JOURNAL OF VIROLOGY, Dec. 1987, p. 3832–3840
0022-538X/87/123832-09$02.00/0

Vol. 61, No. 12

Initiation and Termination of Duck Hepatitis B Virus DNA Synthesis during Virus Maturation

JAU-MIN LIEN, DAVID J. PETCU, CAROL E. ALDRICH, AND WILLIAM S. MASON*
Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Received 18 May 1987/Accepted 13 August 1987

We characterized a number of important features of the structure of the cohesive overlap region of the DNA genome of duck hepatitis B virus. The 5′-terminal nucleotide of minus-strand DNA was localized to nucleotide 2537, a G residue within the 12-base repeat sequence DR1. This G residue was shown to be the site of a covalent linkage to a protein, consistent with speculation that this protein is the primer of minus-strand synthesis, which occurs by reverse transcription. The 3′ terminus of the minus strand was heterogeneous, being mapped to nucleotides 2530 and 2531, indicating that the minus strand is terminally redundant by seven or eight bases and ends at the putative 5′ end of the transcribed RNA template (pregenome) for reverse transcription. We previously demonstrated that the presumptive RNA primer of plus-strand synthesis remains attached to plus-strand DNA during virus maturation; moreover, the sequence of this primer suggested an origin from the 5′ end of the pregenome (J.-M. Lien, C. E. Aldrich, and W. S. Mason, J. Virol. 57:229–236, 1986). We show here that over 75% of plus-strand primers are capped, further supporting the idea that these primers are uniquely derived from the 5′ end of the pregenome. Finally, we found that seemingly mature duck hepatitis B virus genomes are incomplete by at least 12 bases, in that the 12-base repeat sequence DR2 is not copied into plus-strand DNA during virus maturation. Since DR2 in virion DNA is duplexed with the RNA primer of plus-strand synthesis, it is possible that the failure to make complete plus strands is due to an inability of the viral DNA polymerase to carry out a displacement of the bound RNA primer.

Hepadnaviruses are a family of small DNA viruses with a genome size of about 3 kilobase pairs. The duplex DNA genome is held in a circular conformation by a cohesive overlap between the 5′ ends of the two DNA strands (25). One strand, the minus strand, has been considered to be of full length. The other strand, the plus strand, varies from 50 to 80% of full length in the mammalian hepadnaviruses (9, 12, 31) and appears complete in the majority of the virus particles of duck hepatitis B virus (DHBV) (18). Incomplete plus strands can be readily elongated in vitro by a DNA polymerase activity inside the virions (11).

To elucidate the replicative mechanism of the hepadnaviruses, DNA synthesis complexes (immature viral cores) from DHBV-infected livers have been isolated and characterized (30). By analyzing the nucleic acids within these complexes, it was shown that hepadnaviruses replicate asymmetrically, in that minus-strand DNA is synthesized from an RNA template (pregenome) via reverse transcription, while the plus-strand DNA is copied from the minus-strand DNA template. The data suggest a replicative scheme in many ways similar to that of retroviruses (17, 35).

An examination of the intermediates in hepadnavirus DNA synthesis complexes strongly suggests that there are only two sites for initiation of DNA synthesis, corresponding to the 5′ ends of the cohesive overlap. We and others have therefore been carrying out a detailed characterization of the cohesive overlap region of virion DNA. The results of Molnar-Kimber et al. (21) have shown that the cohesive overlap in the DHBV genome falls within the boundaries of a short direct repeat sequence that is 12 bases in length, and similar observations have been reported by Seeger et al. (27) for ground squirrel hepatitis virus and by Will et al. (36) for human hepatitis B virus. The copy flanking the 5′ terminus of the minus strand is designated DR1, and that flanking the 5′ terminus of the plus strand is DR2. These findings were consistent with earlier observations coming from sequence determinations of cloned hepadnaviral DNAs (26) that the location and the separation of these DRs is highly conserved within the hepadnavirus family.

When the 5′ terminus of DHBV virion plus-strand DNA was mapped, we found that a covalently linked oligoribonucleotide of 18 to 19 bases is present on virtually all plus-strand molecules (13). We have inferred that this RNA primes plus-strand DNA synthesis. Sequence analysis has shown that the 3′ 11 to 12 bases of this oligoribonucleotide are identical to the DR2 sequence but that the upstream 6 bases have no homology to the sequence of the corresponding region in viral DNA. Upon further examination, however, it was noted that the oligoribonucleotide shows perfect alignment with DR1 and six bases upstream of it. The sequence of the oligoribonucleotide therefore appears to correspond to the 5′ terminus of a major poly(A)$^+$ RNA species that is thought to be the viral pregenome (5); however, since this RNA is terminally redundant by ca. 270 bases, the oligoribonucleotide sequence also exists at the 5′ end of the 3′-terminal redundancy. These data suggest that the oligoribonucleotide must originate from the 5′ end of the terminal redundancy on the pregenome, and since plus-strand DNA synthesis initiates immediately after the DR2 sequence present upstream of the terminal redundancy, a primer translocation event is implicated. Similar conclusions have been reached for the mammalian hepadnaviruses (27, 36). The conservation of the direct repeats in the hepadnavirus family appears therefore to reflect the mechanism of the plus-strand priming. To further delineate the model of plus-strand primer formation, we have previously shown indirect evidence that a cap structure common to the 5′ end of most eucaryotic mRNAs is present at the 5′ terminus of at least some of the primer molecules, implying their origin from the 5′ terminus of the pregenome (13). Additional

* Corresponding author.

