary structures of the major viral RNAs present in the livers of ground squirrels infected by ground squirrel hepatitis virus (GSHV; 4). Both major classes of viral RNA, the 2.3- and 3.5-kilobase (kb) classes, are plus stranded, unspliced, and polyadenylated at a common position (Fig. 1). The 2.3-kb RNAs are subgenomic in size and display heterogeneous 5′ ends that bracket a methionine codon in the presurface reading frame. This apparently facilitates the expression of two surface antigens (sAgs), designated pre-S2 and S, that differ at their amino termini (4, 18, 22). The 3.5-kb RNAs are slightly longer than genome length and, thus, terminally redundant. Their 5′ ends are clustered at nucleotides −25, −16, and +6, relative to the first methionine codon in the core antigen (cAg) open reading frame. These positions suggest that the 3.5-kb RNAs may serve as mRNAs for cAgs encoded from the first (precore) and second (core) methionine codons in the cAg gene.

Because they include all of the information in viral DNA, the 3.5-kb RNAs are obvious candidates to serve also as templates for synthesis of the viral genome. As noted above, the multiple 5′ ends of the 3.5-kb RNAs confer on these RNAs the potential to direct the synthesis of different viral proteins. Does synthesis of these ends also yield RNAs with different potentials as replicative templates? Two pieces of evidence have emerged that suggest that the shortest 3.5-kb RNA may be the sole functional replicative template. (i) The 3′ end of virion minus-strand DNA is unique and maps to nucleotide +6 (21). This fact suggests that reverse transcription may proceed to the end of the template by using the shortest 3.5-kb RNA. (ii) A capped oligoribonucleotide is attached to the 5′ end of virion plus-strand DNA at position −200 and presumably primes synthesis of the DNA at that site (11, 21). The sequence of this RNA primer corresponds to the 5′-terminal 17 nucleotides of the shortest 3.5-kb RNA. This correspondence suggests that the primer may have been transported to the site of initiation of plus-strand DNA synthesis after cleavage from the 5′ end of the shortest 3.5-kb RNA (11, 21).

These observations, then, suggest that the shortest 3.5-kb RNA is used for synthesis of the DNA structures found in mature virions. It is, however, possible that all three 3.5-kb RNAs are packaged into viral core particles but that only the shortest can be utilized by the replicative machinery or that various DNA forms are synthesized in intracellular core particles with RNAs of different structure but that only cores with the described DNA structure exit the cell.

We sought to clarify these issues by directly examining the disposition of RNAs in cytoplasmic lysates of GSHV-infected hepatocytes to determine which GSHV RNA species are associated with polyribosomes and which are found in core particles (the site of reverse transcription [13, 23]). We present evidence here that most of the 2.3-kb and the two longer 3.5-kb RNAs and a minor fraction of the shortest 3.5-kb RNAs appear to be associated with polyribosomes. This finding provides further evidence that both precore and core proteins are likely to be synthesized in GSHV-infected livers (4). The shortest 3.5-kb RNA is found predominantly within viral core particles. We estimate that the ratio of the

* Corresponding author.

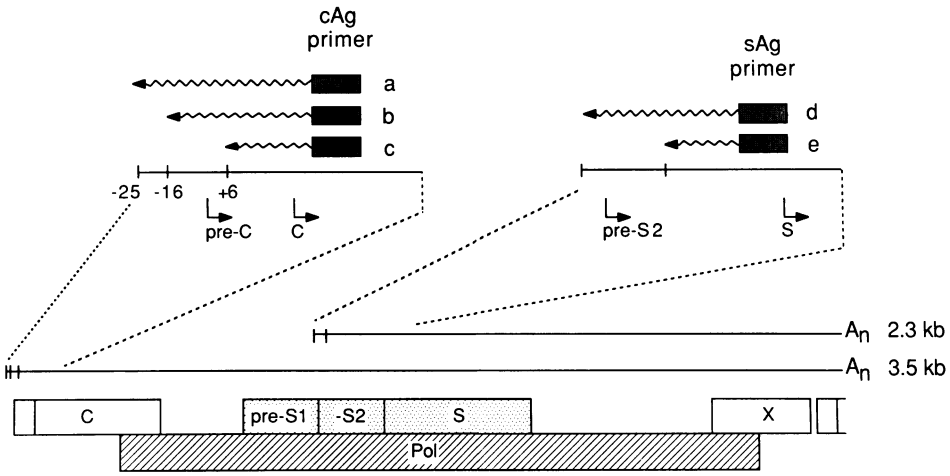

FIG. 1. The major GSHV RNAs and the primers used for primer extension analysis. The two major classes of GSHV RNA are represented by two lines, with short vertical lines marking the positions of the major 5′ ends and A_n designating the poly(A) tails. Depicted below them are the viral long open reading frames (20). Depicted above are enlargements of the regions of the RNAs examined by primer extension analysis. The primers (■) and extension products (〰) are not to scale. The cAg primer is composed of minus-strand nucleotides +167 to +148. The sAg primer is composed of minus-strand nucleotides +1643 to +1624. a, b, and c, cAg primer extension products 192, 183, and 162 (±3) nucleotides long, respectively, that terminate at the three major 5′ ends of the 3.5-kb RNAs, at nucleotides −25, −16, and +6, respectively (4); d and e, sAg primer extension products 158 and 119 (±3) nucleotides long, respectively, that terminate at the two major 5′ ends of the 2.3-kb RNAs (4); ⌊→, translation initiation sites of the putative precore (pre-C) and core (C) proteins (at +1 and +91, respectively) and the pre-S2 and major sAgs (S).

shortest 3.5-kb RNA to all other known viral RNAs is at least 100 to 1 in cores. This finding provides independent evidence that the shortest GSHV 3.5-kb RNA is the template for synthesis of the viral genome and reveals a remarkable selectivity in packaging of viral RNA, in which the extreme 5′ end is a critical feature.

