JOURNAL OF VIROLOGY, Jan. 1986, p. 229–236
0022-538X/86/010229-08$02.00/0

Vol. 57, No. 1

# Evidence that a Capped Oligoribonucleotide Is the Primer for Duck Hepatitis B Virus Plus-Strand DNA Synthesis

JAU-MIN LIEN, CAROL E. ALDRICH, AND WILLIAM S. MASON*
*Institute for Cancer Research, Philadelphia, Pennsylvania 19111*

Received 1 July 1985/Accepted 9 September 1985

**The plus strand of virion DNA of duck hepatitis B virus possessed, at its 5′ terminus, a capped oligoribonucleotide 18 to 19 bases in length. This oligoribonucleotide had a unique 5′ end, the heterogeneity in length reflecting two distinct junctions with plus-strand DNA that were 1 base apart. The sequence of the RNA differed from that predicted by the sequence of duck hepatitis B virus upstream of the 5′ ends of plus-strand DNA but was identical to a downstream sequence corresponding to the 5′ terminus of a major poly(A)$^+$ viral RNA mapped by Büscher and co-workers (Cell 40:717–724, 1985). This RNA transcript is thought to serve as the template (i.e., the pregenome) for minus-strand synthesis via reverse transcription. The results suggest that the pregenome also donates a capped oligoribonucleotide that acts as the primer of plus-strand DNA synthesis, using the minus-strand DNA as template.**

Hepatitis B-like viruses (hepadna viruses) are synthesized in the liver and released into the blood, from which virus is readily isolated. These small DNA viruses have a genome size of about 3 kilobase pairs (kbp). The duplex genome is held in a circular conformation by a short cohesive overlap between the 5′ ends of the two DNA strands. One strand, the viral minus strand, is complete and is believed to have specific 3′ and 5′ termini. The other strand, of plus polarity, is incomplete, with a specific 5′ terminus but with a 3′ terminus that is heterogeneous in location within a virion population, resulting in genomes that have single-stranded regions (33). Virions contain an endogenous DNA polymerase that can repair the single-stranded gaps.

From the cytoplasmic fraction of liver cells, it is possible to isolate immature viral cores. These immature cores have an endogenous DNA polymerase that carries out not only plus-strand but also minus-strand DNA synthesis. Analysis of these two reactions have shown that minus-strand DNA is synthesized by reverse transcription of an RNA intermediate, while plus-strand DNA is synthesized by copying the viral minus strand (32). This pathway is at least superficially similar to the process of proviral DNA synthesis by the RNA tumor viruses (34). Unlike the retroviruses, hepadna viruses appear to use a protein rather than a tRNA as a primer for reverse transcription (20). The primer for hapadna virus plus-strand synthesis had not been identified.

In a previous study, we mapped the ends of the cohesive overlap of duck hepatitis B virus (DHBV) genomic DNA and showed that the 5′ end of the minus strand corresponds, as might be expected, to the origin of reverse transcription (20). As with other members of the hepadna virus family (9, 10), the presumptive protein primer remains bound to the 5′ end of the mature minus strand of DHBV (21). The present work was undertaken to study the priming of plus-strand synthesis and was based upon the assumption that plus-strand synthesis begins at a single site corresponding to the 5′ end of the plus strand of virion DNA. Our initial approach was to test the possibility suggested by an earlier report (10) that the 5′ terminus of the plus strand might be blocked to phosphorylation catalyzed by polynucleotide kinase. A DHBV virion DNA preparation was treated with RNase A and bacterial alkaline phosphatase and then incubated with [$\gamma$-$^{32}$P]ATP and polynucleotide kinase. Subsequent analyses described below indicated that the 5′ end of the plus strand could be phosphorylated; however, the phosphorylated 5′ end of the plus-strand DNA determined in these experiments was mapped ca. 20 bases downstream from the previously determined 5′ plus-strand end of the cohesive overlap (20). In our earlier experiments (20) to map the plus-strand end of the cohesive overlap, the virion DNA had not been pretreated with RNase A. This fact, taken together with the results from the 5′ end labeling, suggested that ca. 20 bases of RNA were present at the 5′ end of the plus strand. Experiments that establish this point and that define the sequence of the RNA are described herein. The presence of an RNA on the 5′ end of the plus strand suggests a role in priming of plus-strand synthesis. A similar oligoribonucleotide that acts as a primer of plus-strand synthesis had recently been described in the retrovirus family (4, 23, 29, 30). The hepadna virus-associated RNA is extremely stable, being present on virtually all viral genomes.

## MATERIALS AND METHODS

**Virion DNA.** Except as indicated, DHBV was obtained from the pooled sera of 2- to 3-week-old congenitally infected Pekin ducks (22) hatched from eggs generously provided by Anna O'Connell (Institute for Cancer Research). Some experiments made use of cloned virus, originally derived by transfection of DHBV-free Pekin ducks (17) with cloned viral DNA (see below). To obtain large amounts of cloned virus, 1-day-old ducklings were inoculated intravenously with 0.1 ml of viremic serum from a duck infected by DNA transfection and exsanguinated 2 to 3 weeks postinfection.

Virus was purified from the bulk of serum components by differential and equilibrium centrifugation as previously described (18). Virus was then collected from appropriate gradient fractions by centrifugation and suspended in 0.225 M NaCl–15 mM Tris hydrochloride (pH 7.5)–5 mM EDTA–0.5 mg of protease K per ml–0.05% sodium dodecyl sulfate (SDS) at a concentration of ca. 25-fold (relative to serum). After a 1-h incubation at 37°C, samples were extracted two times with an equal volume of phenol-chloroform (1:1, vol/vol) (saturated with 0.5 M Tris hydrochloride; pH 8), and nucleic acids were collected by ethanol

* Corresponding author.

precipitation. In general, about 1 μg of DHBV DNA was recovered per 100 ml of serum.

**Recombinant DNAs.** The cloning of the entire DHBV genome into the *Eco*RI sites of pBR322 and of bacteriophage M13 mp7 has been described previously (16, 20). The M13 mp7 DHBV-specific sequences were from the same clone (*Eco*RI DHBV DNA) completely sequenced by Mandart et al. (15). This DNA was originally isolated and cloned from DHBV virions (16). The DHBV DNA cloned into pBR322 was derived from a nuclear DNA preparation enriched for covalently closed circular DNAs (20; J. Summers and C. Rogler, personal communication). The sequence through the cohesive overlap region of this latter DNA was obtained by the technique of Maxam and Gilbert (19), as described by Molnar-Kimber et al. (20). The sequence was identical to that of Mandart et al. (15). Both DNAs have been shown by Sprengel et al. (31) to be infectious when injected into the liver of Pekin ducks.