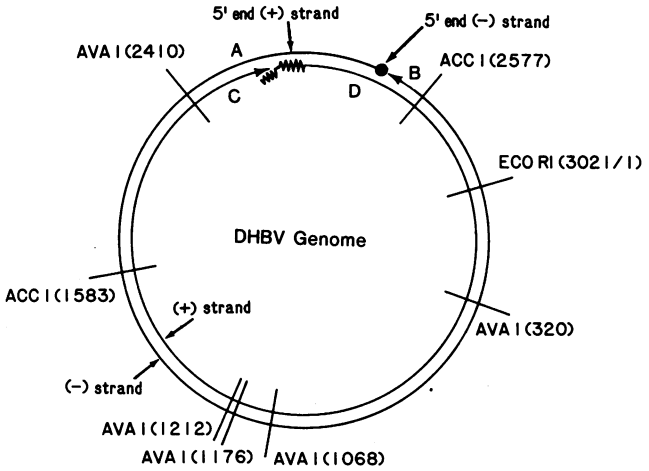


FIG. 1. Characterization of the cohesive overlap region of DHBV. The characterization of the cohesive overlap described in the text was based on analysis of four single-stranded nucleic acid fragments, A, B, C, and D, shown on the circular genome of DHBV (not drawn to scale). Fragment A was a minus-strand fragment extending from the *Ava*I site at nucleotide 2410 to the natural 5′ terminus; this fragment was ^{32}P labeled at the 3′ end for analysis. Fragment B was a minus-strand fragment extending from the *Acc*I site at nucleotide 2577 to the natural 3′ termini; the fragment was ^{32}P labeled at the 5′ end at the *Acc*I site at 2577. Fragment C was a plus-strand fragment extending from the *Ava*I site at nucleotide 2410 to the natural 3′ termini; this fragment was ^{32}P labeled at the 5′ end for analysis. Fragment D was a plus-strand fragment extending from the natural 5′ terminus, including the oligoribonucleotide primer of plus-strand synthesis (13) to the *Acc*I site at nucleotide 2577; the fragment was labeled with ^{32}P at the *Acc*I site at 2577. Restriction endonuclease cleavage sites on the DHBV genome for the enzymes used in this study are taken from Mandart et al. (14). An extra fragment sometimes appears after *Ava*I digestion of virion DNA; this fragment could be explained if a subpopulation of virus in our preparations lacked the *Ava*I site at nucleotide 1068. Symbols: ●, protein at 5′ end of minus strand; ⁓, oligoribonucleotide at 5′ end of plus strand.

evidence will now be presented suggesting that most plus-strand primers originate from this site.

The other aspect of the cohesive overlap structure that had been noted is the presence of a covalently linked protein at or near the 5′ terminus of the minus-strand DNA (8, 22). There has been speculation that the protein, as yet uncharacterized, serves as a primer in minus-strand DNA synthesis (8, 22), but the exact site of the protein-DNA linkage has not been mapped or shown to be the 5′-terminal nucleotide, which would be consistent with the predicted priming role of the protein. A possible linkage to the 5′-terminal nucleotide had only been inferred because the 5′ terminus of the virion minus strand appeared to be blocked to phosphorylation by polynucleotide kinase after treatment with alkaline phosphatase (8). We therefore did experiments to map directly the site of covalent attachment of the protein to the minus strand.

Since plus-strand synthesis apparently initiates at a site about 50 bases away from the 5′ end of the minus-strand DNA template in DHBV, an intramolecular template switch is presumably necessary for DNA elongation to continue around the genome (Fig. 1). We mapped the 3′ end of the virion minus-strand DNA and found, as others did for the mammalian hepadnaviruses (27, 36), that the template switch is probably facilitated by the presence of a terminal redundancy in the minus-strand template. Once plus-strand DNA synthesis is carried across the break in the minus-strand template, it proceeds continuously until either completion or arrest as a result of some step in viral maturation, such as virion release from the cell.

Since the plus-strand primer is not removed during virus maturation (13), it seemed likely that the oligoribonucleotide remains firmly bound to DR2 and thereby blocks plus-strand completion by preventing elongation through DR2. Because the majority of the virion plus-strand DNA in DHBV appeared essentially complete in length (18), we were able to test this hypothesis by mapping the 3′ end(s) of the plus-strand DNA present in the DHBV virion. These and other results described above are summarized in Fig. 2.

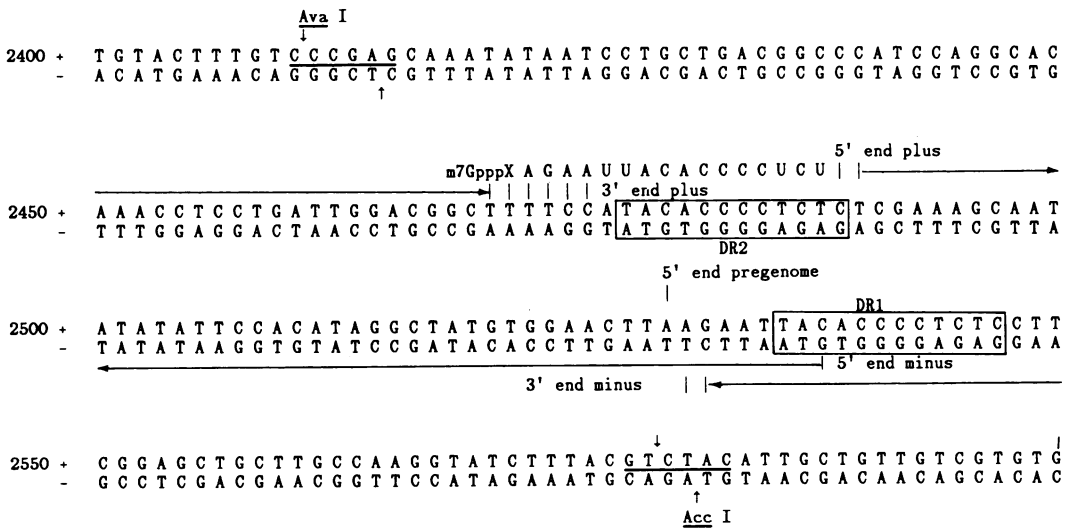


FIG. 2. Sequence of DHBV cohesive overlap region. The mapped DNA initiation and termination sites are summarized and denoted as vertical bars along the sequence of the DHBV cohesive overlap region according to Mandart et al. (14). The 12-base direct repeat sequence is bracketed. The sequence of the capped oligoribonucleotide present at the 5′ terminus of the plus-strand DNA is shown on top of the sequence where it binds. The placement of the 5′ end of the RNA pregenome is different by one to two bases from that of Büscher et al. (5) and is based on the assumption that the plus-strand primer sequence more accurately places this 5′ terminus (see Discussion).

MATERIALS AND METHODS

Virion DNA. DHBV was obtained from the pooled sera of 2- to 3-week-old congenitally infected Pekin ducks hatched from eggs generously provided by Anna O'Connell (Fox Chase Cancer Center). Virus was purified by differential and equilibrium centrifugation as described previously (16), collected from appropriate gradient fractions, and suspended in 20 mM Tris hydrochloride (pH 7.5)–150 mM NaCl. Unless otherwise indicated, virion DNA was isolated by pronase digestion, followed by phenol-chloroform (1:1, vol/vol, saturated with 0.5 M Tris hydrochloride, pH 8.0) extraction and ethanol precipitation (13). Individual labeling protocols described below generally involved the DNA recovered from virus purified from 50 to 100 ml of serum.

Recombinant DNA. The cloning of the entire DHBV genome into *Eco*RI sites of pBR322 and of bacteriophage M13mp7 has been described previously (15, 21). The DHBV sequences in M13mp7 were derived from the same clone that was completely sequenced by Mandart et al. (14). This DNA was originally isolated and cloned from DHBV virions (15). The DHBV DNA cloned into pBR322 was derived from a nuclear DNA preparation enriched for covalently closed circular DNA (21). The sequence through the cohesive overlap region of the latter DNA was determined by Molnar-Kimber et al. (21) and shown to be identical to that of Mandart et al. (14). The position of restriction endonuclease cleavage sites given below refers, in every case, to the first base of the recognition site on the viral plus strand (14).