MATERIALS AND METHODS

Animals. Beechey ground squirrels were trapped in the wild in Palo Alto, Calif. sAg in serum was assayed by the heterologous solid-phase radioimmunoassay for human hepatitis B sAg (Ausria 11; Abbott Laboratories, North Chicago, Ill.) according to the specifications of the manufacturer. Viral DNA in serum was assayed by the dot-blot method (24) after proteinase K digestion (400 μg/ml at 50°C for 4 h) and phenol extraction. The animal used for this study tested positive for sAg in serum when trapped and contained both sAg and viral DNA in its serum when sacrificed, 4 months later. The second animal, from which data not shown here were derived, was infected by direct liver inoculation with a cloned GSHV genome bearing a G-to-C transversion at nucleotide +18 (21). The animal was sacrificed after remaining positive for sAg in serum for 4 months.

Subcellular fractionation. A freshly excised liver was divided into three portions. One portion was immediately frozen at −70°C and used later for RNA isolation by the guanidium isothiocyanate procedure as described previously (4). A second portion was homogenized at 4°C by 2 strokes of a motor-driven Teflon pestle in 2 ml of a solution containing 10 mM Tris hydrochloride (pH 8.0), 0.15 M KCl, 1.5 mM $MgCl_2$, and 2 mM vanadate complexes per g and 100 μg of cycloheximide, included to inhibit ribosomal runoff, per ml. Nuclei and cell debris were removed by centrifugation at 10,600 × *g* at 4°C for 10 min. The supernatant was brought to 1% Nonidet P-40 by the addition of 20% Nonidet P-40, generating the solution referred to as the cytoplasmic lysate.

The sucrose gradient sedimentation protocol was adopted, in modified form, from Katze et al. (10). Lysate (1.3 ml) was loaded onto linear 10 to 50% sucrose gradients containing 20 mM Tris hydrochloride (pH 8.0), 50 mM KCl, and either 5 mM magnesium acetate or 10 mM EDTA (all solutions except the Tris hydrochloride were treated with diethylpyrocarbonate). Centrifugation was carried out at 27,000 rpm (12,400 × *g*) at 4°C for 4 h in an SW41 rotor. Fractions were collected from the top with an Auto Densi-Flo 11-C probe and a monostatic pump (both from Buchler Instruments Div., Nuclear-Chicago Corp., Fort Lee, N.J.). The pellet was suspended in gradient buffer without sucrose. Fractions from the gradient containing Mg^{2+} were brought to 20 mM EDTA by the addition of 0.5 M EDTA, and all fractions were brought to 1% sodium dodecyl sulfate (SDS) by the addition of 20% SDS immediately before the isolation of nucleic acids.

RNA was isolated by extraction with phenol-chloroform-isoamyl alcohol (25:24:1) that had been equilibrated with 0.1 M Tris hydrochloride (pH 8.0). Extractions were repeated until the phenol-water interface cleared, with additions of 10 mM Tris hydrochloride (pH 8.0)–1 mM EDTA as needed to maintain the volume of the aqueous phase. The nucleic acids were then precipitated twice with sodium acetate and ethanol. To degrade RNases in the lysate prior to RNA isolation or to isolate DNA, samples were pretreated with proteinase K (proteolysis removes the protein that is covalently attached to the 5′ end of viral minus-strand DNA, allowing retention in the aqueous phase of minus strands and any plus strands that may be extensively hydrogen bonded to them (7, 8). Yeast tRNA (5 μg) was added as a carrier, and the samples were brought to 10 mM vanadate complexes and 200 μg of proteinase K per ml and were incubated at 37°C for 30 min. The samples were then brought to 1% SDS by adding 20% SDS and 500 μg of proteinase K per ml and were incubated at 37°C for 1 h (14).

Equilibrium density centrifugation was performed as described by Feitelson et al. (5) except that the cesium was buffered with 10 mM Tris hydrochloride (pH 7.5), and the samples were centrifuged at 24,500 rpm (10,000 × *g*) at 10°C

for 70 h in an SW41 rotor. Lysate (1.3 ml) was adjusted to 1.19 g/ml with CsCl and loaded onto a 10-ml CsCl step gradient from 1.25 to 1.40 g/ml. Fractions were collected as described above. The pellet was suspended in gradient buffer. The density of each fraction was determined by weighing duplicate 50-μl aliquots (12 fractions were collected; after their densities were determined, the fractions were pooled by twos). The fractions were then dialyzed exhaustively at 4°C against 10 mM Tris hydrochloride (pH 8.0)–0.15 M NaCl, and nucleic acids were isolated as described above.