Cloned virus was obtained from the DHBV DNA inserted into pBR322, essentially as described by Seeger et al. (26). Briefly, the *Eco*RI insert was excised from pBR322, purified by preparative gel electrophoresis in low-melting-temperature agarose, and religated under conditions (25 μg/ml) giving a mixture of approximately equal parts of monomers, dimers, trimers, and some higher-order multimers. Coprecipitates of DHBV DNA (5 μg) and calcium phosphate (12) were injected into three sites in the liver of 1-day-old Pekin ducks in a total volume of 100 μl. Sera were collected 3 weeks later from those ducks (ca. 25%) that developed a viremia.

**5′-end-labeled plus-strand DNA.** The objective of the following protocol was to obtain the 5′-end-labeled plus-strand fragment mapping from the 5′ end of the plus strand of virion DNA to the *Acc*I site at position 2577 (Fig. 1A). Virion nucleic acids corresponding to ca. 0.25 μg of DHBV DNA were suspended in 100 μl of 1 mM Tris hydrochloride (pH 7.5)–1 mM EDTA and digested with RNase A at a concentration of 100 μg/ml for 1 h at 37°C. After the addition of SDS to 0.1% (wt/vol) final concentration and NaCl to 0.15 M, the solution was extracted with an equal volume of phenol-chloroform (1:1, vol/vol), and DNA was collected by ethanol precipitation. The DNA was then suspended and digested with 100 U of bacterial alkaline phosphatase for 1 h at 65°C in 50 μl of 50 mM NaCl–10 mM Tris hydrochloride (pH 8.3)–1 mM $MgCl_2$. An equal volume of 10 mM Tris hydrochloride (pH 7.5)–10 mM EDTA–0.2% (wt/vol) SDS–2 mg of pronase per ml was added, the mixture was incubated 1 h at 37°C and extracted with an equal volume of phenol-chloroform (1:1, vol/vol), and the DNA was ethanol precipitated. The DNA was then denatured by heating, and phosphorylation was carried out for 1 h at 37°C in 40 μl in 50 mM Tris hydrochloride (pH 7.5)–10 mM $MgCl_2$–5 mM dithiothreitol containing 18 U of T4 polynucleotide kinase and 35 pmol of [$\gamma$-$^{32}$P]ATP (2,900 Ci/mmol). The reaction was subjected to pronase digestion and phenol-chloroform extraction as above, and nucleic acids were ethanol precipitated together with 25 μg of wheat germ rRNA as carrier. Although greater than 10% of the isotope was incorporated into acid-precipitable material, the bulk of the incorporated label was in low-molecular-weight contaminating nucleic acid fragments, and further purification was required.

Partial purification of plus-strand DNA was achieved by annealing the denatured, $^{32}$P-labeled kinase reaction product to 2.5 μg of the M13 mp7 DHBV minus-strand DNA clone (or plus-strand DNA as a control) in 28 μl of 0.6 M NaCl–40 mM Tris hydrochloride (pH 7.5)–2 mM EDTA for 1 h at 68°C, followed by preparative gel electrophoresis in low-melting-temperature agarose containing 0.5 μg of ethidium bromide per ml (16). The DNA was visualized with UV light, and a section of the gel extending from the M13 band, at the position of a 2.3-kbp marker DNA, to a 6.6-kbp marker was removed and extracted. A portion (75%) of each DNA preparation was then cleaved with 9 U of *Acc*I for 1 h at 37°C in 6 mM NaCl–6 mM Tris hydrochloride (pH 7.5)–6 mM $MgCl_2$–6 mM 2-mercaptoethanol–0.01% (vol/vol) Triton X-100 and further analyzed as described below. Even after preparative agarose gel electrophoresis, only about 4% of the radioactivity (selection for plus-strand DNA) was present in the terminal *Acc*I plus-strand fragment. A substantial amount of the contamination present at this stage appeared to be due to binding of nonviral sequences to the bacteriophage M13 sequences, a conclusion supported by annealing of the end-labeled nucleic acids with wild-type M13 mp7 bacteriophage DNA. The recovery of label in the *Acc*I fragment was about 1% of the expected value; the apparent low efficiency of labeling may be due to the high proportion of contaminating low-molecular-weight nucleic acid fragments in the polynucleotide kinase reaction.

**3′-end-labeled plus-strand DNA.** The objective of the following procedures was to obtain the 5′-end-terminal fragment of the virion plus strand, 3′ end labeled at the *Acc*I site at position 2577 (GT ↓ CTAC) (Fig. 1B) by repair synthesis with [$\alpha$-$^{32}$P]dCTP. Total virion nucleic acids, corresponding to ca. 0.25 μg of virion DNA, were suspended and digested with 6 U of *Acc*I in a 30-μl volume of digestion buffer containing 10 mM dithiothreitol and 45 U of placental RNase inhibitor (Bolton Biologicals, Richmond Heights, Mo.). The solution was then brought to a volume of 100 μl with 0.15 M NaCl–10 mM Tris hydrochloride (pH 7.5)–10 mM EDTA–0.1% (wt/vol) SDS and extracted with an equal volume of phenol-chloroform (1:1, vol/vol), and the nucleic acids were collected by precipitation with ethanol. The nucleic acids were then suspended in 10 μl of *Acc*I buffer, supplemented as above, containing 5 pmol of [$\alpha$-$^{32}$P]dCTP (800 Ci/mmol) and 2.5 U of Klenow fragment DNA polymerase I and incubated for 30 min at room temperature. Approximately 30 to 50% of the $^{32}$P was rendered acid precipitable. The reaction was stopped by the addition of 10 mM Tris hydrochloride (pH 7.5)–10 mM EDTA–0.1% (wt/vol) SDS–25 μg of wheat germ rRNA to a volume of 100 μl, and the nucleic acids were collected by ethanol precipitation. The nucleic acids were then suspended, hybridized with either the plus or minus strand of DHBV in bacteriophage M13 mp7, and subjected to preparative gel electrophoresis as described above, except that hybridization and electrophoresis were done in the presence of 10 μM aurintricarboxylic acid, an RNase inhibitor (11, 13). As for the 5′ end labeling, the bulk of radiolabel was incorporated into nonviral nucleic acids, and only about 5% of the radiolabel recovered from the preparative gel (selection for plus-strand DNA) was in the 3′-end-labeled *Acc*I fragment containing the 5′ terminus of the plus strand. Additional details of analysis and further purification of this fragment by polyacrylamide gel electrophoresis are described below.