3′ end labeling of minus-strand DNA fragment A. The objective was to obtain the 5′-terminal *Ava*I fragment of the minus strand of virion DNA that was 3′ end labeled at the *Ava*I site at nucleotide 2410 (Fig. 1). Virus purified through a CsCl density gradient (21) was collected from the gradient fractions by centrifugation. Virion DNA was released by digestion for 1 h at 37°C in 1 mg of proteinase K per ml–0.1% sodium dodecyl sulfate–0.15 M NaCl–0.02 M Tris hydrochloride, pH 7.5. The DNA was collected by phenol-chloroform extraction and ethanol precipitation and then digested with *Ava*I. Labeling at *Ava*I cleavage sites was achieved by carrying out a repair reaction with the Klenow fragment of *Escherichia coli* polymerase I and [α-^{32}P]TTP. A 4- to 5-volume sample of 1 mg of proteinase K or pronase per ml–0.1% sodium dodecyl sulfate–10 mM Tris hydrochloride–10 mM EDTA was then added for 1 h at 37°C, and the DNA was again extracted with phenol-chloroform and collected by ethanol precipitation.

The 931-bp *Ava*I restriction fragment encompassing the cohesive overlap of virion DNA was then purified by preparative gel electrophoresis in low-melting-point agarose and denatured, and the 5′-terminal fragment A of the minus strand labeled at its 3′ end was further purified by urea-polyacrylamide gel electrophoresis (13).

An *Ava*I-*Eco*RI fragment of cloned DHBV DNA that was 3′ end labeled as described above was further digested with *Acc*I to release an *Ava*I-*Acc*I fragment (nucleotides 2410 to 2577). This DNA was then denatured and further purified by urea-polyacrylamide gel electrophoresis. Both the minus-strand terminal *Ava*I fragment of virion DNA and the *Ava*I-*Acc*I fragment of cloned DNA were subjected to DNA sequencing as described below.

5′-end-labeled minus-strand DNA fragment B. The purpose of the following protocol was to obtain minus-strand fragment B (Fig. 1) of virion DNA that had been 5′ end labeled at the *Acc*I site at nucleotide 2577 with T4 polynucleotide kinase. Virion DNA was first digested with *Acc*I, which cleaves at nucleotides 1583 and 2577. The DNA was then extracted twice with phenol-chloroform and subjected to ethanol precipitation. After treatment with 100 U of bacterial alkaline phosphatase for 1 h at 65°C, the mixture was digested with 1 mg of pronase per ml–0.1% sodium dodecyl sulfate and again subjected to phenol-chloroform extraction and ethanol precipitation (13). Phosphorylation was then done with T4 polynucleotide kinase and [γ-^{32}P]ATP (13). The DNA was extracted with phenol-chloroform and collected by ethanol precipitation, and the *Acc*I fragment mapping from nucleotides 1583 to 2577 was purified by preparative gel electrophoresis in low-melting-point agarose (13). This DNA fragment was further fractionated after denaturation by hybrid selection with either M13mp7-DHBV minus-strand or plus-strand DNA (13). The M13-virion DNA hybrids were gel purified and denatured, and radiolabeled DNAs were resolved by urea-polyacrylamide gel electrophoresis (13). The only radiolabeled low-molecular-weight fragments (<100 bases) that should be released by denaturation will map from the natural 3′ end(s) of the minus strand to the *Acc*I site at nucleotide 2577.

A homologous fragment of cloned DHBV DNA 5′ end labeled on the minus strand at the *Acc*I site at nucleotide 2577 and extending from this site to the *Bam*HI site at nucleotide 1658 was cleaved by a modified procedure of Maxam and Gilbert (4, 19, 24, 29) to provide molecular weight markers.

5′-end-labeled plus-strand DNA fragment C. The objective here was to characterize plus-strand fragment C (Fig. 1) of virion DNA that was labeled at its 5′ end at the *Ava*I site at nucleotide 2410. Virion DNA isolation, *Ava*I cleavage, bacterial alkaline phosphatase digestion, and 5′ phosphorylation with [γ-^{32}P]ATP and T4 polynucleotide kinase were done essentially as described above. The labeled *Ava*I fragment extending from nucleotides 2410 to 320 was purified by electrophoresis in a low-melting-point agarose gel and then denatured and subjected to urea-polyacrylamide gel electrophoresis. Alternatively, the agarose gel-purified *Ava*I fragment was denaturated and further fractionated by hybrid selection with either M13mp7 DHBV minus-strand DNA or plus-strand DNA as a negative control (13); products of the annealing reaction were then purified by electrophoresis into and elution from a low-melting-point agarose gel as previously described (13). The gel-purified DNAs were then subjected to electrophoresis in an 8% urea–polyacrylamide gel together with the *Ava*I fragment that had not been subjected to hybrid selection.

A fragment of cloned DHBV DNA mapping from the *Ava*I site at nucleotide 2410 to the *Xba*I site at 2662 and 5′ end labeled on the plus strand at the *Ava*I site was agarose gel purified and cleaved by a modified procedure of Maxam and Gilbert (4, 19, 24, 29) to provide a homologous molecular weight marker.

Modified chemical degradative sequencing. Sequencing reactions were modified (29) from those of Rubin and Schmid (24) and Bencini et al. (4) as follows. Approximately 5,000 to 10,000 cpm of labeled DNA was distributed to each of four microcentrifuge tubes in a volume of 10 μl. About 1.5 μg of carrier DNA was added to each tube, and the following reactions were carried out: A+G, add 1 μl of formic acid (pH 2.0), incubate at 37°C for 40 to 45 min, cool on ice; A+C, add 1 μl of 30% NaOH, incubate at 90°C for 10 to 15 min, cool on ice; C+T, add 15 μl of hydrazine, incubate at room temperature for 15 to 20 min, dehydrate by adding 1 ml of butanol, vortexing, and spinning for 2 min, suspend pellet in 150 μl of H_2O and add 1 ml of butanol, vortex, spin, and dry; T, heat

2 min at 90°C, cool on ice, add 20 μl of 0.04-mg/ml $KMnO_4$, incubate at room temperature for 10 to 15 min, stop by adding 10 μl of allyl alcohol, dehydrate again with 1 ml of butanol as in the C+T reaction. All reaction mixtures were then suspended in 150 μl of 1 M piperidine, incubated at 90°C for 30 min, and cooled on ice. To the A+C reaction was added 150 μl of 70% ethanol. All tubes were then filled with butanol, vortexed, and centrifuged for 2 min. The supernatants were removed, and 150 μl of 1% sodium dodecyl sulfate was added to the pellets. The butanol extractions were then repeated as above. The supernatants were removed, and the pellets were rinsed with 1 ml of ethanol, centrifuged for 2 min, and dried. Generally, about 50 to 75% of the radioactivity was retained to this point. The pellets were suspended in formamide dye mix, heated to 95°C for 5 min, and electrophoresed on an 80-cm 6 or 8% urea–polyacrylamide linear gel.