The third portion of the liver was used for puromycin treatment, as described by Adelman et al. (1), followed by sucrose gradient sedimentation.

Nucleic acid analysis. Primer extensions and agarose gel electrophoresis were performed as described previously (4). DNA dot-blots of extract-derived nucleic acids were performed as described by Thomas (24). RNA was hydrolyzed by incubating each sample, plus a control sample containing 1 ng of GSHV RNA synthesized in vitro with the SP6 polymerase, in 0.2 M NaOH at 65°C for 20 min. For analysis of DNA from the second animal, the nucleic acids were applied to nitrocellulose without prior alkaline hydrolysis and allowed to hybridize with a probe specific for viral minus strands.

RESULTS

Experimental design. Cytoplasmic lysates were prepared from the livers of ground squirrels chronically infected with GSHV and subjected to rate zonal sedimentation in sucrose, nuclease digestion, and equilibrium density centrifugation in CsCl. After each of these procedures, nucleic acids were isolated and tested for the presence of viral RNA and DNA species. Because the major GSHV RNAs differ in the positions of their 5′ ends, they are distinguishable by primer extension analysis (Fig. 1). Extension of a synthetic primer composed of minus-strand nucleotides +167 to +148 (the cAg primer) on RNA from infected liver yielded major products of 192, 183, and 162 (±3) nucleotides (labeled a, b, and c in the figures), corresponding to the three major 5′ ends of the 3.5-kb RNAs (4). Extension of a synthetic primer composed of minus-strand nucleotides +1643 to +1624 (the sAg primer) yielded major products of 158 and 119 (±3) nucleotides (labeled d and e), corresponding to the two major 5′ ends of the sAg RNAs. We have not detected RNAs in infected liver with ends outside these two regions of the genome (4). Furthermore, the use of 5′-proximal primers to identify the RNAs should yield signals that are relatively unaffected by moderate amounts of nonspecific degradation that may occur during fractionation or by degradation of the body of the RNAs that may occur during their use as replicative templates (23). (The cAg primer anneals to sequences just outside of the terminal redundancies in the 3.5-kb RNAs; hence, it anneals only to the 5′ ends of these RNAs.) All the data presented in this paper are derived from experiments performed in parallel on material from a single animal; where noted below, similar experiments were performed on material from a second animal, with qualitatively similar results.

Sucrose gradient sedimentation. As a first step toward distinguishing the RNAs in intracellular viral core particles from those associated with polyribosomes and other ribonucleoprotein particles, we subjected a Nonidet P-40-treated cytoplasmic lysate to sedimentation on linear 10 to 40% sucrose gradients in the presence and absence of

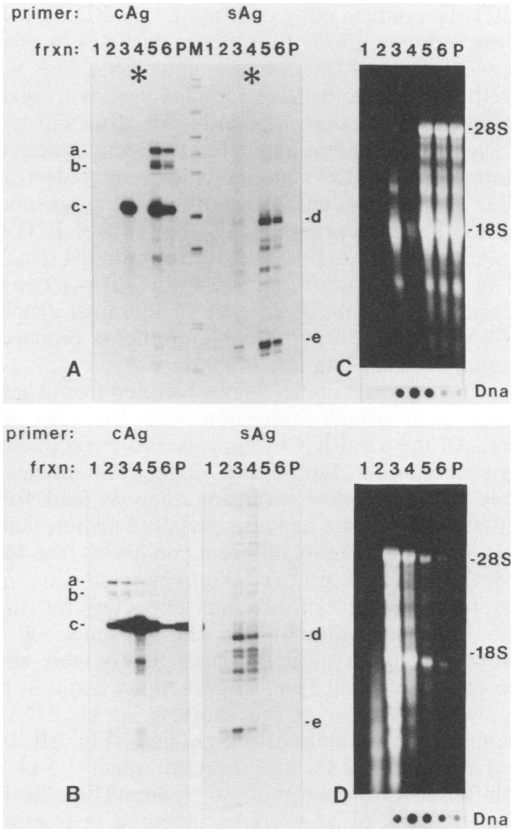

FIG. 2. Sucrose gradient sedimentation of cytoplasmic lysates. A cytoplasmic lysate prepared from a GSHV-infected liver was layered onto linear 10 to 50% sucrose gradients containing either Mg^{2+} (A and C) or EDTA (B and D). Six fractions were collected, beginning at the top of each gradient, plus the pellet (P). Nucleic acids were isolated from each fraction, and equal portions were subjected to three treatments: (i) primer extension analysis with the cAg and sAg primers (Fig. 1), followed by electrophoresis of the products through a 6% polyacrylamide-urea gel (A and B); (ii) electrophoresis in an agarose gel containing ethidium bromide (C and D, top); or (iii) alkaline hydrolysis, followed by dotting onto nitrocellulose and hybridization with a probe containing the GSHV DNA genome cloned in a derivative of pBR322 (C and D, bottom [DNA]). The nucleic acids isolated for primer extension analysis and agarose gel electrophoresis from fraction 4 of the gradient containing Mg^{2+} (asterisks) were lost. However, nucleic acids were also isolated, with prior proteinase K treatment, for the DNA dot-blot analysis presented in C, and a small portion of this sample was included in the ethidium-stained gel. M, Msp fragments of pBR322 (molecular weights, 110, 123, 147, 160, 180, 190, 201, 217, 238, and 242); a, b, c, d, and e, primer extension products diagrammed in Fig. 1. Some sAg primer extension products likely resulted from the primer annealing to fragments of 3.5-kb RNAs.