**Polyacrylamide gel electrophoresis.** End-labeled DNA fragments were denatured and analyzed by gel electrophoresis on 8 or 10% polyacrylamide gels (60 or 80 cm) containing 8 M urea, 50 mM Tris borate (pH 8.3), 1 mM EDTA (19), and 0.1% (wt/vol) SDS. Two different molecular weight markers were used: (i) bacteriophage φX174 DNA was digested with *Hin*fI, and the restriction endonuclease fragments were 3′ end labeled essentially as described above, except that three

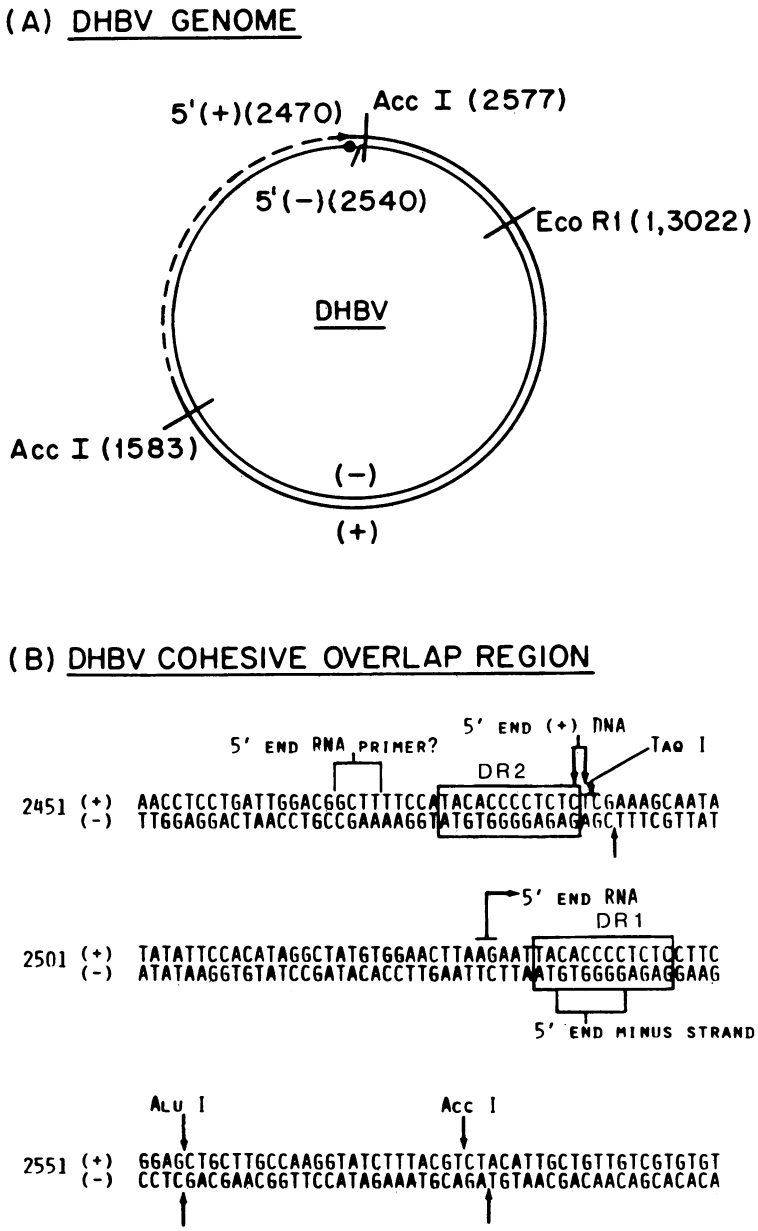


FIG. 1. Structure of the DHBV genome. (A) Location of the cohesive overlap relative to the *Acc*I cleavage site at position 2577, which has been used to characterize the 5′ end of the plus strand (15, 20). (B) DNA sequence through the cohesive overlap region, indicating the 12-bp direct repeat (boxed), *Taq*I, *Alu*I, and *Acc*I cleavage sites, the approximate position of the 5′ end of the minus strand, and the previously reported position of the 5′ end of the plus strand (here indicated as 5′ end of RNA primer) (20). The actual 5′ ends of plus-strand DNA are also indicated. The previous estimate for the 5′ end of the plus strand was based upon the length of the plus-strand fragment from the *Acc*I site to the 5′ end, including the oligoribonucleotide (see text).

deoxynucleotides were added to the 3′ termini; and (ii) cloned DHBV DNA 3′ end labeled at the *Acc*I site at position 2577 with [α-$^{32}$P]dCTP was cleaved by the method of Maxam and Gilbert (19, 24) to create a sequence ladder corresponding to the 5′-terminal *Acc*I plus-strand fragment of virion DNA which is under analysis. After electrophoresis, gels were transferred to a sheet of plastic film, covered with Saran Wrap, and subjected to autoradiography at −70°C with Kodak XAR film and Du Pont Cronex Lightning-Plus intensifying screen.

As indicated earlier, the plus-strand DNA fragment 3′ end labeled at the *Acc*I site at position 2577 was given a final purification by polyacrylamide gel electrophoresis. The labeled fragment was located by autoradiography, and the region of the gel containing this fragment was cut out. The fragment was eluted by the addition of 0.2 ml of 0.5 M NaCl–10 mM Tris hydrochloride (pH 7.4)–1 mM EDTA–0.1% (wt/vol) SDS to the gel slice and incubation for 16 h at 37°C with gentle agitation. The fragment was then collected from the soluble phase by ethanol precipitation.