Periodate oxidation of the 5′ cap on virion plus-strand fragment D. The 5′-terminal *Acc*I fragment D of the virion plus strand containing the oligoribonucleotide was 3′ end labeled at the *Acc*I site at nucleotide 2577 (Fig. 1) by repair synthesis with [α-^{32}P]dCTP as described previously (13), precipitated with ethanol, and suspended in 2.5 mM Tris hydrochloride (pH 7.4)–2.5 mM EDTA containing 5 μM aurintricarboxylic acid. Periodate oxidation was done after a fivefold dilution of the *Acc*I fragment to a final volume of 50 μl with 10 mM sodium acetate (pH 4.6) and 25 mM $NaIO_4$ for 60 min at 0°C in the dark (10). This reaction will convert the *cis*-diol of the cap to the aldehyde form, which can be distinguished from the *cis*-diol by affinity gel electrophoresis (see below). A control reaction was done without $NaIO_4$. The reactions were stopped by the addition of 50 μl of ethylene glycol, and the incubations continued for another 30 min. Before ethanol precipitation, $NaIO_4$ was removed by dialysis against 1 mM EDTA for 3 h at 4°C.

Affinity gel electrophoresis of plus-strand fragment D. Affinity electrophoresis of plus-strand fragment D containing a 5′-terminal oligonucleotide (13) (Fig. 1) was done on boronate-substituted polyacrylamide gels. These gels should specifically retard migration of fragment D species with a cap by interaction with the *cis*-diol group of the cap. *N*-Acryloyl-3-aminophenylboronic acid (APB) was synthesized by the Organic Synthesis Laboratory of the Institute for Cancer Research by the protocol described by Igloi and Kössel (10). An affinity gel was prepared by the addition of 5% (wt/wt) APB (with respect to the total amount of acrylamide plus APB) to the standard urea-polyacrylamide gel components. In particular, for a 6% urea–polyacrylamide gel containing APB, 120 mg of APB was added to 2.28 g of acrylamide, 0.08 g of bisacrylamide, and 16.8 g of urea, and the mixture was dissolved in 0.1 M Tris acetate (pH 9.0)–0.1% ammonium persulfate in a final volume of 40 ml. Polymerization was initiated by addition of 10 μl of *N, N, N′, N′*-tetramethylethylenediamine. The running buffer was 0.1 M Tris acetate (pH 9.0). Recirculation of the running buffer to maintain the alkaline pH was required after samples had migrated into the gel. Control gel electrophoresis was prepared and run as above except no APB was added. Bacteriophage φX174 DNA digested with *Hin*fI and 3′ end labeled at the restriction sites by repair synthesis with three deoxyribonucleotides provided molecular weight markers.

RESULTS

Evidence that protein is linked to the 5′-terminal nucleotide of the minus strand of DHBV. Gerlich and Robinson (8) first demonstrated that the DNA genome of the prototypic hepadnavirus, human hepatitis B virus, is covalently linked to a protein. The attachment site of the protein was mapped to restriction fragments that span the 5′ end of the minus strand. From the observation that the 5′ terminus of the minus strand could not be phosphorylated by polynucleotide kinase after incubation with alkaline phosphatase, it was inferred that the protein might be bound to the terminal nucleotide of the virion minus strand. Previous experiments from our laboratory have shown that protein is covalently bound at or near the 5′ end of the minus strand of DHBV; moreover, nascent minus strands as small as 30 bases in length that were radiolabeled in the endogenous reaction of DNA synthesis complexes were shown to be bound to protein (22). All these observations taken together suggest that protein serves as a primer for minus-strand DNA synthesis.

To test the idea further, we asked whether protein is bound to the exact 5′ end of the minus strand of the virion DNA. Virion DNA of DHBV treated with proteinase K and pronase was digested with *Ava*I, which cleaves at multiple locations including a site about 120 bases downstream from the 5′ end of the minus strand (14) (Fig. 1 and 2). The idea was that the fragment of attached protein surviving this treatment would be sufficient to cause a mobility shift when the viral DNA was analyzed by gel electrophoresis. The proteinase treatment was required to remove most of the attached protein, since fragments of viral DNA with the entire protein attached will not enter a polyacrylamide gel under conditions used in DNA sequencing (22). After 3′ end labeling at the *Ava*I sites, the 5′ terminal fragment of the minus strand (fragment A, Fig. 1) was gel purified and subjected to sequencing by a modified procedure of Maxam and Gilbert (see Materials and Methods). A homologous restriction fragment from a cloned DNA was also 3′ end labeled at this *Ava*I site and then sequenced as a control. The results are shown in Fig. 3. The two sequence ladders aligned perfectly except that the uncleaved virion DNA did not align with the predicted cleavage fragment of the cloned DNA. While the first cleavage of the virion DNA indicated a 5′-terminal guanosine residue on the minus strand at nucleotide 2537 within DR1 (asterisk, Fig. 3), a wide spacing corresponding to a mobility shift of about three bases existed between this band and the uncleaved fragment. (A fuzzy band was also present in the T lane very close to the uncleaved fragment. This band was considered to be an artifact for two reasons. First, no comparable band was present in the C+T lane, and second, a similar band was also present in the T lane of the control DNA despite the absence of a thymidine residue at the terminus of the cloned DNA fragment being sequenced.) These results indicate that minus-strand DNA synthesis is initiated with the G residue at nucleotide 2537 and that a proteinase K- and pronase-resistant oligopeptide fragment is attached to the 5′-terminal G residue, accounting for the anomalous slow migration of the uncleaved fragment. Further evidence for the conclusion that the anomolous migration of the uncleaved fragment is due to a covalently bound peptide was obtained when we analyzed DNA that had been subjected only to proteinase K treatment. In this case, the fragment migrated about one base slower than after incubation with both proteinases (data not shown).

Mapping the 3′ end of the minus strand: evidence that the minus strand is terminally redundant. To map the termination sites of minus-strand DNA synthesis, virion DNA was digested with *Acc*I and 5′ end labeled at the *Acc*I sites, one

of which was located at nucleotide 2577 about 40 bases upstream from the approximate 3′ end of minus-strand DNA (Fig. 1). The labeled restriction fragments were then hybrid selected to separate minus-strand from plus-strand DNA. Since only two *Acc*I sites exist on the virus genome, at nucleotides 1583 and 2577 (14), the only short minus-strand fragment(s) (fragment[s] B, Fig. 1) that is 5′ end labeled must derive from the 3′ end of the minus-strand DNA. When the labeled *Acc*I fragments were analyzed in a denaturing gel, two short minus-strand fragments differing by one base were resolved (Fig. 4, lane 1). Corresponding fragments were not present after hybrid selection of plus-strand DNA (Fig. 4, lane 2). When aligned with the sequence ladder of a cloned viral DNA 5′ end labeled at the *Acc*I site at nucleotide 2577,