EDTA. EDTA dissociates polyribosomes, whereas GSHV core particles remain intact (13). We divided each gradient into six fractions, resuspended the material that pelleted during sedimentation, and isolated nucleic acids by extraction with phenol and precipitation with ethanol. To confirm that EDTA had achieved the desired disruption of polyribosomes, we subjected nucleic acid from each fraction to electrophoresis in an agarose gel in the presence of formaldehyde and stained the gel with ethidium bromide. The results demonstrate the expected shift of 28S and 18S rRNAs toward the top of the gradient containing EDTA (Fig. 2D), relative to their positions in the gradient containing Mg^{2+}

(Fig. 2C). To confirm our expectation that EDTA would not affect the sedimentation of viral core particles in this experiment, we examined fractions for viral DNA (see Materials and Methods). RNA in these samples was hydrolyzed with base, and the DNA was applied in dots to a nitrocellulose filter. The filter was then incubated with a radioactive probe containing the GSHV genome cloned in a derivative of pBR322. The results demonstrate that, in the presence (Fig. 2D, bottom) or absence (Fig. 2C, bottom) of EDTA, viral DNA sedimented to the middle of the gradients (fractions 3, 4, and 5), the region expected from previous work to contain viral cores (23; unpublished results). Identical distributions of DNA were seen in comparable gradients prepared from the second animal (data not shown).

We then subjected nucleic acids isolated from each gradient fraction to primer extension with the cAg and sAg primers. All the viral RNAs appeared to cosediment in the gradient containing Mg^{2+} (Fig. 2A). (The nucleic acids prepared for this primer extension analysis from fraction 4 were lost; however, we had also prepared nucleic acids from this fraction under slightly different conditions [see Materials and Methods], and primer extensions on this material yielded the amounts of product expected for the peak fraction [data not shown].) In the presence of EDTA, however, most of two longer 3.5-kb RNAs were shifted to the top of the gradient (Fig. 2B, fractions 2 and 3, bands a and b), whereas most of the shortest 3.5-kb RNAs again sedimented to the middle of the gradient (Fig. 2B, band c). Most of the sAg RNAs, like the two longer 3.5-kb RNAs, were shifted toward the top of the gradient (Fig. 2B, bands d and e). The peak of sAg RNAs appeared in fraction 3, as opposed to fraction 2 for the longer 3.5-kb RNAs, for reasons that are unclear. Results similar to those depicted in Fig. 2A and B were also obtained with material from the second animal (data not shown).

Taken together, these data suggested that the RNAs shifted toward the top of the gradient containing EDTA, i.e., most of the 2.3-kb and the two longer 3.5-kb RNAs but only a small fraction of the shortest 3.5-kb RNA, were released from structures, presumably largely polyribosomes, that sediment farther in the gradient containing Mg^{2+}. To be certain that the differences in recovery of the extension products using the cAg primer were not due to the variable copurification of a diffusible inhibitor in the different samples, we mixed samples from fractions 1 and 4 of the gradient containing EDTA (Fig. 2B) and repeated the primer extension reaction. The recovery of products from this reaction was not inhibited by the sample from fraction 1, thereby ruling out this potential artifact (data not shown). We then serially diluted the products from fractions 2 and 4 (Fig. 2B) prior to electrophoresis to quantify the ratios of the 5′ ends in these samples. The results demonstrate that the shortest 3.5-kb RNA is at least 15-fold enriched relative to the two longer 3.5-kb RNAs in fraction 4 relative to fraction 2 (data not shown).

RNase sensitivity. From these data, we predicted that the sAg and the two longer 3.5-kb RNAs would be preferentially sensitive to digestion by endogenous and exogenous RNases, whereas the shortest 3.5-kb RNA would be relatively resistant. Strong indirect evidence has been obtained previously that RNA in hepadnavirus core particles is resistant to digestion with 0.1 to 0.5 μg of RNase A per ml (13, 23). We therefore isolated nucleic acids from the total lysate with and without prior incubation with proteinase K and exposure to increasing amounts of RNase A. Nucleic acids isolated from the lysate simply by phenol extraction in the presence

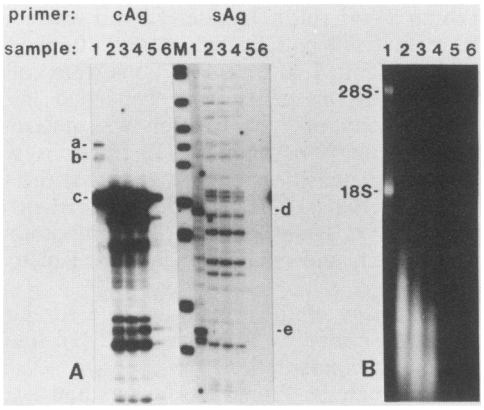

FIG. 3. RNase sensitivity of GSHV RNAs in the cytoplasmic lysate. Six samples taken from the cytoplasmic lysate were used for isolation of nucleic acids by phenol-chloroform extraction. Sample 1, no prior treatment; sample 2, prior proteinase K digestion; samples 3 through 6, prior incubation with 0.05, 0.5, 5, and 25 μg of RNase A per ml, respectively, followed by proteinase K digestion. Isolated nucleic acids were subjected to primer extension analysis with the cAg and sAg primers (A) or were electrophoresed through an agarose gel containing ethidium bromide (B). M, Msp fragments of pBR322; a, b, c, d, and e, primer extension products diagrammed in Fig. 1.