**RNA sequencing.** RNA sequencing was done essentially as described by Donis-Keller et al. (7). The terminal plus-strand fragment, 3′ end labeled at the *Acc*I site, was isolated and purified as above. A cloned plus-strand DNA fragment, 3′

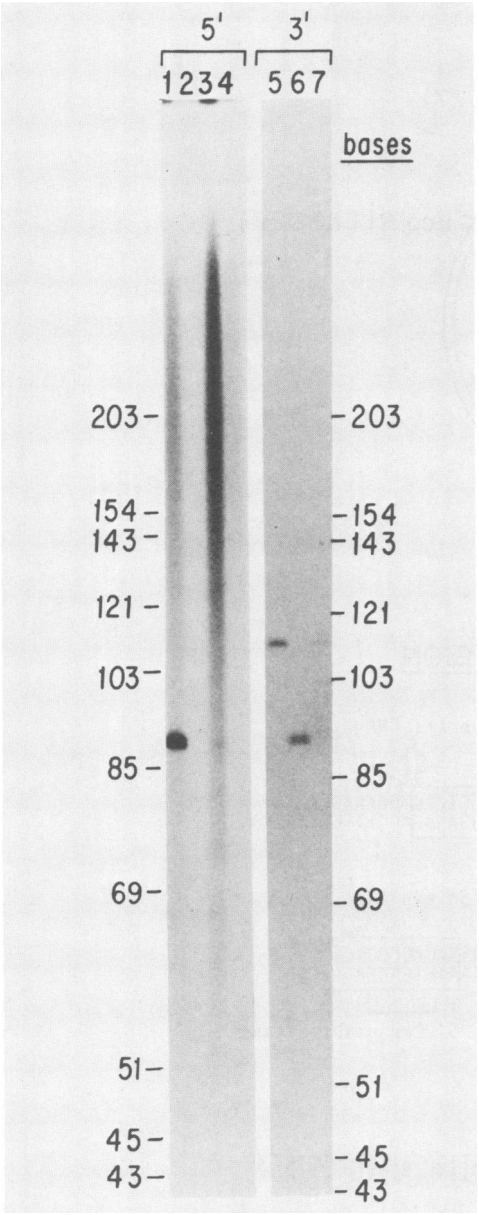


FIG. 2. Evidence for an oligoribonucleotide at the 5′ end of the plus strand of virion DNA. Virion DNA was denatured and either 5′ end labeled with [γ-$^{32}$P]ATP and T4 polynucleotide kinase or 3′ end labeled with [α-$^{32}$P]dCTP after *Acc*I cleavage, as described in Materials and Methods. The DNA was digested with RNase A and bacterial alkaline phosphatase before the 5′ end labeling but not the 3′ end labeling. Plus and minus strands were then partially purified by preparative gel electrophoresis in low-melting-temperature agarose. The purified fragments were subjected to polyacrylamide gel electrophoresis on a 60-cm 10% polyacrylamide gel and detected by autoradiography. Lanes: 1 to 4, 5′ end labeled; 5 to 7, 3′ end labeled; 1, 75% of purified plus strand, *Acc*I cut; 2, 25% of purified plus strand, not cut with *Acc*I; 3, 75% of purified minus strand, *Acc*I cut; 4, 25% of purified minus strand, not *Acc*I cut; 5, purified *Acc*I-cleaved plus strand; 6, purified *Acc*I-cleaved plus strand, incubated for 5 min at 100°C in 0.3 N NaOH before electrophoresis; 7, same as lane 6, except with only 0.5 min of incubation in alkali.

end labeled at the same *Acc*I site, was cleaved by the method of Maxam and Gilbert (19, 24) to provide a size marker for the virion DNA fragment from which the RNA had been totally removed. About 600 cpm of the purified virion fragment per lane was subjected to partial digestions at 50°C by five different RNases, namely, T1 (7), U2 (7), Physarum M (6), *Bacillus cereus* (14) (P-L Biochemicals), and chicken liver III (2) (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) under the incubation buffers as described previously (5). Partial alkaline hydrolysis was done by incubation in 50 mM $Na_2CO_3$-$NaHCO_3$ (pH 9.0)–1 mM EDTA at 90°C for 15 min. Tobacco acid pyrophosphatase (TAP) (Bethesda Research Laboratories) digestion was performed by incubation in 50 mM sodium acetate (pH 5.0)–10 mM 2-mercaptoethanol–1 mM EDTA for 30 min at 37°C (8, 28).

## RESULTS

**Presence of RNA at the 5′ end of the plus strand of DHBV virion DNA.** Previous studies (20) have determined the location of the cohesive overlap region of DHBV virion DNA relative to the known deoxynucleotide sequence (15). The 5′ ends of the minus and plus strands were located approximately at positions 2537 to 2542 and 2468 to 2471, respectively (Fig. 1). Direct repeat sequences of 12 bp (DR1 and DR2) are found near the termini of the minus and plus strands (15, 20). The 5′ end of the presumptive pregenomic RNA recently mapped by Büscher et al. (3) is located just a few nucleotides from the minus-strand DNA initiation site (20) (Fig. 1B).

As a first step in studying the priming events of plus-strand DNA synthesis, we assessed the availability of the 5′ terminus of the plus strand of DHBV virion DNA to phosphorylation by polynucleotide kinase. The virus genome was digested with RNase A to completion, treated with alkaline phosphatase, denatured, and incubated with [γ-$^{32}$P]ATP and polynucleotide kinase. Genomic plus-strand DNA was partially purified by agarose gel electrophoresis after hybridization with cloned minus-strand DNA or plus-strand DNA as a control. After digestion with *Acc*I, which cuts ca. 110 bases downstream from the previously mapped 5′ end of the plus strand, the radiolabeled DNA was denatured and resolved by electrophoresis in a urea-polyacrylamide gel (Fig. 2). When this DNA was hybrid selected with cloned minus-strand DNA and digested with *Acc*I, two species ca. 91 to 92 bases in length were detected (lane 1). (These species were not well resolved in this reproduction and are more clearly demonstrated in Fig. 3.) The species were not detected when *Acc*I digestion was omitted (lane 2). Furthermore, when the hybrid selection was carried out with cloned plus-strand DNA followed by *Acc*I digestion, only a faint signal was detected at the same position (lane 3). Again, no distinctive band was present without *Acc*I treatment (lane 4). Since our protocol yielded two plus-strand DNA fragments differing in length by 1 base and about 20 bases shorter than was expected from our previous studies (20), the possibility was raised that an oligoribonucleotide was present at the 5′ end of plus-strand DNA. This hypothesis was compatible with previous results since virion DNA used in the earlier characterization of the cohesive overlap region had not been pretreated with either RNase A or alkali (20).