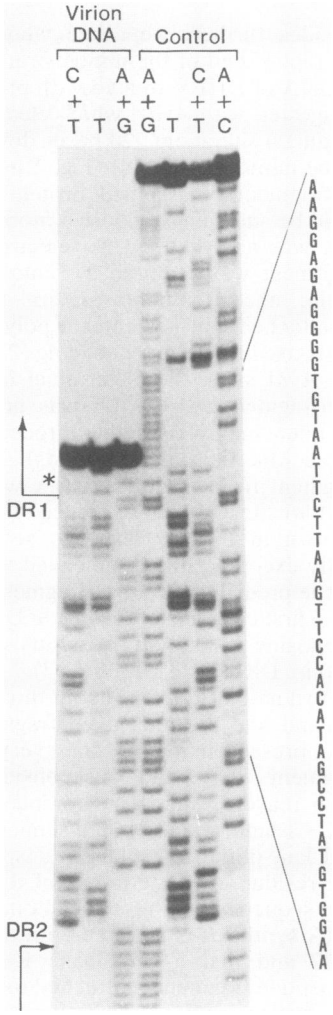


FIG. 3. Evidence that protein is linked to the 5′-terminal nucleotide of the minus strand of DHBV DNA. A virion minus-strand 5′-terminal fragment (fragment A, Fig. 1) 3′ end labeled at the *Ava*I site at nucleotide 2410 was subjected to DNA sequencing by a modified chemical degradative method (see Materials and Methods) (lanes A+G, T, and C+T at the left). As a control size marker, a cloned *Acc*I-*Ava*I minus-strand fragment 3′ end labeled at the same *Ava*I site was also sequenced (lanes A+G, T, C+T, and A+C at the right). The first cleavage fragment in the virion DNA sequence ladder is marked by an asterisk, and the nucleotide sequence through this region is shown at the right. The regions DR1 and DR2 are denoted by arrows and brackets at the left.

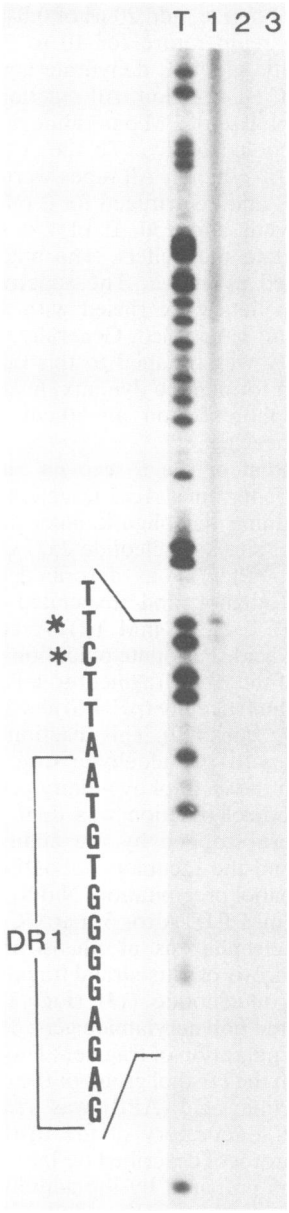


FIG. 4. Evidence that virion minus-strand DNA is terminally redundant. Virion DNA was 5′ end labeled with [γ-^{32}P]ATP and polynucleotide kinase at the *Acc*I site at nucleotide 2577 as described in Materials and Methods. The plus- and minus-strand DNAs were partially purified by hybrid selection with M13mp7-DHBV DNA clones of either minus- or plus-strand polarity. The purified M13mp7-DHBV–virion DNA hybrids were denatured and subjected to electrophoresis on an 80-cm 12.5% urea–polyacrylamide gel. As a size marker, a cloned DHBV DNA was 5′ end labeled at the same *Acc*I site at nucleotide 2577 in the minus strand and cleaved by a modified procedure of Maxam and Gilbert (lane T, for T-specific cleavage). It should be noted that in the marker lane, the designation of a given band in the sequence ladder depends on removal of the corresponding terminal base; thus, fragments in the experimental lanes actually terminate at the next base in the ladder. In addition, the marker fragments have a 3′-terminal phosphate not present on virion DNAs; this phosphate will cause a fragment to migrate about one half base faster than the corresponding fragment lacking the phosphate (32). Lanes: 1, virion minus-strand DNA; 2, virion plus-strand DNA; 3, virion DNA hybrid selected with wild-type M13mp7 as a control.

as exemplified by a T ladder, the two minus-strand fragments were found to terminate at nucleotides 2530 and 2531 (Fig. 2). This result, together with the data from the mapping of the initiation site of the minus-strand DNA, indicates a terminal redundancy seven to eight bases in length upon completion of minus-strand DNA synthesis, compatible with recent data from studies with other hepadnaviruses (27, 36). The same termination sites were also observed by mapping the 3′ ends of the minus-strand DNA radiolabeled by the endogenous polymerase activity of the DNA synthesis complexes isolated from infected livers (data not shown).

Initiation and termination of plus-strand synthesis: evidence that the plus-strand primer is capped and blocks the completion of plus-strand synthesis. We have previously shown that an oligoribonucleotide is present on the 5′ end of all virion plus-strand DNAs (13). This oligoribonucleotide is believed to serve as a primer for plus-strand DNA synthesis. The sequence of this oligoribonucleotide is identical to that of the 5′ end of the presumptive RNA pregenome (see Discussion). Since the pregenome is terminally redundant each RNA contains two copies of this sequence. If the copy present at the 5′ end of the pregenome is the primer, a cap structure that is common to most eucaryotic mRNAs may be present on the oligoribonucleotide. A preliminary experiment with a decapping enzyme, tobacco acid pyrophosphatase, to remove the terminal structure was consistent with this hypothesis (13). However, the enzyme had no effect on about half the primers, possibly because of inhibitors in the reaction or because some of the primers lack the cap. An alternative approach was taken to resolve this issue.

A cap structure contains a *cis*-diol group in its terminal 7-methylated guanosine residue (1, 28), and this *cis*-diol group is known to form a complex with borate. Recently, Igloi and Kössel (10), extending the idea of borate affinity chromatography, developed a method of affinity electrophoresis by copolymerizing a synthetic acrylamide derivative, APB, into an acrylamide gel. The polyacrylamide gel thus formed behaves physically the same as the underivatized gel, except that RNAs presenting the *cis*-diol groups at their termini will interact with the boronate and migrate slower relative to DNA size markers. This technique therefore provides a means to characterize the termini of large RNA fragments without destroying them as in the classical chromatography methods. A plus-strand 5′-terminal *Acc*I fragment of DHBV DNA containing the oligoribonucleotide (fragment D, Fig. 1) was gel purified as described before (13) and subjected to analysis by APB-derivatized polyacrylamide gel electrophoresis. As a control, the same terminal fragment was oxidized by periodate, which is known to convert the *cis*-diol to a dialdehyde (1, 28) and thus would abrogate its ability to complex with boronate (10). Both the untreated and oxidized fragments migrated at the same position in the underivatized gel, while in the presence of 5% APB, the untreated fragment was retarded and migrated 10 bases slower than the oxidized control that remains at the same position relative to marker DNAs (Fig. 5). Since the oligoribonucleotide is covalently linked to the 5′ terminus of the *Acc*I DNA fragment, the result cannot be attributed to a *cis*-diol group present at the 3′ end of the RNA molecule; moreover, there is no evidence that a hypermodified Q base is present in the RNA sequence (10). We therefore concluded that a *cis*-diol group is present at the 5′ end of the oligoribonucleotide, consistent with the presence of a cap structure. The complete removal of the cap structure by periodate oxidation followed by β elimination (1, 28) was also consistent with this conclusion (data not shown).