of 1% SDS followed by ethanol precipitation yielded primer extension products corresponding to all of the major viral RNAs (Fig. 3A, sample 1). The rRNAs in this preparation appeared largely intact (Fig. 3B, sample 1). After incubation of the lysate at 37°C for 30 min with proteinase K and then for another 90 min with proteinase K and 1% SDS, recovery of the sAg and the two longer 3.5-kb RNAs was virtually abolished, whereas recovery of the shortest 3.5-kb RNA was actually enhanced ca. threefold (Fig. 3A, sample 2 and other exposures not shown). The loss of the sAg and the two longer 3.5-kb RNAs was due to endogenous RNase activity in the lysate; under these conditions, there was substantial degradation of rRNA (Fig. 3B, sample 2), despite the presence of carrier tRNA and concentrations of vanadate complex RNase inhibitor higher than those included in the sucrose gradients. (The degradation was not due to RNase activity in the proteinase K, because identical treatments of both RNA synthesized in vitro and material from the middle of the sucrose gradient containing Mg^{2+} resulted in undiminished recovery of all the 5′ ends [data not shown].) Note that the primer extension products migrating more rapidly than the major products in Fig. 3 are absent in lane 1 and were not observed after identical treatment of fractions from both the sucrose and cesium gradients (data not shown), indicating that the corresponding RNAs were generated in vitro in this experiment.

These data indicate that the shortest 3.5-kb RNA is preferentially resistant to nuclease attack. We found further that recovery of the shortest 3.5-kb RNA was only slightly diminished by prior incubations with 0.05 and 0.5 μg of RNase per ml (Fig. 3A, samples 3 and 4 and other exposures not shown). Substantial loss of the shortest 3.5-kb RNA was not seen until the concentration of RNase A reached 5 μg/ml, a concentration at which we were no longer able to detect RNA by ethidium bromide staining (Fig. 3A and B, sample 5). To confirm these findings, we treated a portion of fraction 5 from the sucrose gradient containing Mg^{2+} (Fig. 2A) with 0.1 μg of RNase A per ml. This treatment resulted

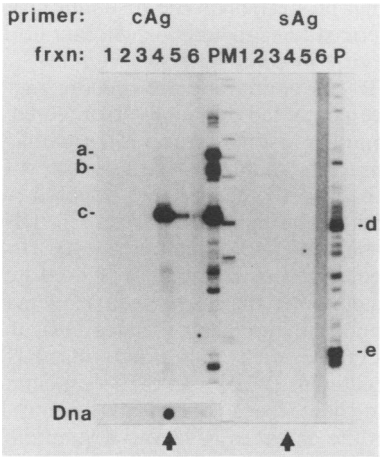

FIG. 4. Equilibrium density sedimentation of the cytoplasmic lysate. The cytoplasmic lysate was loaded onto a cesium chloride step gradient from 1.25 to 1.40 g/ml and subjected to equilibrium density sedimentation. Six fractions were collected, beginning at the top of the gradient, plus the pellet (P). After dialysis, nucleic acids were isolated and subjected to primer extension with the cAg and sAg primers. A portion of the nucleic acids from each fraction was also subjected to alkaline hydrolysis, followed by dotting onto nitrocellulose and hybridization with a probe containing the GSHV DNA genome cloned in a derivative of pBR322 (bottom [DNA]). M, Msp fragments of pBR322; a, b, c, d, and e, primer extension products diagrammed in Fig. 1; arrows, fraction with a density of 1.31 to 1.36 g/ml, where GSHV core particles are expected to band (5).

in the preferential loss of the sAg and the two longer 3.5-kb RNAs (data not shown).

Equilibrium density sedimentation. To provide further evidence that the shortest 3.5-kb RNA is in particles with the properties expected of viral cores, we subjected the cytoplasmic lysate to equilibrium density sedimentation in cesium chloride. GSHV core particles are known to band in cesium gradients at a density of 1.34 g/ml (5), whereas free RNA, with a density of ca. 1.7 g/ml, pellets in such gradients. We divided the gradient into six fractions and suspended the pellet; after dialysis to remove the cesium, we isolated nucleic acids and performed primer extension analyses as outlined above for the sucrose gradients. The results demonstrate that, whereas all the viral RNA species were recovered in the pellet fraction, the only viral RNA in the gradient was the shortest 3.5-kb RNA (Fig. 4). This RNA displayed a sharp peak in signal in the fraction with a density range from 1.31 to 1.36 (Fig. 4, fraction 4 [arrow]). A DNA dot-blot, performed as described for Fig. 2, revealed a coincident peak of viral DNA (Fig. 4, bottom, fraction 4 [arrow on left]).