To test the hypothesis, virion nucleic acids were extracted, digested with *Acc*I, and 3′ end labeled with [α-$^{32}$P]dCTP by using the Klenow fragment DNA polymerase I in the presence of placental RNase inhibitor. Genomic plus strands were again purified by hybridization with cloned minus-strand DNA. Unlike the results described above, a single species with the expected size of ca. 110 bases was now resolved in a denaturing gel (Fig. 2, lane 5). Complete

alkaline hydrolysis of this species resulted in two fragments ca. 20 bases shorter (lane 6) and differing in length by 1 base. Again, better resolution of these species is illustrated in Fig. 3. The same results have been obtained with RNase A digestion of the 110-base fragment (data not shown). Moreover, partial alkaline hydrolysis with sodium bicarbonate implied that every base, except perhaps the penultimate, was a ribonucleotide (Fig. 3). The data indicated that there was an oligoribonucleotide attached to the 5′ end of the plus strand of DHBV virion DNA. Furthermore, the 5′ end of this species appeared to be unique, corresponding to what was previously estimated to be the 5′ end of the plus-strand DNA.

**Sequencing of the 5′-terminal oligoribonucleotide.** As a first step in characterizing the 5′-terminal oligoribonucleotide, the 110-base plus-strand *Acc*I fragment, 3′ labeled at the *Acc*I site, was purified and subjected to digestion, to completion, with RNases A, U2, and T1. Although both RNases A and U2, which cleave on the 3′ side of pyrimidine and adenine residues, respectively, consistently gave the results expected from the viral DNA sequence (Fig. 1B), RNase T1, which cleaves on the 3′ side of guanine residues, reproducibly converted the 110-base plus-strand fragment to a 107-base fragment. This result cannot be explained by the DNA sequence upstream of the start site of plus-strand DNA (Fig. 1B).

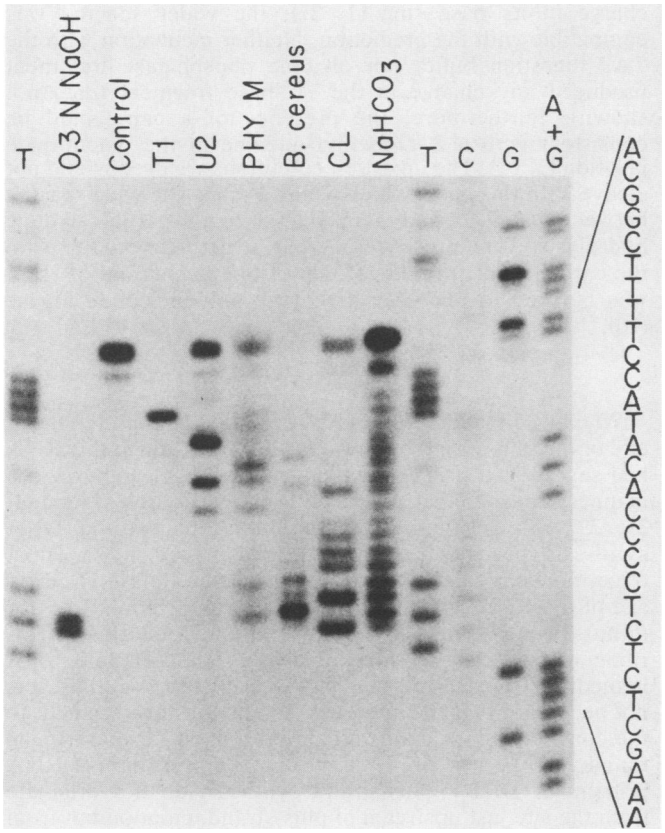


FIG. 3. Determination of the sequence of the oligoribonucleotide at the 5′ end of the plus strand of virion DNA. Equal amounts of the terminal plus-strand fragment, 3′ end labeled at the *Acc*I site at position 2577, were subjected to polyacrylamide gel electrophoresis after incubation with the indicated RNases under conditions giving partial hydrolysis. Incubation with 0.3 N NaOH was for 5 min at 100°C, followed by neutralization and ethanol precipitation to eliminate the high salt. As a size marker, the plus strand of cloned DHBV DNA was 3′ end labeled at the *Acc*I site at position 2577 and subjected to cleavage by the procedure of Maxam and Gilbert (19) (lanes C, G, A+G) with the modifications of Rubin and Schmid (24) (lane T). It should be noted that cleavage by the Maxam and Gilbert procedure (19) is 3′ of the base(s) indicated at the top of the lanes and leaves a 5′ phosphate at the end of the labeled fragment. In contrast, cleavage by alkali or the RNases illustrated here removes the 5′-terminal phosphate from the labeled fragment. Removal of the terminal phosphate causes a fragment to migrate slower (ca. 1 base) than the corresponding phosphorylated fragment (29). The sequence of the DNA cleaved by the modified Maxam and Gilbert procedure (19, 24) is shown at the right.

To eliminate the possibility that the virus we characterized differed from that sequenced by Mandart et al. (15) and Molnar-Kimber et al. (20), we carried out subsequent experiments with molecularly cloned virus, as described in Materials and Methods. Complete RNA sequencing was then performed. The polyacrylamide gel-purified 110-base plus-strand fragment, 3′ end labeled at the *Acc*I site, is shown in the control lane of Fig. 3. A band migrating faster in the control lane is believed to be the degradation product of the main band. Again, complete alkaline hydrolysis produced two bands with about the same intensity, while partial hydrolysis with sodium bicarbonate gave 19 bands, which serve to align the fragments produced by enzymatic hydrolysis. Partial RNase digestions were adjusted to give optimal results, utilizing five RNases of differing sequence specificity. In addition to RNases T1 and U2 described above, RNase *B. cereus* cleaves specifically at pyrimidine residues, RNase chicken liver III cleaves at cytidine residues, and RNase Phy M cleaves at both adenine and uridine residues. By this approach, we determined the 17 bases immediately adjacent to the distal 5′ terminus of plus-strand DNA and located the two 5′ termini of plus-strand DNA in C and T, at positions 2488 and 2489, respectively. It should be noted that none of the digestion conditions appeared to produce a fragment corresponding to a cleavage between bases 2 and 3 from the upstream end of the oligoribonucleotide. The location of adenine in the expected third position from the end was deduced from the results of the U2 digestion.

The RNA sequence determined from the data in Fig. 3 is summarized in Fig. 4. The RNA sequence corresponding to the region upstream of the 5′ end of the plus strand of virion DNA is shown in the upper row. The two sequences aligned completely in the 12-base direct-repeat region but diverged in the nucleotides upstream of the direct repeat. However, a perfect alignment was obtained when the RNA sequence was compared with the plus-strand sequence in the region where minus-strand DNA synthesis initiates (cf. Fig. 1B and the lower row in Fig. 4).