To estimate the percentage of RNA primers that were capped, it was necessary to consider that ca. 25% of the untreated material (Fig. 5, APB gel) was found to have the same electrophoretic mobility as the oxidized sample in the derivatized gel; however, alkali treatment to remove the primer affected the mobility of less than half of the minor species while apparently shifting the retarded (capped) fragment entirely to the 90-base position (data not shown). This result implied that a significant portion of the minor species lacked the primer RNA and was a contaminant. Thus, at least 75% and probably 90% or more of the plus-strand primers are capped.

The presence of the capped oligoribonucleotide on virtually all virion plus-strand DNAs raised the possibility that the RNA exists in an RNA-DNA hybrid structure with the minus-strand DNA in the DR2 region. If the plus-strand DNA is complete in length in some virions, this region could even have a triple-strand structure. Alternatively, the oligoribonucleotide may cause premature termination of plus-strand DNA synthesis by blocking the elongation through the DR2 region. To analyze the termination site(s) of plus-strand DNA synthesis, the *Ava*I restriction fragment of virion DNA (from nucleotides 2410 to 320) spanning the cohesive overlap region (Fig. 1) was isolated and 5′ end labeled with T4 polynucleotide kinase. Since the end of the restriction fragment at the *Ava*I site at nucleotide 2410 is about 80 bases upstream from the mapped 5′ end of plus-strand DNA (13), denaturation releases small 3′-terminal fragments (fragment[s] C, Fig. 1) of the plus strand. These

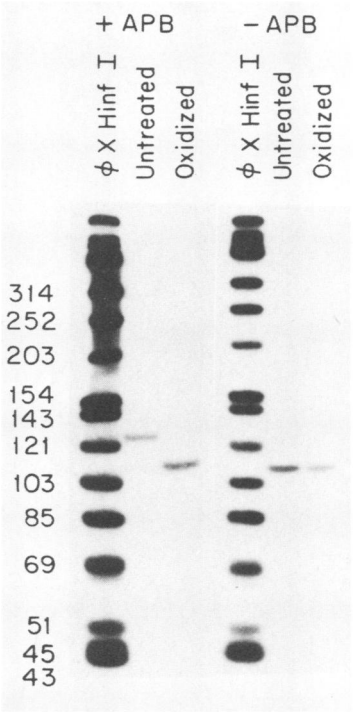
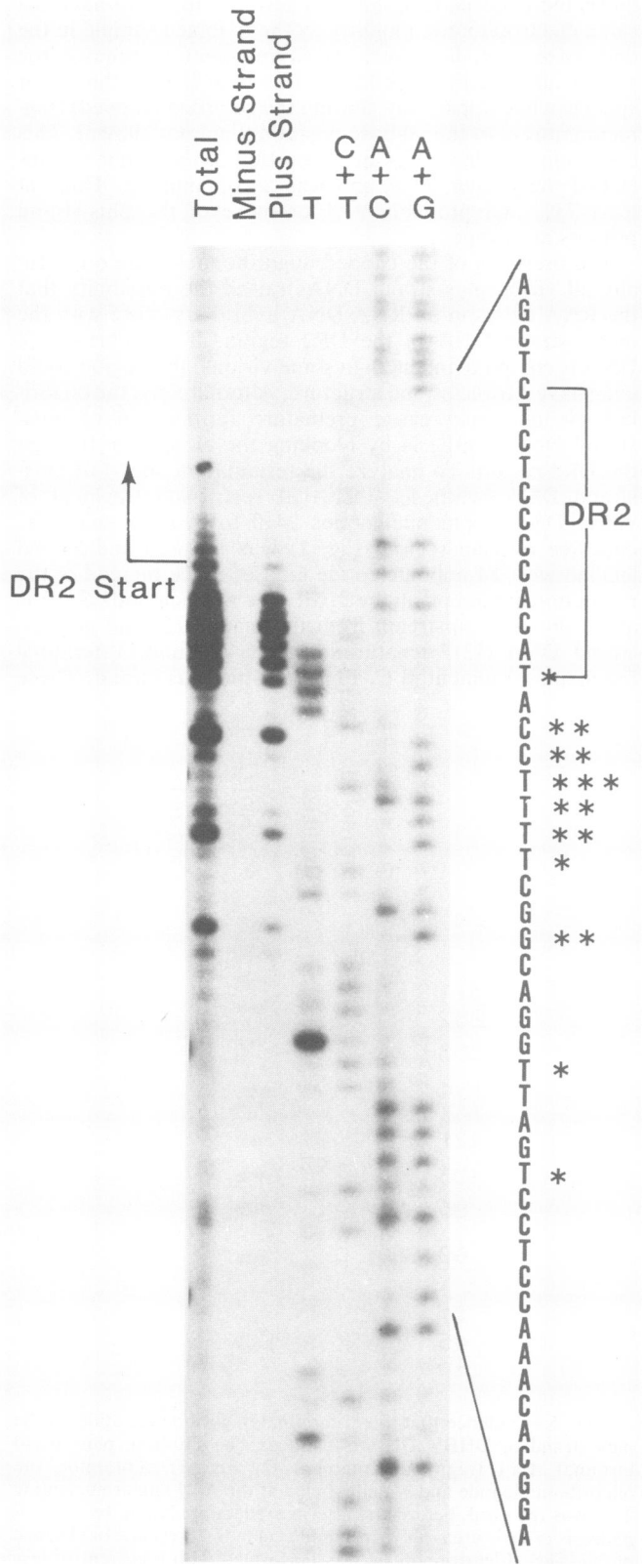


FIG. 5. Evidence that the 5′-terminal oligoribonucleotide on the plus strand of DHBV DNA is capped. The 110-base plus-strand terminal *Acc*I fragment (fragment D, Fig. 1) containing the oligoribonucleotide and 3′ end labeled at the *Acc*I site at nucleotide 2577 was resolved, before and after periodate oxidation, by electrophoresis in a 6% urea–polyacrylamide gel in the presence or absence of 5% APB. Electrophoresis was performed with a current of less than 5 mA overnight with recirculation of the running buffer to maintain an alkaline pH. Bacteriophage φX174 DNA cleaved with *Hin*fI and 3′ end labeled was used as a size marker (φX).