We sought evidence that the viral DNA in fraction 4 was characteristic of the forms expected for viral core particles. We used primer extension to examine nucleic acids isolated from the gradient after treatment with proteinase K; proteolysis permits recovery of nucleic acids that are extensively hydrogen bonded to protein-linked viral minus-strand DNA (7, 8). Use of the cAg primer with these nucleic acids yielded a sharp peak in fraction 4 of an extension product of 370 nucleotides (data not shown). This band corresponds precisely to that expected from extension of the cAg primer to the 5′ end of plus-strand DNA (including its terminal oligoribonucleotide). In addition, prior proteinase K treatment increased recovery of the shortest 3.5-kb RNA from fraction 4 about twofold relative to recovery of this RNA from the pellet. These data demonstrate that the shortest 3.5-kb RNA cosediments on cesium density gradients with viral core particles and suggest that some of this RNA may be hydrogen bonded to protein-linked viral minus strands. Equilibrium density sedimentation of a lysate from the second animal yielded similar results.

DISCUSSION

Evidence for selective packaging. We have presented evidence that, of the major viral RNAs present in the livers of GSHV-infected ground squirrels, only the shortest 3.5-kb RNA is efficiently packaged into viral core particles. Thus, viral RNAs with only 21 additional or 1,480 fewer nucleotides at their 5′ termini are not detectably packaged. These conclusions were derived by determining which RNAs were associated with viral core particles by three independent criteria. We also examined the ability of the viral RNAs to sediment through sucrose after treatment with puromycin under conditions known to dissociate polyribosomes (1). We found that the sedimentation of the shortest 3.5-kb RNA was preferentially resistant to such treatment (data not shown).

Because the major GSHV RNAs differ in their subcellular localizations, the relative recovery of these RNAs can differ among isolation procedures. For example, hepadnaviral cores are known to be somewhat unstable in cesium (16), and we saw an overall decrease in yield of shortest 3.5-kb RNAs from the cesium gradients relative to the yields achieved by other isolation procedures (compare Fig. 4 to Fig. 2 and 3; data not shown). The relative yield of RNA from fraction 4 and from the pellet from the cesium gradient (Fig. 4), therefore, probably underestimates the ratio of RNA in cores to RNA associated with other structures. Conversely, disruption of polyribosomes with puromycin reduced the overall recovery of polyribosome-associated RNAs (data not shown). We found, however, that the ratios of the RNA ends recovered directly from the lysate or from the sucrose gradients were similar to the ratios of ends in RNA recovered by the guanidinium isothiocyanate procedure from another portion of the same liver (data not shown). This finding indicates that these latter procedures did not result in significant preferential loss of RNAs.

To be able to assess the relative abundances of the 2.3- and 3.5-kb RNAs, we compared the efficiency of extension of the cAg and sAg primers by using an RNA preparation with which we had independently established the ratio of these two RNA classes (4). We found that extension of the two primers was about equally efficient (data not shown).

These findings allow us to make some quantitative statements about the approximate relative subcellular distribution of the major viral RNAs. At steady state, most of the sAg and the two longer 3.5-kb RNAs are associated with polyribosomes. In contrast, most of the shortest 3.5-kb RNA, which is more abundant than either of the former sets of RNAs, is in core particles. We could not detect bands corresponding to sAg or the two longer 3.5-kb RNAs in primer extensions from the fraction of the cesium gradient containing cores, even with overexposures of the gel presented in Fig. 4. We estimate, therefore, that the ratio of the shortest 3.5-kb RNA to all other known viral RNAs is at least 100 to 1 in cores.

Two earlier pieces of evidence had indirectly implicated the shortest 3.5-kb RNA as the replicative template. (i) The 3′ end of virion minus-strand DNA is unique, and it is the

precise complement (±3 nucleotides) of the 5′ end of the shortest 3.5-kb RNA (21). (ii) The sequence of the oligoribonucleotide attached to the 5′ end of virion plus-strand DNA matches, within a couple bases of the ends, that of the 5′-terminal 17 nucleotides of the shortest 3.5-kb RNA (21). A similar correspondence in sequence has been detected between an oligoribonucleotide attached to the duck hepatitis B virus (DHBV) plus-strand DNA and the 5′ end of a greater-than-genome-length DHBV RNA (11). The data presented in this paper provide independent evidence that the GSHV 3.5-kb RNA with its 5′ end at +6 is the sole template employed in synthesis of the viral genome. The homogeneity of the relevant features of the virion DNA can thus be explained by selective RNA packaging, without invoking selection at other steps of viral maturation.

We anticipated that our primer extension assays might reveal novel RNA 5′ ends generated by cleavage of the packaged RNAs. This could occur if the plus-strand primer was cleaved from viral RNA prior to degradation of 3′ sequences. In addition, synthesis of minus-strand DNA apparently begins at nucleotide +15, which is present at both ends of the template (21). If initiation occurs at the 5′ end of the template, it could be accompanied by RNase H degradation of the extreme 5′ terminus of the RNA. We did not observe such cleavages, which suggests that, if they occur, the resultant RNAs must be short-lived.