To confirm that the plus-strand DNA fragment with this 5′-terminal RNA does, in fact, initiate from positions 2488 and 2489 rather than in the sequence immediately following the direct repeat at the minus-strand DNA initiation site, a *Taq*I digestion of this fragment was carried out. *Taq*I digestion gave a 90-base band (Fig. 5), as predicted from the known sequences (cf. Fig. 1). *Alu*I digestion also gave the predicted restriction fragment (data not shown).

**Presence of a cap structure on the 5′-terminal oligoribonucleotide.** Büscher et al. (3) have recently mapped the 5′ end of a 3.5-kb poly(A)$^+$ RNA from DHBV-infected liver at positions 2530/2531 ± 1 and suggested that this RNA might serve both as an mRNA and as the template for reverse transcription of the viral genome, i.e., the pregenome (32). These positions were very near the 5′ end (position 2529) of the oligoribonucleotide that we sequenced.

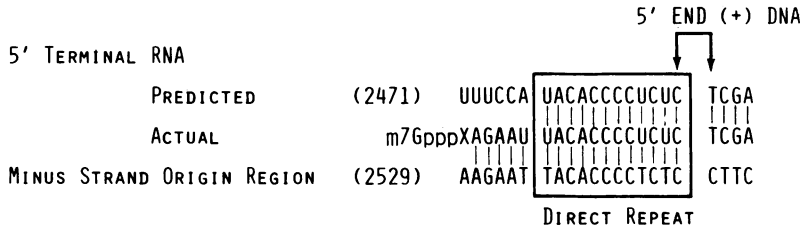


FIG. 4. Sequence of the oligoribonucleotide located at the 5′ end of the plus strand of virion DNA. The RNA sequence, determined as described in the legend to Fig. 3 and the text, is shown in the center. Evidence for the two 5′-terminal bases is presented subsequently. The top line shows the RNA sequence corresponding to the region upstream of the 5′ end of plus-strand DNA (15, 20) (cf. Fig. 1B). The bottom line shows the DNA sequence through the minus-strand origin (Fig. 1). The 12-base direct-repeat sequence (Fig. 1) is boxed.

To further test the possibility that the oligoribonucleotide might be derived directly from the 5′ end of the pregenomic RNA, we treated the 110-base plus-strand fragment, 3′ end labeled at the *Acc*I site, with the decapping enzyme TAP, which has been shown to remove terminal cap structures without cleaving the polynucleotide chain (28). TAP treatment shortened the 110-base fragment by approximately 1 base (Fig. 5). Since the cap structure contains a positive charge in its base ring (1, 27), the wider spacing was compatible with the prediction. Neither incubation with the TAP digestion buffer nor alkaline phosphatase treatment produced any change in the 110-base fragment (data not shown). Furthermore, the presence of a cap would be consistent with a 2′-*O*-methyl adenine in the penultimate position (1, 27). The digestion conditions we used would not cleave 3′ to this base, which would explain the wide spacing between bands 2 and 3 in Fig. 3 (e.g., partial alkaline hydrolyses). The presence of a cap structure would explain the earlier report that the 5′ end of the plus-strand of HBV was blocked to phosphorylation by polynucleotide kinase (10).

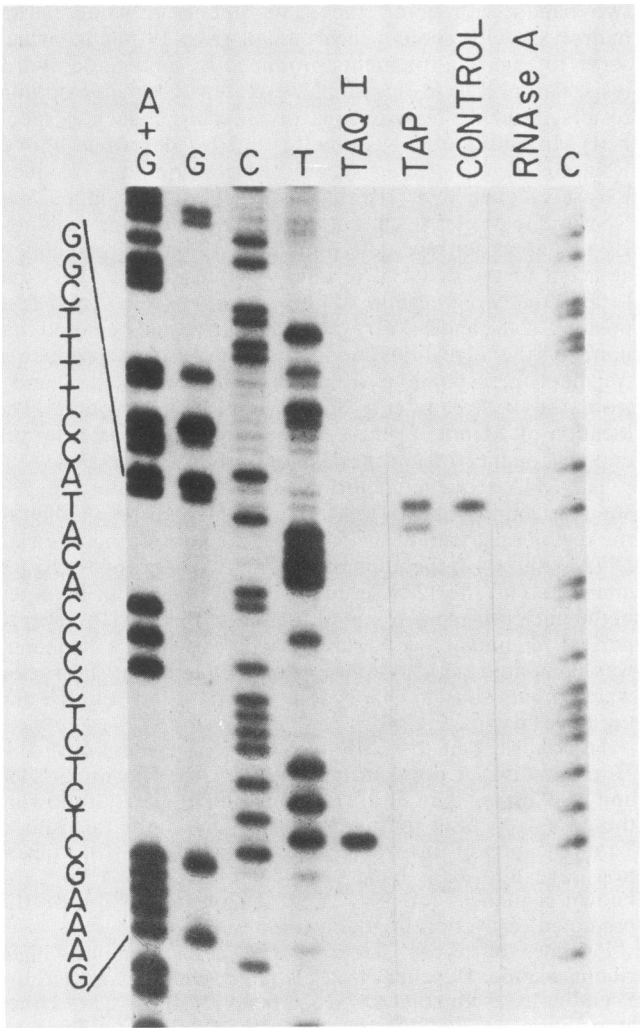


FIG. 5. Evidence that the 5′-terminal oligoribonucleotide on the DHBV plus strand is capped. The 110-base plus-strand fragment, 3′ end labeled at the *Acc*I site, was digested with various enzymes as described in the text. The undigested fragment (control) and the fragment from which RNA was completely removed by RNase A digestion are shown for comparison with the TAP- and *Taq*I-treated fragment. The molecular weight marker, cloned plus-strand DNA 3′ end labeled at the *Acc*I site and cleaved by the modified Maxam and Gilbert procedure (19, 24), is described in the legend to Fig. 3. The sequence at the left is for this cloned DNA.