were resolved by urea-polyacrylamide gel electrophoresis and aligned with a sequence ladder derived from a homologous *Ava*I fragment of cloned viral DNA to determine their 3′ ends. Multiple bands were seen in the expected size range (Fig. 6). These were also observed if the DNA was hybrid selected for plus-strand DNA, but not in the minus-strand fraction. Since there was no sequence length heterogeneity observed in this region of DHBV DNA when the *Ava*I fragment of the complementary minus strand isolated from the same virion preparation was sequenced (Fig. 3), we conclude that heterogeneity in the length of the plus-strand fragments must originate from multiple termination sites of plus-strand DNA synthesis. All the termination sites in Fig. 6 are marked by asterisks next to the sequence through this region, and it is readily noted that the predominant sites are located right before the DR2 sequence. Quantitation of these results indicates that <1% of plus strands, if any, extend past the first base of DR2. Since the oligoribonucleotide at the 5′ terminus of plus-strand DNA is complementary to the DR2 sequence of minus-strand DNA, but diverges further upstream, the result suggests that the DR2 region exists in virions as an RNA-DNA hybrid and that the viral polymerase is unable to displace the RNA primer during genome maturation within the immature cores.

DISCUSSION

We completed a characterization of the structure of the cohesive overlap region of the DNA genome of DHBV. By using direct DNA sequencing, both the 5′ and 3′ ends of the two DNA strands were precisely mapped. As discussed below, these data have interesting implications for critical events of hepadnavirus replication, including not only the initiation but also the elongation and termination of viral DNA synthesis.

Minus-strand DNA synthesis. The major poly(A)$^+$ RNA species that is thought to be the template for minus-strand DNA synthesis via reverse transcription has been mapped for several hepadnaviruses (5–7, 23, 36). This RNA begins within the cohesive overlap and, for DHBV, has been shown to be terminally redundant by about 270 bases. The DR1 sequence containing the start site for reverse transcription resides within this long terminal redundancy and is, therefore, reiterated twice in the presumptive pregenomic RNA. Our sequencing data have pinpointed the first nucleotide of minus-strand DNA to the third base within DR1 (at nucleotide 2537) and are in agreement with the results of Molnar-Kimber et al. (21) obtained by the technique of primer extension. Minus-strand DNA synthesis can thus, in theory, initiate within the DR1 sequence of either the 5′ or 3′ ends of

FIG. 6. Plus strand of DHBV DNA terminates before DR2. Virion DNA was cleaved with *Ava*I and 5′ end labeled with polynucleotide kinase and [γ-^{32}P]ATP. The DNA fragment that spans the cohesive overlap was gel purified and hybrid selected to separate and isolate either minus- or plus-strand DNA. About 1,000 cpm of each hybrid-selected DNA was run together with the unselected virion DNA (lane designated as total). A sequence ladder, shown to the right, was generated by a modified chemical degradative method (see Materials and Methods) with a cloned *Ava*I-*Xba*I fragment 5′ end labeled at the same *Ava*I site at nucleotide 2410. The bases at which plus-strand DNAs terminate are marked by asterisks alongside the sequence spanning the region. The number of asterisks approximates the relative intensity of each stop site as seen by autoradiography. The DR2 region is denoted by arrows and brackets. The comments made in the legend to Fig. 4 should be kept in mind when comparing the sequencing ladder with the fragments of virion DNA.

the RNA template (5). If initiation occurs at the 5′ end, a template switch event must be implicated so that nascent minus-strand DNA can be transferred to the 3′ end of the pregenome to carry out further strand elongation. There is no evidence at present to rule out this possibility. However, when the 3′ end of the virion minus-strand DNA was mapped, a small terminal redundancy was noted (Fig. 4) and the 3′ end(s) coincided with the 5′ end of the pregenome template as approximately mapped by Büscher et al. (5) to nucleotide 2530/2531 (±1). If the plus-strand primer is assumed to define the exact 5′ end of the pregenome, then the pregenome actually begins at nucleotide 2529 and minus-strand elongation stops just before the penultimate base of the RNA, which is likely to be a 2′-O-methylated adenosine residue (13). (In this regard, it is important to note that the mapping of the 5′ end of the pregenome by primer extension with reverse transcriptase [5] might have failed to reveal the penultimate nucleotide for the same reason that viral minus-strand synthesis stops before this base.) This result supports the hypothesis that the 5′ end of the pregenome RNA template is preserved until the completion of minus-strand DNA synthesis. The data favor a model in which minus-strand DNA synthesis initiates within that copy of DR1 located in the 3′-terminal redundancy of the pregenome.

One useful observation from our data is that a proteinase-resistant oligopeptide fragment is covalently linked to a G residue at nucleotide 2537 within DR1; that is, at the exact 5′ end of minus-strand DNA. Since a protein has been shown previously to be covalently linked to the minus-strand DNA at or close to its terminus (8, 22), the result provides further support for the idea that protein is the primer for minus-strand DNA synthesis. The interesting question is now to ask how the protein primer recognizes and initiates specifically at the start site we mapped and what the identity of this protein is.

Plus-strand DNA synthesis. We have previously demonstrated that an oligoribonucleotide is linked to the 5′ ends of almost all the virion plus-strand DNAs examined (13). The RNA is thought to be the primer for plus-strand DNA synthesis. Further analysis has shown that the oligoribonucleotide sequence is identical to that of the 5′-terminal 17 to 18 bases of the pregenome RNA as determined by Büscher et al. (5), with the exception that these investigators did not detect the penultimate base of the primer, which is probably part of the pregenome, as well. The results have been confirmed in the mammalian hepadnaviruses (27, 36), which reveal a remarkable similarity in the size and the structure of the oligoribonucleotide. In this report, we showed that the *cis*-diol group characteristic of the cap structure present in eucaryotic mRNAs can be demonstrated in over 75% of the oligoribonucleotides studied. The results argue for a common origin of all the plus-strand primers from the 5′ end of the pregenome, as opposed to the 5′ end of the 3′-terminal redundancy, and imply a translocation mechanism of this primer that is probably similar to the priming of transcription in influenza virus (33) and coronavirus (2, 3).

After the oligoribonucleotide primer is translocated to the DR2 sequence on minus-strand DNA, which provides a complementary binding site, plus-strand DNA synthesis initiates immediately adjacent to the DR2 sequence and continues to the 5′ end of the template 48 to 49 bases away (13). A structural stop in plus-strand DNA elongation is proposed, and a template switch event is needed. Our data show that a terminal redundancy is present in the minus-strand DNA template of DHBV and are compatible with the hypothesis that a template switch is mediated by this short redundant sequence (27). However, we have been unable to demonstrate a plus-strand DNA fragment indicative of a strutural stop (20, 34) in strand elongation at the 3′ end of the minus-strand template in the replicative complexes from the infected livers (unpublished data), which probably indicates that the template switch event is not a kinetic block in plus-strand DNA synthesis in vivo.