Our results do not rule out processive degradation of the RNAs from a site 3′ of the cAg primer. Consistent with this possibility is our observation that the 5′ end of the shortest 3.5-kb RNA is enriched at least twofold, relative to the other RNA ends, in the nonpolyadenylated fraction of liver RNA (data not shown). Our finding that prior proteinase K treatment of the lysate (Fig. 3) or isolated core particles (data not shown) preferentially increased recovery of the 5′ end of the shortest 3.5-kb RNA two- to threefold, relative to that achieved by SDS and phenol alone, suggests that the majority of the packaged RNAs may be associated with protein-linked viral minus strands and hence undergoing reverse transcription.

Messenger function. Although most of the shortest 3.5-kb RNAs are packaged into core particles, the data suggest that a fraction of these RNAs may be associated with polyribosomes (compare fraction 2 in Fig. 2A and B). In contrast, most of the two longer 3.5-kb RNAs were associated with polyribosomes, as judged by their release from rapidly sedimenting structures upon treatment with EDTA (Fig. 2) and puromycin (data not shown). Furthermore, the exclusion of the two longer 3.5-kb RNAs from core particles suggests that their function is confined to translation. If the first in-frame methionine codon in each of the three 3.5-kb RNAs is used for translation initiation, then both precore and core proteins are presumably synthesized in the infected liver.

The amino termini of the cAgs produced in the infected liver have not been identified for any of the hepadnaviruses. The homologous human hepatitis B virus (HBV) precore protein contains a signal peptide that can be cleaved after synthesis of the protein in vitro (P. Garcia, J.-H. Ou, W. J. Rutter, and P. Walter, personal communication), and constructs expressing this protein in transfected cells result in the secretion of immunoreactive products of the cAg frame (17, 19). The homologous HBV core protein synthesized under the same in vitro conditions is not detectably cleaved (Garcia et al., personal communication), and its immunoreactive products remain largely within the transfected cell (17, 19). These data indicate that the precore region confers distinct properties on the resultant protein. We have initiated genetic studies to address whether both proteins are required for viral infectivity.

Thus far, RNAs containing the precore region at their 5′ ends have been detected only in livers infected by GSHV (4) and woodchuck hepatitis virus (15; unpublished results), whereas transcripts analogous to the shortest GSHV 3.5-kb RNAs have been detected in livers infected by woodchuck hepatitis virus (15; unpublished results), DHBV (3), and HBV (H. Will, personal communication). However, given that the precore reading frame is preserved in all infectious hepadnaviral genomes that have been sequenced and is well conserved among mammalian isolates (17), it seems likely that the precore region is expressed via mRNAs that have yet to be detected in DHBV- and HBV-infected livers. Such mRNAs, like their GSHV counterparts, may be more susceptible than the detected RNAs to degradation by endogenous RNases and, thus, may be preferentially lost during isolation.

Packaging models. The GSHV sAg RNAs, like retroviral *env* mRNAs, appear not to be efficiently packaged into core particles, presumably because they both lack the *cis*-acting sequences required for packaging (12, 26). This explanation cannot, however, account for the absence of the longer 3.5-kb RNAs from core particles, because these RNAs possess all of the primary sequences present in the packaged RNA. Several potential models to explain this observation can be envisioned. The models we offer invoke (i) recognition of a 5′-terminal sequence in conjunction with a cap structure, (ii) recognition of a 5′-proximal secondary structure, (iii) *cis*-packaging, and (iv) translational interference.

(i) Sequence plus cap. The packaging apparatus may recognize the precise sequence of the 5′ end of the shortest 3.5-kb RNA, perhaps in conjunction with the cap structure putatively present on this RNA. The apparatus performing this recognition could be that which chooses the site of initiation of reverse transcription. In this model, formation of the initiation complex for minus-strand DNA synthesis dictates packaging and can occur only on the 5′ end of the shortest 3.5-kb RNA. Polymerization on the 5′ end of the template would then yield an AT-rich deoxyribonucleotide that could be transferred to the 3′ end of the template to allow completion of minus-strand DNA synthesis (4, 21). Alternatively, the recognition apparatus could be one that also recognizes the oligoribonucleotide primer for cleavage or transfer of the primer to the site of initiation of plus-strand DNA synthesis (11, 21).

(ii) Secondary structure. The additional bases found on the longer 3.5-kb RNAs may alter the secondary structure of the 5′ end of the RNA in a manner that prevents recognition of a packaging signal by *trans*-acting factors. We have examined predicted secondary structures for these 5′ ends but can detect no obvious basis for such an effect.

(iii) *Cis* packaging. In the *cis* packaging model, each RNA that is packaged is recognized by a protein translated from the same molecule. The precore protein would be incapable of initiating the packaging of its mRNA, perhaps as a result of translocation of the protein into the endoplasmic reticulum. The core protein, synthesized only from the shortest 3.5-kb RNA, would be able to initiate packaging, perhaps by virtue of remaining cytoplasmic. Not every translation event would result in packaging of the RNA, allowing a number of core proteins to be made from a single RNA prior to packaging of the RNA into a core particle. In a variation of this model, a core-pol fusion protein, putatively generated by ribosomal frameshifting near the carboxy

terminus of core (4, 9), would initiate the packaging of its mRNA.

(iv) Translational interference. The 5′ ends of the longer 3.5-kb RNAs may be bound by translation machinery in a manner that renders a packaging signal in the precore region inaccessible to packaging factors. Similarly, translation of the precore sequence may target the RNAs to a location in the cell where packaging does not occur efficiently.