## DISCUSSION

When we first detected an RNA covalently bound to the 5′ end of the plus strand of virion DNA, we assumed this RNA had served as a primer for plus-strand DNA synthesis in a manner similar to that described in the retrovirus system (4, 23, 29, 30). In the retrovirus family, plus-strand DNA synthesis has been shown to be primed by a short oligoribonucleotide generated by a specific RNase H cleavage of the RNA template for reverse transcription (i.e., for minus-strand synthesis). The sequence of this oligoribonucleotide is the same as that of the genomic RNA immediately upstream of the plus-strand start site; thus, it is not necessary that the hydrogen bonding of this fragment to the complementary minus strand be broken for the priming function to be served. By contrast, the presumptive RNA primers of DHBV plus-strand synthesis do not originate from the site just upstream of plus-strand initiation but from almost 50 bases downstream. Our data, together with those of Büscher et al. (3), suggest that the primers are the terminal 18 or 19 bases of pregenomic RNA (32).

A model of DHBV DNA synthesis (20, 32) incorporating these recent results is shown in Fig. 6. The presumptive structure of pregenomic RNA, taken from the work of Büscher et al. (3), is shown at the top. R signifies the 270-base terminal redundancy on this 3.5-kb poly(A)$^+$ RNA. As shown, reverse transcription initiates within the 12-base direct repeat (Fig. 1), presumably from a protein primer (20). The RNA template is degraded during minus-strand elonga-

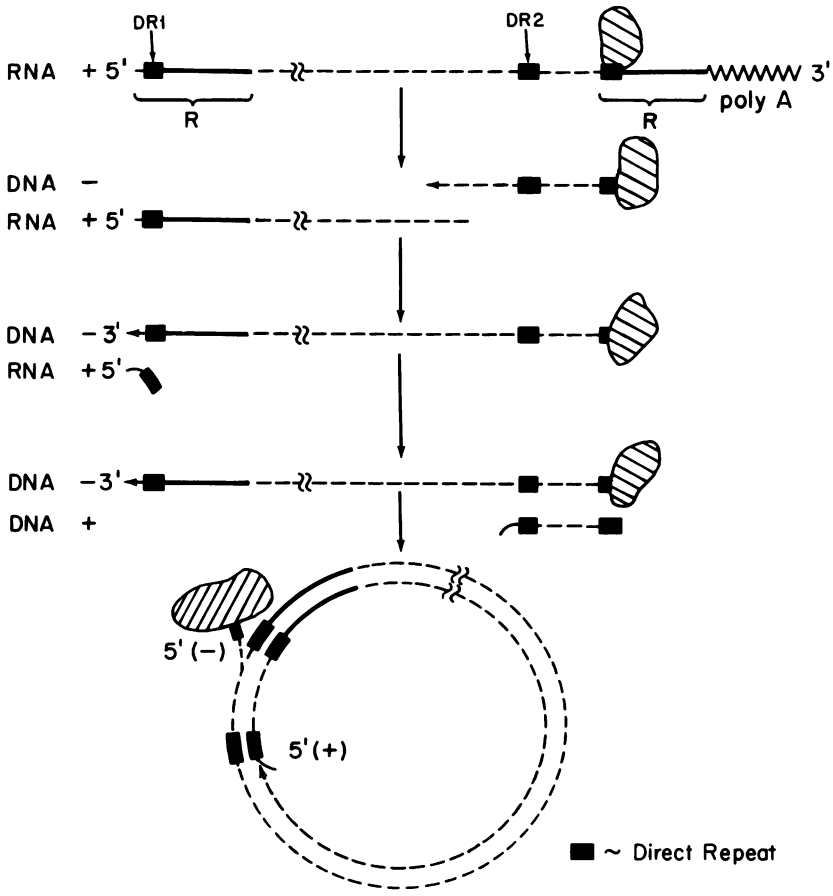


FIG. 6. Mechanism of DHBV DNA synthesis. Details are presented in the text. The 12-base direct-repeat sequence is shown as a solid box (not to scale). R, 270-base terminal redundancy on presumptive pregenomic RNA (3).

tion, possibly by an RNase H activity, so that intermediates in minus-strand elongation are essentially single stranded (32). After completion of the minus strand, the 5′-terminal pregenome oligoribonucleotides bind to the minus strand just upstream of the plus-strand start site and act as a primer. Our data suggest that the pregenome is capped. The 12-base direct-repeat sequence (DR2) serves to align the primers at this site, and plus-strand synthesis begins. This mechanism of plus-strand priming may explain the significance for conservation of a 12- to 13-base direct repeat in the cohesive overlap region of all members of the hepadna virus family (25). Circularization of the viral genome is essential to complete plus-strand synthesis.

Three points need to be made concerning this model. First, the mechanism for circularization during plus-strand elongation past the 5′ end of the minus strand is unclear; however, such a mechanism may be provided by preliminary results from our laboratory, suggesting that the minus strand of DHBV contains a short terminal redundancy of ca. 7 to 8 bp, apparently resulting from synthesis to the 5′ end of the pregenome (D. Petcu and W. Mason, unpublished observations). Second, the proposal that minus-strand synthesis initiates at the 3′ DR1 sequence, rather than the 5′ DR1, of the pregenome is made for simplicity of illustration; initiation might equally well occur at the 5′ DR1. Third, the model proposes that the plus-strand primer is created by translocation of the terminal oligoribonucleotide of the pregenome to DR2 from what would appear to be a more stable base pairing with the minus strand at DR1. This dissociation and translocation within viral replication complexes may be facilitated by protein-nucleic acid interactions involving the viral core subunit or the protein primer of minus-strand synthesis (or both), which may have the ability to specifically interact with the viral RNA in the vicinity of DR1. An alternative to our model is that the oligoribonucleotide primer of plus-strand synthesis is not derived from the reverse-transcribed pregenome but rather is captured from a separate viral RNA and packaged into replication complexes before initiation of viral DNA synthesis. Further experimentation will be required to resolve these alternatives.

In conclusion, it is important to note that preliminary results also suggest for human hepatitis B virus a terminal redundancy of the DNA minus strand and an oligoribonucleotide linked to the plus strand (H. Will and colleagues, personnel communication). Likewise, a 5′-terminal oligoribonucleotide has been demonstrated on the plus strand of ground squirrel hepatitis virus DNA, and sequencing has led to the conclusion that the oligoribonucleotide is equivalent to that described herein (C. Seeger and H. Varmus, personnel communication).

## ACKNOWLEDGMENTS

We are pleased to acknowledge Pei-Jer Chen, Katherine Molnar-Kimber, and John Taylor for helpful suggestions during the course of this work. We also thank Thomas London, Anna O'Connell, Jesse Summers, and John Taylor for a critical reading of

the manuscript, Laura Coates for technical assistance, and Helene McGinley for help in manuscript preparation.