Agarose gel electrophoretic analysis has suggested that most of the plus-strand DNAs in the DHBV virions are complete in length (18). When the 3′ ends of these DNAs were mapped, however, most of the ends terminated immediately upstream of the DR2 sequence, which is where the oligoribonucleotide primer binds. Since the primer is found attached to most if not all of the plus strands in a preparation of virion DNA, the results suggest that DR2 exists as an RNA-DNA hybrid in the mature viral genome and that DNA polymerase inside the virion is unable to displace the RNA primer from DR2. It is not known whether an inherent property of the viral polymerase or a structural feature of DNA synthesis complexes that is relieved later upon initiation of infection would account for this block. Since the covalently closed circular viral DNA found in infected liver lacks ribonucleotides (D. Petcu, unpublished observations), either a host or viral polymerase must eventually copy the DR2 sequence. It may be possible to use polymerase-specific inhibitors of DNA synthesis to distinguish between these two possibilities.

ACKNOWLEDGMENTS

We are grateful to A. Hampton, J. Summers, and J. Taylor for helpful suggestions during the course of this work and to L. Coates for technical assistance. We thank C. Bergman (Organic Synthesis Laboratory, Fox Chase Cancer Center) for preparation of APB. We acknowledge A. Capriotti and M. Piatek for assistance in the preparation of this manuscript.

This work 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. D. Petcu is a Medical Scientist Training Program trainee of the National Institutes of Health in the Department of Genetics, University of Pennsylvania.

LITERATURE CITED

1. **Banerjee, A. K.** 1980. 5′-terminal cap structure in eucaryotic messenger ribonucleic acids. Microbiol. Rev. **44**:175–205.
2. **Baric, R. S., C.-K. Shieh, S. A. Stohlman, and M. M. C. Lai.** 1987. Analysis of intracellular small RNAs of mouse hepatitis virus: evidence for discontinuous transcription. Virus Res. **6**: 342–354.
3. **Baric, R. S., S. A. Stohlman, M. K. Razavi, and M. M. C. Lai.** 1985. Characterization of leader-related small RNAs in coronavirus-infected cells: further evidence for leader-primed mechanism of transcription. Virus Res. **3**:19–33.
4. **Bencini, D. A., G. A. O'Donovan, and J. R. Wild.** 1984. Rapid chemical degradation sequencing. Biotechniques **2**:4–5.
5. **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.
6. **Cattaneo, R., H. Will, and H. Schaller.** 1984. Hepatitis B virus transcription in the infected liver. EMBO J. **3**:2191–2196.
7. **Enders, G. H., D. Ganem, and H. E. Varmus.** 1985. Mapping the major transcripts of ground squirrel hepatitis virus: the presumptive template for reverse transcriptase is terminally redundant. Cell **42**:297–308.
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. **Hruska, J. F., D. A. Clayton, J. L. R. Rubenstein, and W. S. Robinson.** 1977. Structure of hepatitis B Dane particle DNA

polymerase reaction. J. Virol. **21**:666–672.
10. **Igloi, G. L., and H. Kössel.** 1985. Affinity electrophoresis for monitoring terminal phosphorylation and the presence of quenosine in RNA. Application of polyacrylamide containing a covalently bound boronic acid. Nucleic Acids Res. **13**:6881–6898.
11. **Kaplan, P. M., R. L. Greenman, J. L. Gerin, R. H. Purcell, and W. S. Robinson.** 1973. DNA polymerase associated with human hepatitis B antigen. J. Virol. **12**:995–1005.
12. **Landers, T. A., H. B. Greenberg, and W. S. Robinson.** 1977. Structure of hepatitis B Dane particle DNA and nature of the endogenous DNA polymerase reaction. J. Virol. **23**:368–376.
13. **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.
14. **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.
15. **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.
16. **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.
17. **Mason, W. S., J. M. Taylor, and R. Hull.** 1987. Retroid virus genome replication. Adv. Virus Res. **32**:35–96.
18. **Mason, W. S., J. M. Taylor, G. Seal, and J. Summers.** 1981. An HBV-like virus of domestic ducks, p. 107–116. *In* W. Szmuness, H. J. Alter, and J. E. Maynard (ed.), Viral hepatitis. Franklin Institute Press, Philadelphia.
19. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
20. **Mitra, S. W., S. Goff, E. Gilboa, and D. Baltimore.** 1979. Synthesis of 600-nucleotide-long plus strand DNA by virions of Moloney leukemia virus. Proc. Natl. Acad. Sci. USA **76**: 4355–4359.
21. **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.
22. **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.
23. **Möröy, T., J. Etiemble, C. Trépo. P. Tiollais, and M.-A. Buendia.** 1985. Transcription of woodchuck hepatitis virus in the chronically infected liver. EMBO J. **4**:1507–1514.
24. **Rubin, C. M., and C. W. Schmid.** 1980. Pyrimidine-specific chemical reactions useful for DNA sequencing. Nucleic Acids Res. **8**:4613–4619.
25. **Sattler, F., and W. S. Robinson.** 1979. Hepatitis B viral DNA molecules have a cohesive overlap. J. Virol. **32**:226–233.
26. **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.
27. **Seeger, C., D. Ganem, and H. E. Varmus.** 1986. Biochemical and genetic evidence for the hepatitis B virus replication strategy. Science **232**:477–484.
28. **Shatkin, A. J.** 1976. Capping of eukaryotic mRNAs. Cell **9**:645–653.
29. **Shlomchick, M. J., D. A. Nemazee, V. L. Sato, J. VanSnick, D. A. Carson, and M. G. Weigert.** 1986. Variable region sequences of murine IgM anti-IgG monoclonal autoantibodies (rheumatoid factors)—a structural explanation for the high frequency of IgM anti-IgG B cells. J. Exp. Med. **164**:407–427.
30. **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.
31. **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. USA **72**:4597–4601.
32. **Tapper, D. P., and D. A. Clayton.** 1981. Altered mobility of polynucleotides in high resolution polyacrylamide gels due to removal of terminal phosphates. Nucleic Acids Res. **9**:6787–6794.
33. **Ulmanen, I., B. A. Broni, and R. M. Krug.** 1981. Role of two of the influenza virus core P proteins in recognizing cap 1 structures (m^7GpppNm) on RNAs and in initiating viral RNA transcription. Proc. Natl. Acad. Sci. USA **78**:7355–7359.
34. **Varmus, H. E., S. Heasley, H.-J. Kung, H. Opperman, V. C. Smith, J. M. Bishop, and P. R. Shank.** 1978. Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. J. Mol. Biol. **120**:55–82.
35. **Weiss, R., N. Teich, H. Varmus, and J. Coffin.** 1983. RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Will, H., W. Reiser, T. Weimer, E. Pfaff, M. Büscher, R. Sprengel, R. Cattaneo, and H. Schaller.** 1987. Replication strategy of human hepatitis B virus. J. Virol. **61**:904–911.