These models have different implications for the potential of hepadnaviruses to package cellular RNAs and to support the replication of defective genomes. Discrimination between the models will likely require analysis of the behavior of viral mutants in tissue culture cells that permit normal packaging; partially defective phenotypes could be then examined biochemically.

ACKNOWLEDGMENTS

We thank Dan Trueba for help with the animal work, Robin Colgrove for RNA secondary-structure predictions, Peter Walter for advice on characterization of polyribosome-associated RNA, Titia de Lange and David Persing for comments on the manuscript, and members of the Rutter and Walter laboratories for communication of unpublished results.

This work was supported by Public Health Service grants from the National Institutes of Health. G.H.E. is a Medical Scientist Training Program trainee, award GM07618; D.G. is a fellow of the John L. and George H. Hartford Foundation; and H.E.V. is an American Cancer Society Research Professor.

LITERATURE CITED

1. **Adelman, M. R., D. D. Sabatini, and G. Blobel.** 1973. Ribosome-membrane interaction. J. Cell. Biol. **56**:206–229.
2. **Beasley, R. P., C.-C. Lin, L.-Y. 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. **Buscher, M., W. Reiser, H. Will, and H. Schaller.** 1985. Transcripts and the putative RNA pregenome of duck hepatitis B virus: implications for reverse transcription. Cell **40**:717–724.
4. **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.
5. **Feitelson, M. A., P. L. Marion, and W. S. Robinson.** 1982. Core particles of hepatitis B virus and ground squirrel hepatitis B virus. I. Relationship between hepatitis B core antigen- and ground squirrel hepatitis core antigen-associated polypeptides by sodium dodecyl sulfate-polyacrylamide gel electrophoresis and tryptic peptide mapping. J. Virol. **43**:687–696.
6. **Ganem, D.** 1982. Persistent infection of humans with hepatitis B virus: mechanisms and consequences. Rev. Infect. Dis. **4**:1026–1047.
7. **Ganem, D., L. Greenbaum, and H. E. Varmus.** 1982. Virion DNA of ground squirrel hepatitis virus: structural analysis and molecular cloning. J. Virol. **44**:374–383.
8. **Gerlich, W. H., and W. S. Robinson.** 1980. Hepatitis B virus contains protein attached to the 5′ terminus of its complete strand. Cell **21**:801–809.
9. **Jacks, T., and H. E. Varmus.** 1985. Expression of the Rous sarcoma virus *pol* gene by ribosomal frameshifting. Science **230**:1237–1242.
10. **Katze, M. G., Y.-T. Chen, and R. M. Krug.** 1984. Nuclear-cytoplasmic transport and VA1-independent translation of influenza viral messenger RNAs in late adenovirus-infected cells. Cell **37**:483–490.
11. **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.
12. **Mann, R., and D. Baltimore.** 1985. Varying the position of a retrovirus packaging sequence results in the encapsidation of both unspliced and spliced RNAs. J. Virol. **54**:401–407.
13. **Miller, R. H., P. L. Marion, and W. S. Robinson.** 1984. Hepatitis B viral DNA-RNA hybrid molecules in particles from infected liver are converted to viral DNA molecules during an endogenous DNA polymerase reaction. Virology **139**:64–72.
14. **Miller, R. H., C.-T. Tran, and W. S. Robinson.** 1984. Hepatitis B virus particles from plasma and liver contain viral DNA-RNA hybrid molecules. Virology **139**:53–63.
15. **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.
16. **Ohori, H., M. Tamaki, 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.
17. **Ou, J.-H., O. Laub, and W. J. Rutter.** 1986. Hepatitis B 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.
18. **Persing, D. H., H. E. Varmus, and D. Ganemo.** 1986. Antibodies to pre-S and X determinants arise during natural infection with ground squirrel hepatitis virus. J. Virol. **60**:177–184.
19. **Roossinck, M. J., S. Jameel, S. H. Loukin, and A. Siddiqui.** 1986. Expression of hepatitis B viral core region in mammalian cells. Mol. Cell. Biol. **6**:1393–1400.
20. **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.
21. **Seeger, C., D. Ganem, and H. E. Varmus.** 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science **232**:477–484.
22. **Standring, D. N., W. J. Rutter, H. E. Varmus, and D. Ganem.** 1984. Transcription of the hepatitis B surface antigen gene in cultured murine cells initiates within the presurface region. J. Virol. **50**:563–571.
23. **Summers, J., and W. S. Mason.** 1982. Replication of the genome of a hepatitis B virus by reverse transcription of an RNA intermediate. Cell **29**:403–415.
24. **Thomas, P. S.** 1980. Hybridization of denatured RNA and DNA fragments to nitrocellulose. Proc. Natl. Acad. Sci. USA **77**:5201–5205.
25. **Varmus, H., and R. Swanstrom.** 1985. Replication of retroviruses, p. 75–134. *In* R. Weiss, N. Teich, H. E. Varmus, and J. Coffin (ed.), The molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Watanabe, S., and H. Temin.** 1982. Encapsidation sequences for spleen necrosis virus, an avian retrovirus, are between the 5′ long terminal repeat and the start of the *gag* gene. Proc. Natl. Acad. Sci. USA **79**:5986–5990.