This investigation was supported by Public Health Service grants AI-18641, CA-06927, and RR-05539 from the National Institutes of Health and by an appropriation from the Commonwealth of Pennsylvania.

## LITERATURE CITED

1. **Banerjee, A. K.** 1980. 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. Microbiol. Rev. **44**:175–205.
2. **Boguski, M. S., P. A. Hieter, and C. C. Levy.** 1980. Identification of a cytidine-specific ribonuclease from chicken liver. J. Biol. Chem. **255**:2160–2163.
3. **Büscher, M., W. Reiser, H. Will, and H. Schaller.** 1985. Transcripts and putative RNA pregenome of duck hepatitis B virus: implication for reverse transcription. Cell **40**:717–724.
4. **Champoux, J. J., E. Gilboa, and D. Baltimore.** 1984. Mechanism of RNA primer removal by the RNase H activity of avian myeloblastosis virus reverse transcriptase. J. Virol. **49**:686–691.
5. **Chen, P.-J., A. Cywinski, and J. M. Taylor.** 1985. Reverse transcription of 7S L RNA by an avian retrovirus. J. Virol. **54**:278–284.
6. **Donis-Keller, H.** 1980. Phy M: an RNase activity specific for U and A residues useful in RNA sequence analysis. Nucleic Acids Res. **8**:3133–3142.
7. **Donis-Keller, H., A. M. Maxam, and W. Gilbert.** 1977. Mapping adenines, guanines and pyrimidines in RNA. Nucleic Acids Res. **4**:2527–2538.
8. **Efstratiadis, A., J. N. Vournakis, H. Donis-Keller, G. Chaconnas, D. K. Dougall, and F. C. Kafatos.** 1977. End labeling of enzymatically decapped mRNA. Nucleic Acids Res. **4**:4165–4174.
9. **Ganem, D., L. Greenbaum, and H. E. Varmus.** 1982. Viral DNA of ground squirrel hepatitis virus: structural analysis and molecular cloning. J. Virol. **44**:374–383.
10. **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.
11. **Givens, J. F., and K. F. Manly.** 1976. Inhibition of RNA-directed DNA polymerase by aurintricarboxylic acid. Nucleic Acids Res. **3**:405–418.
12. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
13. **Hallick, R. B., B. K. Chelm, P. W. Gray, and E. M. Orozco, Jr.** 1977. Use of aurintricarboxylic acid as an inhibitor of nucleases during nucleic acid isolation. Nucleic Acids Res. **4**:3055–3064.
14. **Lockard, R. E., B. Alzner-Deweerd, J. E. Heckman, J. MacGee, M. W. Tabor, and U. L. RajBhandary.** 1978. Sequence analysis of 5' $^{32}$P labeled mRNA and tRNA using polyacrylamide gel electrophoresis. Nucleic Acids Res. **5**:37–56.
15. **Mandart, E., A. Kay, and F. Galibert.** 1984. Nucleotide sequence of a cloned duck hepatitis virus genome: comparison with the human and woodchuck hepatitis B virus sequences. J. Virol. **49**:782–792.
16. **Mason, W. S., C. Aldrich, J. Summers, and J. M. Taylor.** 1982. Asymmetric replication of duck hepatitis B virus DNA in liver cells: free minus-strand DNA. Proc. Natl. Acad. Sci. USA **79**:3997–4001.
17. **Mason, W. S., M. S. Halpern, J. M. England, G. Seal, J. Egan, L. Coates, C. Aldrich, and J. Summers.** 1983. Experimental transmission of duck hepatitis B virus. Virology **131**:375–384.
18. **Mason, W. S., G. Seal, and J. Summers.** 1980. Virus of Pekin ducks with structural and biological relatedness to human hepatitis B virus. J. Virol. **36**:829–836.
19. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
20. **Molnar-Kimber, K. L., J. Summers, and W. S. Mason.** 1984. Mapping of the cohesive overlap of duck hepatitis B virus DNA and of the site of initiation of reverse transcription. J. Virol. **51**:181–191.
21. **Molnar-Kimber, K. L., J. Summers, J. M. Taylor, and W. S. Mason.** 1983. Protein covalently bound to minus-strand DNA intermediates on duck hepatitis B virus. J. Virol. **45**:165–172.
22. **O'Connell, A. P., M. K. Urban, and W. T. London.** 1983. Naturally occurring infection of Pekin duck embryos by duck hepatitis B virus. Proc. Natl. Acad. Sci. USA **80**:1703–1706.
23. **Resnick, R., C. A. Omer, and A. J. Faras.** 1984. Involvement of retrovirus reverse transcriptase-associated RNase H in the initiation of strong stop (+) DNA synthesis and the generation of the long terminal repeat. J. Virol. **51**:813–821.
24. **Rubin, C. M., and C. W. Schmid.** 1980. Pyrimidine-specific chemical reactions useful for DNA sequencing. Nucleic Acids Res. **8**:4613–4619.
25. **Seeger, C., D. Ganem, and H. E. Varmus.** 1984. The nucleotide sequence of an infectious, molecularly cloned genome of the ground squirrel hepatitis virus. J. Virol. **51**:367–375.
26. **Seeger, C., D. Ganem, and H. E. Varmus.** 1984. The cloned genome of ground squirrel hepatitis virus is infectious in the animal. Proc. Natl. Acad. Sci. USA **81**:5849–5852.
27. **Shatkin, A. J.** 1976. Capping of eukaryotic mRNAs. Cell **9**:645–653.
28. **Shinshi, H., M. Miwa, and T. Sugimura.** 1976. Enzyme cleaving the 5'-terminal methylated blocked structure of messenger RNA. FEBS Lett. **65**:254–257.
29. **Smith, J. K., A. Cywinski, and J. M. Taylor.** 1984. Initiation of plus-strand DNA during reverse transcription by an avian retrovirus genome. J. Virol. **49**:200–204.
30. **Smith, J. K., and J. M. Taylor.** 1984. The specificity of initiation of plus-strand DNA by Rous sarcoma virus. J. Virol. **52**:314–319.
31. **Sprengel, R., C. Kuhn, C. Manso, and H. Will.** 1984. Cloned duck hepatitis B virus DNA is infectious in Pekin ducks. J. Virol. **52**:932–937.
32. **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.
33. **Tiollais, P., P. Charnay, and G. N. Vyas.** 1981. Biology of hepatitis B virus. Science **213**:406–411.
34. **Weiss, R., N. Teich, H. Varmus, and J. Coffin.** 1982. RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.